

NCKU 112.1
Introduction to Analysis

Eric Liu

CONTENTS

CHAPTER 0 BASIC SET THEORY PAGE 2

0.1 ZF	2
0.2 Ordinal	10
0.3 Axiom of Choice and its equivalents	16
0.4 Cardinality and Cardinal	26
0.5 Ordinal Arithmetic	33
0.6 Cardinal Arithmetic	34
0.7 Set Theoretic Limit	35

CHAPTER 1 THE REAL AND COMPLEX NUMBER SYSTEM PAGE 37

1.1 Introduction	37
1.2 Ordered Fields	41
1.3 Real Numbers Field	44
1.4 Irrational Power	51
1.5 Euclidean Space	59
1.6 Complex Numbers	62
1.7 Existence of Real Numbers - Dedekind Cut*	70
1.8 Existence of Real Numbers - Decimal*	71
1.9 Uniqueness of Real Numbers*	72
1.10 Exercises	73
1.11 Complicated Exercises	75

CHAPTER 2 BASIC TOPOLOGY PAGE 76

2.1 Different but Equivalent Ways to Define A Topology	76
--	----

2.2 Closure and Interior Operator	81
2.3 Different Types of Connected	85
2.4 Continuous Function	92
2.5 Homeomorphism	98
2.6 Three Non-Equivalent Compactness and Separable	99

CHAPTER 3 IMPORTANT NOTIONS IN METRIC SPACE PAGE 102_____

3.1 Basic Fact of Sequence	102
3.2 Closed	105
3.3 Open	112
3.4 Connected and Convex	118
3.5 Bounded	120
3.6 Compactness	122
3.7 Metric Space: Some Properties Implied by Compact	128
3.8 Complete	130
3.9 Euclidean Metric	131
3.10 Cauchy Sequence	136
3.11 Equivalent Definition of Continuity and Two Kinds of Discontinuity for $\mathbb{R} \rightarrow X$	141

CHAPTER 4 NUMERICAL SEQUENCE AND SERIES PAGE 149_____

4.1 Special Real Sequence and Series	149
4.2 Sequence in Normed Space	154
4.3 Comparison Test, Geometric Series and p-Series	155
4.4 Sub-Sequential Limit	160
4.5 Euler Number	168
4.6 Tests for Convergence of Series	172
4.7 Cauchy Product and Mertens' Theorem	177
4.8 Reimann Rearrangement Theorem	181

CHAPTER 5 CALCULUS AND RELEVANT NOTIONS PAGE 183

5.1	Derivative	183
5.2	Discontinuity and Convex	186
5.3	Uniformly Continuous	191
5.4	Tests for Uniform Continuity	196
5.5	Absolute Continuous and Lipschitz Continuous	202

CHAPTER 6 RUDIN: DIFFERENTIATION PAGE 208

6.1	Derivative	208
6.2	Mean Value Theorem	211
6.3	The Continuity of Derivative	213
6.4	L'Hospital Rule	214
6.5	Taylor's Theorem	216

CHAPTER 7 RUDIN: RIEMANN-STIELTJES INTEGRAL PAGE 219

7.1	Definition and Existence of The Integral	219
-----	--	-----

CHAPTER 8 VECTOR ANALYSIS PAGE 228

8.1	Important Inequalities	228
8.2	Definition of Differential-able on Euclidean Space	231
8.3	Example Exercises for Differentiable	233
8.4	Directional Derivative	236
8.5	Gradient	237
8.6	Integral	238

CHAPTER 9 VE 2 PAGE 240

9.1	Line Integral	240
9.2	Fundamental Theorem of Line Integral	248
9.3	Green's Theorem	253

CHAPTER 10	VE 3	PAGE 261
-------------------	-------------	-----------------

10.1Flux and Surface Integral	261
10.2Gauss Divergence Theorem	270

CHAPTER 11	ADVANCED CALCULUS HW	PAGE 272
-------------------	-----------------------------	-----------------

11.1HW1	272
11.2HW2	280
11.3HW3	291
11.4HW4	309
11.5HW5	327
11.6HW6	332
11.7HW7	339

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:

$$x \in A$$

and

$$A = B$$

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)

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

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

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

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)

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

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)

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

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)

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

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)

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

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)

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

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

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

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

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

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

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

Before reading the proof of the following theorem, notice that the set $\{y \in \bigcup x : \forall z(z \in x \longrightarrow y \in z)\}$ 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.

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

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

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

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

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

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

The theorem states:

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

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

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

$P_1(Y)$ 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)

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

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)

$$\forall x, \neg x \in x$$

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

$$\cdots \in x_3 \in x_2 \in x_1$$

Formally, we say

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$\langle S, < \rangle$ 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:

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

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

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

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

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

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

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, WOLG, $y = a$, then $x \in a = y$.

We now prove the transitive part.

Let $x, y, z \in a \cup \{a\}$ and $x \in y$ and $y \in z$. By axiom of foundation, it is impossible more than one of x, y, z equals to a . If $x = a$, then we have $y \in a$, so we see

$\cdots a = x \in y \in a \in y \in a$, which implies it is impossible $x = a$. Similarly, we can deduce $y \neq a$. So far, we have proven $x \in a$ and $y \in a$. If $z \in a$, then because a is an ordinal, our proof 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

$$a \in b \iff a \subset b$$

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

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

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

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

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

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

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

$$0, 1, 2, 3, \dots; \omega, \omega + 1, \omega + 2, \dots$$

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

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

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

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

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

$$\forall x \in X, x \leq t$$

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

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

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

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

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

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

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

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

$$\{\{a\}, \{a, b\}\}$$

as (a, b) and we interpret the set

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

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.

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

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

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

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

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

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

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

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

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

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

$$R := \bigcup \{T : T \subseteq' A \text{ and } T \subseteq' B\}$$

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

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

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

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

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

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

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

$$g \leq h \text{ if } g \subseteq h$$

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

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

and we define a partial order \leq for P by

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

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

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

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 $P(f)$

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

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

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

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

$$A \leq_c B \text{ or write } B \geq_c A$$

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

$$A =_c B$$

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

$$A \leq_c B \text{ and } B \leq_c A \iff A =_c B$$

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

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

Also, Define

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

Define

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

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

$$\begin{aligned} x \in A \setminus C &\implies x \in A \setminus (A \setminus g(B)) \\ &\iff x \in A \text{ and } x \notin A \setminus g(B) \\ &\iff x \in A \text{ and } (x \notin A \text{ or } x \in g(B)) \\ &\iff (x \in A \text{ and } x \notin A) \text{ or } (x \in A \text{ and } x \in g(B)) \\ &\iff x \in A \text{ and } x \in g(B) \\ &\iff x \in g(B) \text{ (done)} \end{aligned}$$

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

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

So we just prove

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

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

$$\begin{aligned} x &\notin g^{-1}(A \setminus C) \\ &\iff g(x) \notin A \setminus C \\ &\iff g(x) \in C \\ &\iff \exists n \in \mathbb{N}, g(x) \in C_n \\ &\iff \exists n \in \mathbb{N}, \exists u \in C_{n-1}, g(x) = g(f(u)) \\ &\iff \exists n \in \mathbb{N}, \exists u \in C_{n-1}, x = f(u) \\ &\iff x \in f(C) \text{ (done)} \end{aligned}$$

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

$$A \leq_c B \text{ or } B \leq_c A$$

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

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

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

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

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 $A \not\sim_c B$, then we write

$$A >_c B \text{ or write } B <_c A$$

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

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

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

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

$$|X| = \min\{\sigma \in \text{Ord} : \text{and } \sigma =_c X\}$$

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

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

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

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

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

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

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 $|g(n)| = n = \left| \{x \in \mathbb{N} : \frac{(n-1)n}{2} < x \leq \frac{n(n+1)}{2}\} \right|$, 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

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

■

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

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

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

$$A \leq B$$

if

$$A \subseteq B$$

The class of sets form a partial order.

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

$$\sigma = \{A_i : i \in I\}$$

The collection of lower bounds of σ is exactly

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

and

$$A \text{ is an upper bound of } \sigma \iff \bigcup \sigma \subseteq A$$

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

$$\inf \sigma = \bigcap \sigma \text{ and } \sup \sigma = \bigcup \sigma$$

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

$\inf \sigma$ is the greatest set smaller than all sets in σ

$\sup \sigma$ is the smallest set greater than all sets in σ

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

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

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

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

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

Theorem 0.7.7. (Meaning Behind Lim-Sup)

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

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

Corollary 0.7.8. (Verification of Theorem 0.7.6) We have

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

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

And

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

Proof. Suppose

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

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

Suppose

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

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)

$$\inf\{A_n\} \subseteq \liminf_{n \rightarrow \infty} A_n \subseteq \limsup_{n \rightarrow \infty} A_n \subseteq \sup\{A_n\}$$

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

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

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

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

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

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

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

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

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

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 .

$$\begin{aligned} q - p &= \frac{2 - p^2}{a} \\ q^2 - 2 &= \frac{p^2 - 2}{b} \end{aligned}$$

Now we solve for a, b

$$\begin{aligned} q - p &= \frac{2 - p^2}{a} \text{ and } q^2 - 2 = \frac{p^2 - 2}{b} \implies q = \frac{2 - p^2}{a} + p \text{ and } q^2 = \frac{p^2 - 2}{b} + 2 \\ &\iff \left[\frac{2 - p^2}{a} + p \right]^2 = \frac{p^2 - 2}{b} + 2 \\ &\iff \frac{(2 - p^2)^2}{a^2} + \frac{2p(2 - p^2)}{a} + p^2 = \frac{p^2 - 2}{b} + 2 \\ &\iff (p^2 - 2)^2 \left(\frac{1}{a^2} \right) + (p^2 - 2) \left(-\frac{2p}{a} + 1 - \frac{1}{b} \right) = 0 \\ &\iff \frac{p^2 - 2}{a^2} - \frac{2p}{a} + 1 - \frac{1}{b} = 0 \text{ (because } p^2 - 2 \neq 0) \\ &\iff \frac{p^2 - 2 - 2ap + a^2}{a^2} = \frac{1}{b} \\ &\iff b = \frac{a^2}{p^2 - 2 - 2ap + a^2} = \frac{a^2}{(p - a)^2 - 2} \end{aligned}$$

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

$$\forall x \in S, x \sim x$$

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

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

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

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

$$\exists a \in S, \forall b \in E, a \geq b$$

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

$$\forall b \in E, c \leq b$$

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

$$\sup E := \min A$$

and

$$\inf E := \max B$$

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

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

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

$$E \subseteq S \text{ and } E \neq \emptyset \implies \sup E \text{ exists}$$

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

$$y < z \implies x + y < x + z$$

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

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

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

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

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

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

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

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

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

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

$$x^{p+q} = x^p x^q$$

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

$$x^{pq} = (x^p)^q = (x^q)^p$$

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

$$x^p y^p = (xy)^p$$

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

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

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 impor-

tant 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 $m \geq x - k \longrightarrow x^n - m^n \leq x^n - y$, so we can draw the contradiction $x - k$ is an upper bound of E .

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

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

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

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

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

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

Proof. Observe

$$\begin{aligned} x &= (a^m)^{\frac{1}{n}} \\ \iff x^n &= a^m \\ \iff (x^n)^q &= (a^m)^q \\ \iff x^{nq} &= a^{mq} = a^{np} \text{ (because } n, m, q \in \mathbb{Z}) \\ \iff (x^q)^n &= (a^p)^n \text{ (again, because } n, p, q \in \mathbb{Z}) \\ \iff x^q &= a^p \text{ (by Theorem 1.3.6)} \\ \iff x &= (a^p)^{\frac{1}{q}} \text{ (by Theroem 1.3.6)} \\ \iff (a^m)^{\frac{1}{n}} &= x = (a^p)^{\frac{1}{q}} \end{aligned}$$

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

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

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

$$\begin{aligned} a^x < a^y &\implies a^{\frac{q}{p}} < a^{\frac{n}{m}} \\ &\implies a^q = (a^{\frac{q}{p}})^p < (a^{\frac{n}{m}})^p \text{ (by Theorem 0.2.10)} \\ &\implies a^{mq} = (a^q)^m < ((a^{\frac{n}{m}})^p)^m = ((a^{\frac{n}{m}})^m)^p = (a^n)^p = a^{np} \\ &\implies a > 1 \text{ (by Theorem 0.2.9)} \end{aligned}$$

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

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

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

$$a^{r+s} = a^r a^s$$

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

$$\begin{aligned}(a^{r+s})^{nq} &= (a^{\frac{np+mq}{nq}})^{nq} \\ &= a^{np+mq}\end{aligned}$$

Then observe

$$\begin{aligned}(a^r a^s)^{nq} &= (a^{\frac{p}{q}} a^{\frac{m}{n}})^{nq} \\ &= (a^{\frac{p}{q}})^{nq} (a^{\frac{m}{n}})^{nq} \text{ (by Theorem 0.2.14)} \\ &= ((a^{\frac{p}{q}})^q)^n ((a^{\frac{m}{n}})^n)^q \\ &= (a^p)^n (a^m)^q \\ &= a^{np+mq} = (a^{r+s})^{nq} \text{ (Theorem 0.2.12 and Theorem 0.2.13)}\end{aligned}$$

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

$$(a^r)^s = a^{rs}$$

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

$$\begin{aligned}(a^{rs})^{nq} &= (a^{\frac{mp}{nq}})^{nq} \\ &= a^{mp}\end{aligned}$$

Then observe

$$\begin{aligned}((a^r)^s)^{nq} &= (((a^{\frac{p}{q}})^{\frac{m}{n}})^n)^q \\ &= ((a^{\frac{p}{q}})^m)^q \\ &= (a^{\frac{p}{q}})^{mq} \\ &= ((a^{\frac{p}{q}})^q)^m \\ &= (a^p)^m = a^{mp} = (a^{rs})^{nq}\end{aligned}$$

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

$$a^r b^r = (ab)^r$$

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

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

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 an lower bound of $B(x)$. Notice that there exists rational between $x < q$, so b^q can not be the greatest lower bound of $B(x)$. Then we can deduce $\inf B(y) \leq b^q < \inf B(x)$.

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

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

$$\sup SU = \sup S \sup U$$

and have

$$\inf SU = \inf S \inf U$$

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

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

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

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

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

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

$$a^{r+s} = a^r a^s$$

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

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

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

$$\sup S \sup U = \sup US_\phi$$

and have

$$\inf S \inf U = \inf US_\phi$$

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

$$b^z c^z = (bc)^z$$

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

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

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

$$(a^r)^s = a^{rs}$$

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

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

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

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

So, we have

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

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

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

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 .

$$\begin{aligned}(c, d) &:= \{\{c\}, \{c, d\}\} \\ (b, c, d) &:= (b, (c, d)) = \{\{b\}, \{b, (c, d)\}\} = \{\{b\}, \{b, \{\{c\}, \{c, d\}\}\}\} \\ (a, b, c, d) &:= (a, (b, c, d)) = \dots\end{aligned}$$

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

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

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

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

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

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

and equipped with 2-norm

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

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

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

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

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

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

$f(t)$ 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

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

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

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

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

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

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

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

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

■

The fact that the field in Theorem 1.6.1 is isomorphic to the customary \mathbb{C} can be easily verified by checking $(a, b) \mapsto a + bi$ is field isomorphism. We now define \mathbb{C} in this way. Notice that this definition is an abuse of notation, since in this way, $\mathbb{R} \not\subset \mathbb{C}$. 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

$$\begin{aligned}\overline{(\bar{z})} &= z \\ \overline{z + w} &= \bar{z} + \bar{w} \\ \overline{zw} &= (\bar{z})(\bar{w}) \\ \overline{z^n} &= (\bar{z})^n \\ z + \bar{z} &= 2\text{Re}(z) \text{ and } z - \bar{z} = 2i\text{Im}(z) \\ |\bar{z}| &= |z| \\ |zw| &= |z| |w| \\ |z^n| &= |z|^n\end{aligned}$$

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

$$\begin{aligned}\overline{(\bar{z})(\bar{w})} &= \overline{(z_1 - z_2i)(w_1 - w_2i)} \\ &= \overline{z_1w_1 - z_2w_2 - (z_2w_1 + z_1w_2)i} \\ &= z_1w_1 - z_2w_2 + (z_2w_1 + z_1w_2)i \\ &= zw\end{aligned}$$

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

$$\begin{aligned}|zw| &= |z_1w_1 - z_2w_2 + (z_2w_1 + z_1w_2)i| \\ &= [(z_1w_1 - z_2w_2)^2 + (z_2w_1 + z_1w_2)^2]^{\frac{1}{2}} \\ &= [(z_1w_1)^2 + (z_2w_2)^2 - 2z_1z_2w_1w_2 + (z_2w_1)^2 + (z_1w_2)^2 + 2z_1z_2w_1w_2]^{\frac{1}{2}} \\ &= [z_1^2w_1^2 + z_2^2w_1^2 + z_1^2w_2^2 + z_2^2w_2^2]^{\frac{1}{2}} \\ &= [(z_1^2 + z_2^2)(w_1^2 + w_2^2)]^{\frac{1}{2}} = (z_1^2 + z_2^2)^{\frac{1}{2}}(w_1^2 + w_2^2)^{\frac{1}{2}} = |z| |w|\end{aligned}$$

■

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

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

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

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

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

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

$$\begin{aligned} |\mathbf{x} + \mathbf{y}|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) \\ &= \mathbf{x} \cdot (\mathbf{x} + \mathbf{y}) + \mathbf{y} \cdot (\mathbf{x} + \mathbf{y}) \\ &= |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2(\mathbf{x} \cdot \mathbf{y}) \end{aligned}$$

and observe

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

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

Observe

$$\begin{aligned} |\lambda \mathbf{x}| &= \left(\sum (\lambda x_i)^2 \right)^{\frac{1}{2}} \\ &= (\lambda^2 \sum x_i^2)^{\frac{1}{2}} \\ &= |\lambda| \left(\sum x_i^2 \right)^{\frac{1}{2}} \\ &= |\lambda| |\mathbf{x}| \end{aligned}$$

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:

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

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

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

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

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

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

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

■

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

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

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

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

■

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

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

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

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

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

So we have

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

This **CaC** to Lemma 1.6.8.

Observe

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

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

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

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

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

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

$$\begin{aligned}
\sum |Bv_j - Cw_j|^2 &= \sum (Bv_j - Cw_j) \overline{(Bv_j - Cw_j)} \\
&= \sum (Bv_j - Cw_j) (B\overline{v_j} - \overline{C} \overline{w_j}) \\
&= \sum B^2 |v_j|^2 - BC\overline{v_j}w_j - B\overline{C}v_j\overline{w_j} + C\overline{C} |w_j|^2 \\
&= B^2A - BC\overline{C} - B\overline{C}C + C\overline{C}B \\
&= B^2A - BC\overline{C} \\
&= B(AB - |C|^2)
\end{aligned}$$

Because $B > 0$, then we can deduce

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

So we can deduce

$$\sum |v_j|^2 \sum |w_j|^2 \geq \left| \sum v_j \overline{w_j} \right|^2$$

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

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

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

$$\begin{aligned}
\|B\mathbf{v} - C\mathbf{w}\|^2 &= \langle B\mathbf{v} - C\mathbf{w}, B\mathbf{v} - C\mathbf{w} \rangle \\
&= \langle B\mathbf{v}, B\mathbf{v} \rangle - \langle C\mathbf{w}, B\mathbf{v} \rangle - \langle B\mathbf{v}, C\mathbf{w} \rangle + \langle C\mathbf{w}, C\mathbf{w} \rangle \\
&= B^2\langle \mathbf{v}, \mathbf{v} \rangle - CB\langle \mathbf{w}, \mathbf{v} \rangle - B\overline{C}\langle \mathbf{v}, \mathbf{w} \rangle + C\overline{C}\langle \mathbf{w}, \mathbf{w} \rangle \\
&= B^2A - CB\overline{C} - B\overline{C}C + C\overline{C}B \\
&= B(AB - |C|^2)
\end{aligned}$$

Because $B > 0$, we can deduce

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

So we can deduce

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

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

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

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

Proof. Observe

$$\begin{aligned} \|\mathbf{v} + \mathbf{w}\|^2 &= \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle \\ &= \|\mathbf{v}\|^2 + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle + \|\mathbf{w}\|^2 \\ &= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + \langle \mathbf{v}, \mathbf{w} \rangle + \overline{\langle \mathbf{v}, \mathbf{w} \rangle} \\ &= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\operatorname{Re}(\langle \mathbf{v}, \mathbf{w} \rangle) \\ &\leq \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2|\langle \mathbf{v}, \mathbf{w} \rangle| \\ &\leq \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2(\|\mathbf{v}\|)(\|\mathbf{w}\|) \\ &= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2 \end{aligned}$$
■

1.7 Existence of Real Numbers - Dedekind Cut*

1.8 Existence of Real Numbers - Decimal*

1.9 Uniqueness of Real Numbers*

1.10 Exercises

Question 1

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

Question 2

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

Question 3

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

Question 4

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

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

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

$$\left| \sum z_j \right| \leq \sum |z_j|$$

Question 8

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

$$||x| - |y|| \leq |x - y|$$

Question 9

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

$$|1 + z|^2 + |1 - z|^2$$

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:

$$y < z \longrightarrow x + y < x + z$$

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

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

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

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 , called a topological space, if we say a family τ of subsets of X is a topology on X , then we mean

$$\begin{aligned} X, \emptyset &\in \tau \\ A, B \in \tau &\implies A \cap B \in \tau \\ \mathfrak{U} \subseteq \tau &\implies \bigcup \mathfrak{U} \in \tau \end{aligned}$$

If a subset of X belong to τ , 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

$$\begin{aligned} X, \emptyset &\in \mathfrak{F} \\ A, B \in \mathfrak{F} &\implies A \cup B \in \mathfrak{F} \\ \mathfrak{B} \subseteq \mathfrak{F} &\implies \bigcap \mathfrak{B} \in \mathfrak{F} \end{aligned}$$

Proof.

$$\begin{aligned} X = \emptyset^c \text{ and } \emptyset = X^c &\in \mathfrak{F} \\ A, B \in \mathfrak{F} &\implies A^c, B^c \in \mathfrak{D} \implies A^c \cap B^c = (A \cup B)^c \in \mathfrak{D} \implies A \cup B \in \mathfrak{F} \end{aligned}$$

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

■

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

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

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

$$f = f^{-1}$$

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.

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

- (a) The set of neighborhood $\mathcal{N}(x)$ for x is non-empty.
- (b) A neighborhood around x must contain x
- (c) A set containing a neighborhood around x is a neighborhood around x
- (d) The intersection of two neighborhoods around x is a neighborhood around x
- (e) Every neighborhood N around x contain an neighborhood M around x such that N is an neighborhood of all $m \in M$

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

$$(g(\tau))(x) = \{N \subseteq X : \exists O \in \tau, x \in O \subseteq N\}$$

In English, g define

N is an neighborhood around $x \iff$ there exists an open set O s.t. $x \in O \subseteq N$

The inverse of g will be proved to be

$$h(\mathcal{N}) = \{O : \forall x \in O, O \in \mathcal{N}(x)\}$$

In English, h define

O is an open set $\iff O$ is an neighborhood around all its element

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

Try to prove that X is an neighborhood around x . The fact an neighborhood around x must contain x and the fact follows from definition.

For proving intersection of two neighborhood around x is an neighborhood around x , use the fact the intersection of two open set is open.

For proving each neighborhood N around x contain an neighborhood M around x such that N is an neighborhood of all $m \in M$, verify for $M = O$, where O is the open set such that $x \in O \subseteq N$ (done)

We now prove **h do map neighborhood function to topology.**

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

Let A, B be open. To prove $A \cap B$ is open, first realize A and B are both neighborhood around $x \in A \cap B$, and use the fact the intersection of neighborhood is an neighborhood.

Let \mathfrak{A} be a class of open sets. To prove $\bigcup \mathfrak{A}$ is open, first use the fact $x \in \mathfrak{A}$ implies x is in some open set $A \in \mathfrak{A}$. Then use the fact A is neighborhood around x , and deduce the super set $\bigcup \mathfrak{A}$ must also be a neighborhood. (done)

Lastly, we prove **h is the inverse of g .**

To prove such, we have to first prove **$h(g(\tau)) = \tau$.**

We first have

$$\begin{aligned} Z \in h(g(\tau)) &\iff \forall z \in Z, Z \in (g(\tau))(z) \\ &\iff \forall z \in Z, \exists O_z \in \tau, z \in O_z \subseteq Z \end{aligned}$$

It is now quite clear that $\tau \subseteq h(g(\tau))$. To see $h(g(\tau)) \subseteq \tau$, collect O_z for each $z \in Z$ and use the fact the union of open set is open. (done)

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

Let $x \in X$. Observe

$$Y \in g(h(\mathcal{N}))(x) \iff \exists O \in h(\mathcal{N}), x \in O \subseteq Y$$

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

Arbitrarily pick $N \in \mathcal{N}(x)$. We wish to prove $N \in g(h(\mathcal{N}))$. Let M be the largest neighborhood around x contained by N such that N is an neighborhood around all $m \in M$. Such greatest M exists because the union of neighborhood around x is again an neighborhood around x .

We now have

$$M = \{m \in N : N \in \mathcal{N}(m)\}$$

since if $M \subset \{m \in N : N \in \mathcal{N}(m)\}$, we can pick any m not in M to see $M \cup \{m\}$ is even larger.

We wish to show the neighborhood $M \subseteq N$ is open to close the proof, by using $M = O$ in equation 2.1.

Arbitrarily pick $m \in M$. We know N is an neighborhood around m , then there exists an neighborhood Q_m around m contained by N such that N is an neighborhood around all $q \in Q_m$. We see $Q_m \subseteq M$ as equation 2.1 suggest. By super set property, we are done. (done) (done) ■

We now give the formal definition of neighborhood while a topology is given, and translate Theorem 2.1.5 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 m .

Theorem 2.1.8. (Property of Neighborhood, Part 2)

$$E \text{ is open} \iff E \text{ is an neighborhood around all points in } E$$

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

From right to left, use the definition of neighborhood for each $x \in E$, and the fact that union of open set is open. ■

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

$$\overline{E} := \bigcap \{F \in \mathfrak{F} : E \subseteq F\}$$

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

$$E^\circ := \bigcup \{O \in \tau : O \subseteq E\}$$

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

$$\begin{aligned} \overline{E} = E &\iff E \in \mathfrak{F} \text{ and } E^\circ = E \iff E \in \tau \\ \overline{(\overline{E})} &= \overline{E} \text{ and } (E^\circ)^\circ = E^\circ \\ S \subseteq T &\implies S^\circ \subseteq T^\circ \text{ and } \overline{S} \subseteq \overline{T} \end{aligned}$$

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

$$\begin{aligned} \overline{E \cup F} &= \overline{E} \cup \overline{F} \text{ and } (E \cap F)^\circ = E^\circ \cap F^\circ \\ \overline{E \cap F} &\subseteq \overline{E} \cap \overline{F} \text{ and } E^\circ \cup F^\circ \subseteq (E \cup F)^\circ \\ (E^\circ)^c &= \overline{E^c} \end{aligned}$$

Proof. Assume C is closed and $E \cup F \subseteq C \subseteq \overline{E} \cup \overline{F}$. WOLG, $E \subseteq C \cap \overline{E} \subseteq \overline{E}$ CaC. Assume O is open and $E^\circ \cap F^\circ \subset O \subseteq E \cap F$. WOLG, $E^\circ \subset O \cup E^\circ \subseteq E$ CaC

Because $E^\circ \subseteq E$, we have $E^c \subseteq (E^\circ)^c$. Assume C is closed and $E^c \subseteq C \subseteq (E^\circ)^c$. We see $E^\circ \subset C^c \subseteq E$ CaC ■

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)

$$p \text{ is an interior point of } E \iff E \text{ is a neighborhood around } p$$

Definition 2.2.5. (Definition of Limit and Isolated Point)

p is limit point of E

\iff every open neighborhood around p contain some $q \in E$ that isn't p

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

Theorem 2.2.6. (Interior Contain Exactly All Interior Points)

Interior contain exactly all interior point

Proof. All interior points has an open neighborhood contained by E , which is contained by E° by definition. Each point in interior is an interior point because interior itself is an open neighborhood. ■

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

$$\overline{E} = E' \cup E$$

Proof. Let $p \in (E' \cup E)^c$. Because p is not a limit point of E , we know there exists an open neighborhood N such that

$$N \cap E \subseteq \{p\}$$

Because $p \notin E$, we then know

$$N \in E^c$$

We have proved $E' \cup E$ is closed. We wish to show E' is contained by all closed set F containing E . Let $p \in F^c$. Because F^c is open, we see p is not a limit point of E . ■

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

$$\overline{E} \setminus E \subseteq E'$$

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

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

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

$$p \in E \text{ or } p \in E'$$

There are two possibilities for point have no disjoint neighborhood

$$p \in E \text{ or } p \in E'$$

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

$$E' = \{p : p \in \overline{E \setminus \{p\}}\}$$

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 $E \cup \{p\}$. ■

We now introduce boundary.

Definition 2.2.11. (Definition of Boundary) We define the boundary of S by the closure of S minus the interior of S

$$\partial S = \overline{S} \setminus S^\circ$$

Theorem 2.2.12. (Equivalent Definition for Boundary) We have

$$\partial S = \overline{S} \cap \overline{S^c}$$

and

$$\partial S = \{p : \text{for every open neighborhood } O \text{ around } p, O \cap S \neq \emptyset \neq O \cap S^c\}$$

Proof. Notice $\overline{S^c} = (S^\circ)^c$ from Theorem 2.2.3 and we are done. The second result follows from $\partial S = \overline{S} \cap \overline{S^c}$. ■

Corollary 2.2.13. (Boundary is Closed) We have

$$\partial S \text{ is closed}$$

Corollary 2.2.14. (Boundary of Boundary is Subset of Boundary) For all S , we have

$$\partial \partial S \subseteq \partial S$$

Proof. ∂S is closed tell us

$$\partial \partial S = \partial S \setminus (\partial S)^\circ$$

Theorem 2.2.15. (Boundary is also the Boundary of Complement)

$$\partial S = \partial(S^c)$$

Proof. This follows from $\partial S = \overline{S} \cap \overline{S^c}$ ■

We now introduce exterior

Definition 2.2.16. (Definition of Exterior) We define the exterior by

$$\text{ext}(S) := (\overline{S})^c$$

Theorem 2.2.17. (Equivalent definition of Interior) We have

$$X = S^\circ \cup \partial S \cup \text{ext}(S)$$

and have

$$\text{ext}(S) = (S^c)^\circ$$

Proof. The first result follows straight from definition, as $\partial S \cup S^\circ = \overline{S}$. Rewrite Theorem 2.2.3 into

$$(E^\circ)^c = \overline{E^c} \implies ((E^c)^\circ)^c = \overline{E}$$

Then we have

$$\text{ext}(S) = (\overline{S})^c = (S^c)^\circ$$

■

Theorem 2.2.18. (Basic Property of Exterior) We have

$$S \subseteq T \implies \text{ext}(T) \subseteq \text{ext}(S)$$

$$S^\circ \subseteq \text{ext}(\text{ext}(S))$$

Proof.

$$S \subseteq T \implies \overline{S} \subseteq \overline{T} \implies \text{ext}(T) = (\overline{T})^c \subseteq (\overline{S})^c = \text{ext}(S)$$

Deduce

$$S \subseteq \overline{S} \implies (\overline{S})^c \subseteq S^c \implies \overline{(\overline{S})^c} \subseteq \overline{S^c} = (S^\circ)^c \implies S^\circ \subseteq (\overline{(\overline{S})^c})^c = \text{ext}(\text{ext}(S))$$

■

2.3 Different Types of Connected

Definition 2.3.1. (Definition of Path) By a path from a to b in a subset $S \subseteq X$, we mean a

Continuous function $f : [0, 1] \rightarrow S$ such that $f(0) = a$ and $f(1) = b$

Definition 2.3.2. (Definition of Path-Connected) We say a set S is path-connected if

For each $a, b \in S$, there exists a path f in S from a to b

Definition 2.3.3. (Definition of Separated) We say A, B are separated if

$$\overline{A} \cap B = \emptyset = A \cap \overline{B} \text{ and } A, B \text{ are non-empty}$$

Definition 2.3.4. (Definition of Connected) We say a set V is connected if

V can not be expressed as union of two non-empty separated sets

Theorem 2.3.5. (Equivalent Definition of Connected) The following are equivalent

- (a) V is connected
- (b) V has no proper non-empty subset U relatively clopen to V

Proof. We first prove V is connected $\implies V$ has no proper non-empty subset U relatively clopen to V .

Assume U is a proper non-empty subset of V relatively clopen to V . Define

$$W := V \setminus U$$

We know W is also a proper non-empty subset of V relatively clopen to V . WOLG, we only wish to prove

$$\overline{U} \cap W = \emptyset$$

Because U is relatively closed to V , we know there exists a close set F such that

$$U = F \cap V$$

Then by definition of closure, we know

$$\overline{U} \subseteq F$$

Notice that F and W are disjoint, otherwise there exists some $x \in F \cap W$, and we can deduce

$$x \in W \implies x \in V \implies x \in F \cap V = U (\because x \in F \cap W)$$

which is clearly impossible, since W and U are disjoint.

Then because F and W are disjoint, we see

$$\overline{U} \cap W \subseteq F \cap W = \emptyset$$

WOLG, we see U and W are separated and $V = U \cup W$ **CaC** (done)

We now prove V has no proper non-empty subset U relatively clopen to $V \implies V$ is connected.

Assume V is disconnected. Then we can express V into union of non-empty separated subsets A, B

$$V = A \cup B \text{ and } \overline{A} \cap B = \emptyset = A \cap \overline{B}$$

We wish to prove

$$A \text{ is relatively clopen to } V$$

We first prove

$$A = \overline{A} \cap V$$

It is clear that

$$A \subseteq \overline{A} \cap V$$

Let $x \in \overline{A} \cap V$. Because $\overline{A} \cap B = \emptyset$, we know

$$x \notin B$$

Then because $x \in V$, we know

$$x \in A \text{ (done)}$$

WOLG, from teal part, we also know B is relatively closed to V , then we can deduce A is relatively clopen to V .

Because B is non-empty, we know A is proper. We have proved A is an non-empty proper subset of V relatively clopen to V **CaC**

■

Theorem 2.3.6. (Union of non-Disjoint Connected Sets is Connected)

$$A, B \text{ are both connected and } A \cap B \neq \emptyset \implies A \cup B \text{ is connected}$$

Proof. Assume $A \cup B$ is disconnected. Let

$$V \cup W = A \cup B \text{ and } V, W \text{ be separated}$$

Let $x \in A \cap B$. Because $A \cup B = V \cup W$, we see x must belong to one of V, W . WOLG, let $x \in V$. We know W must intersect with one of A, B . WOLG, let $W \cap A \neq \emptyset$. We see

$$x \in A \cap B \text{ and } x \in V \implies V \cap A \neq \emptyset$$

Notice

$$\overline{W \cap A} \subseteq \overline{W} \cap \overline{A} \text{ and } \overline{V \cap A} \subseteq \overline{V} \cap \overline{A} \text{ and } (V \cap A) \cup (W \cap A) = A \text{ CaC}$$

■

Theorem 2.3.7. (Path Connected is Stronger than Connected)

$$S \text{ is path-connected} \implies S \text{ is connected}$$

Proof. Assume S is not connected. Express S into union of two proper non-empty separated subsets A, B

$$S = A \cup B \text{ where } A, B \text{ are separated}$$

Let

$$a \in A \text{ and } b \in B$$

Because S is path-connected, we know there exists a path f in S from a to b . Consider

$$c := f(\sup f^{-1}[A])$$

and denote $x := \sup f^{-1}[A] \in \mathbb{R}$. Remember

$$c = f(x)$$

Because $c \in S$, we know

$$c \text{ belong to either } A \text{ or } B$$

To cause contradiction, we wish to prove

$$\begin{cases} c \in B \implies c \in \overline{A} \\ c \in A \implies c \in \overline{B} \end{cases}$$

We first prove

$$c \in B \implies c \in \overline{A}$$

Let C be an open neighborhood around c , we only wish to prove

$$\text{There exists } c' \neq c \text{ such that } c' \in A \cap C$$

We know

$$f(x) = c \in C$$

Then because f is continuous, we know there exists an open set O in $[0, 1]$ such that

$$x \in O \text{ and } f [O] \subseteq C$$

Because O is an open neighborhood around $x \in \mathbb{R}$, we know there exists an open ball $B_r (x)$ contained by O . Clearly,

$$x - \frac{r}{2} < x = \sup f^{-1} [A]$$

Then we know there exists some $y \in f^{-1} [A]$ such that

$$x - \frac{r}{2} < y \leq x$$

Clearly,

$$y \in \left(x - \frac{r}{2}, x \right] \subseteq B_r (x) \subseteq O$$

We now prove $f (y)$ is the wanted c' to finish the partial proof.

Notice that

$$y \in f^{-1} [A] \implies f (y) \in A \implies f (y) \neq c \quad (\because c \notin A)$$

Because $x - \frac{r}{2} < y \leq x$, we can deduce

$$y \in B_r (x) \subseteq O$$

and deduce

$$f (y) \in f [O] \subseteq C$$

We have proved

$$f (y) \neq c \text{ and } f (y) \in A \cap C \text{ (done) (done)}$$

We now prove

$$c \in A \implies c \in \overline{B}$$

Let C be an open neighborhood around c , we only wish to prove

$$\text{There exist } c' \neq c \text{ such that } c' \in B \cap C$$

We know

$$f (x) = c \in C$$

Then because f is continuous, we know there exists an open set O in $[0, 1]$ such that

$$x \in O \text{ and } f [O] \subseteq C$$

Because $f (x) = c \in A$, we know $x \neq 1$, otherwise $f (x) \in A \cap B$, which is clearly impossible.

Then we know there exists $y \in O$ such that

$$x < y$$

Because $x = \sup f^{-1}[A]$, we know

$$y \notin f^{-1}[A]$$

Then we can deduce

$$f(y) \in B$$

and from $y \in O$, deduce

$$f(y) \in f[O] \subseteq C$$

Notice that from $f(y) \in B$ and $c \in A$, we know

$$f(y) \neq c$$

We have proved

$$f(y) \neq c \text{ and } f(y) \in B \cap C \text{ (done)}$$

■

We now introduce connected component.

Theorem 2.3.8. (Path-Connected Component) Define a relation \sim in S by

$$a \sim b \text{ if there exists a path in } S \text{ from } a \text{ to } b$$

Then

$$\sim \text{ is an equivalence relation}$$

Proof. To see $a \sim a$

$$\text{define } f(x) = a$$

The inverse image of open set is either \emptyset or $[0, 1]$. Both are open in $[0, 1]$.

We now prove

$$a \sim b \implies b \sim a$$

Let f be a path in S from a to b . It is clear that $g : [0, 1] \rightarrow [0, 1]$

$$g(x) = 1 - x \text{ is continuous}$$

Then we see

$$f \circ g \text{ is a path in } S \text{ from } b \text{ to } a \text{ (done)}$$

We now prove

$$a \sim b \text{ and } b \sim c \implies a \sim c$$

Let f be a path in S from a to b , and let g be a path in S from b to c . Define $h : [0, 1] \rightarrow S$ by

$$h(x) = \begin{cases} f(2x) & \text{if } x \in [0, \frac{1}{2}) \\ g(2(x - \frac{1}{2})) & \text{if } x \in [\frac{1}{2}, 1] \end{cases}$$

Suppose $x \in [0, \frac{1}{2})$. We wish to prove

h is continuous at x

Let $W \subseteq S$ be a neighborhood around $h(x)$. We wish to find a neighborhood $V \subseteq [0, 1]$ around x such that

$$h[V] \subseteq W$$

Because f is continuous at $2x$, we know there exist a neighborhood $T \subseteq [0, 1]$ around $2x$ such that

$$h(x) = f(2x) \in f[T] \subseteq W$$

Let

$$T_{\frac{1}{2}} = \{\frac{t}{2} : t \in T\}$$

Clearly, $T_{\frac{1}{2}}$ is a neighborhood around x , such that $T_{\frac{1}{2}} \subseteq [0, \frac{1}{2}]$, and we can see

$$h[T_{\frac{1}{2}}] = f[T] \subseteq W \text{ (done)}$$

Suppose $x \in [\frac{1}{2}, 1]$. We wish to prove

h is continuous at x

Let $W \subseteq S$ be a neighborhood around $h(x)$. We wish to find a neighborhood $V \subseteq [0, 1]$ around x such that

$$h[V] \subseteq W$$

Because g is continuous at $2(x - \frac{1}{2})$, we know there exists a neighborhood $T \subseteq [0, 1]$ around $2(x - \frac{1}{2})$ such that

$$h(x) = g(2(x - \frac{1}{2})) \in g[T] \subseteq W$$

Let

$$T'_{\frac{1}{2}} = \{\frac{t}{2} + \frac{1}{2} : t \in T\}$$

Clearly, $T'_{\frac{1}{2}}$ is a neighborhood around x , such that $T'_{\frac{1}{2}} \subseteq [\frac{1}{2}, 1]$, and we can see

$$h[T'_{\frac{1}{2}}] \subseteq g[T] \subseteq W \text{ (done)}$$

■

Definition 2.3.9. (Definition of Path-Connected Component) Partition S by equivalence relation \sim defined in Theorem 2.3.8 into

$$\{S_\lambda : \lambda \in \Lambda\}$$

We say each S_λ is an path-connected component.

Theorem 2.3.10. (Connected Components are Mutually Separated)

$$A, B \text{ are non-equal connected component of } S \implies A, B \text{ are separated}$$

Proof. ■

Theorem 2.3.11. (Topologist's Sine Curve) In standard topology of \mathbb{R}^2

$$\left\{ \left(x, \sin \frac{1}{x} \right) : x \in (0, 1) \right\} \cup \{(0, 0)\} \text{ is connected}$$

Proof. Denote $S = \left\{ \left(x, \sin \frac{1}{x} \right) : x \in (0, 1) \right\} \cup \{(0, 0)\}$, and arbitrarily partition S into non-empty disjoint A, B

$$S = A \cup B$$

We wish to prove

$$\overline{A} \cap B \neq \emptyset \text{ or } A \cap \overline{B} \neq \emptyset$$

WOLG, suppose $(0, 0) \in A$. Denote $(0, 0)$ by g . Consider the set

$$\{x : (x, \sin \frac{1}{x}) \in B\}$$

and let

$$\alpha = \inf \{x : (x, \sin \frac{1}{x}) \in B\}$$

If $\alpha = 0$, there are two possibilities

$$\begin{cases} \exists t \in \mathbb{R}^+, \forall x \in (0, t), (x, \sin \frac{1}{x}) \in B \\ \forall t \in \mathbb{R}^+, \exists x \in (0, t), (x, \sin \frac{1}{x}) \notin B \end{cases}$$

If $\exists t \in \mathbb{R}^+, \forall x \in (0, t), (x, \sin \frac{1}{x}) \in B$, we can construct a sequence

$$\left\{ \left(\frac{1}{n\pi}, (\sin n\pi) \right) \right\}_{n \in \mathbb{N}}$$

which converge to g , and its tail is in B . It is now easy to see

$$(0, 0) \in \overline{B} \text{ (done)}$$

If $\forall t \in \mathbb{R}^+, \exists x \in (0, t), (x, \sin \frac{1}{x}) \notin B$, arbitrarily pick a path connected component B_1 of B . Consider $f(\inf p[B_1])$

Theorem 2.3.12. (Conjecture) In \mathbb{R}^n , two connected subset A, B are disjoint implies A, B are separated.

2.4 Continuous Function

Definition 2.4.1. (Definition of Continuous Function) We say a function $f : X \rightarrow Y$ is continuous if

$$V \text{ is open} \implies f^{-1}[V] \text{ is open}$$

Definition 2.4.2. (Definition of Continuous at a point) We say f is continuous at x if

For all neighborhood V around $f(x)$, there exists neighborhood U around x such that $f[U] \subseteq V$

Lemma 2.4.3. Given $f : X \rightarrow Y$, we have

$$f[f^{-1}[V]] \subseteq V$$

where the equality hold true when $V \subseteq f[X]$

Proof. Deduce

$$\begin{aligned} y \in f[f^{-1}[V]] &\implies \exists x \in f^{-1}[V], y = f(x) \\ &\implies \exists y' \in V, y' = f(x) = y \end{aligned}$$

Suppose $V \subseteq f[X]$ and $y \in V$. We know

$$f^{-1}[y] \text{ is non empty}$$

Then We know there exists x such that

$$y = f(x)$$

Because $y \in V$, we know $x \in f^{-1}[V]$, then we see

$$y \in f[f^{-1}[V]]$$

■

Theorem 2.4.4. (Equivalent Definition of Continuous) Given $f : X \rightarrow Y$. The following are equivalent

- (a) f is continuous
- (b) every inverse image of closed set is closed
- (c) f is continuous at all $x \in X$
- (d) for all $A \subseteq X$, we have $f[\overline{A}] \subseteq \overline{f[A]}$

Proof. From (a) to (b), deduce from V is closed to $f^{-1}[V^c]$ is open, and deduce $f^{-1}[V] = (f^{-1}[V^c])^c$.

From (b) to (a), deduce from V is open to $f^{-1}[V^c]$ is closed, and deduce $f^{-1}[V] = (f^{-1}[V^c])^c$.

We now prove f is continuous $\implies f$ is continuous at all $x \in X$. Let V be an neighborhood around $f(x)$. We wish to find an neighborhood U around x such that $f[U] \subseteq V$. By definition of neighborhood, we know there exists an open set O such that

$$f(x) \in O \subseteq V$$

Because f is continuous, we know

$$f^{-1}[O] \text{ is open}$$

Observe that by lemma 2.4.3, we have

$$x \in f^{-1}[O] \text{ and } f[f^{-1}[O]] \subseteq O \subseteq V \text{ (done)}$$

We now prove f is continuous at all $x \in X \implies f$ is continuous. Let $V \subseteq Y$ be open. We wish to prove $f^{-1}[V]$ is open.

Notice that V is an neighborhood around all $y \in V$, then because f is continuous at all $x \in X$, we know for all $x \in f^{-1}[V]$, there exists an neighborhood U_x around x such that

$$f[U_x] \subseteq V$$

Then for all $x \in f^{-1}[V]$, there exists an open neighborhood O_x around x such that

$$f[O_x] \subseteq V$$

This give us

$$O_x \subseteq f^{-1}[V]$$

Then if we collect all such open neighborhood O_x , we see

$$\bigcup_{x \in f^{-1}[V]} O_x = f^{-1}[V] \text{ (done)}$$

We now prove

$$f \text{ is continuous at all } x \in X \implies \forall A \subseteq X, f[\overline{A}] \subseteq \overline{f[A]}$$

It is clear that

$$f[A] \subseteq \overline{f[A]}$$

Let $a' \in \overline{A} \setminus A$, we reduce the problem into proving

$$f(a') \in \overline{f[A]}$$

In fact, we will only have to prove

$$f(a') \in (f[A])'$$

Let N be an neighborhood around $f(a')$, we reduce the problem into finding an $p \in f[A]$ such that

$$p \neq f(a') \text{ and } p \in N$$

Because $a' \notin A$, we don't have to worry about $p \neq f(a')$. Because f is continuous at a' , we know there exists a neighborhood N_X around a' such that

$$f[N_X] \subseteq N$$

Then because $a' \in \overline{A} \setminus A$ is a limit point of A , we know there exists $a \in A$ such that

$$a \in N_X$$

This give us

$$f(a) \in N \text{ (done)}$$

We now prove

$$\forall A \subseteq X, f[\overline{A}] \subseteq \overline{f[A]} \implies \text{every inverse image of closed set is closed}$$

Let $B \subseteq Y$, and $A = f^{-1}[B]$. Suppose

$$\overline{B} = B$$

We reduce the problem into proving

$$\overline{A} \subseteq A$$

From premise, we have

$$f[\overline{A}] \subseteq \overline{f[A]} = B$$

This implies

$$\overline{A} \subseteq f^{-1}[B] = A \text{ (done)}$$

■

We now introduce other concepts we are familiar.

Theorem 2.4.5. (Composition of Continuous Function is Continuous) Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be continuous. We have

$$g \circ f \text{ is continuous}$$

Proof. Let $V \subseteq Z$ be open. We know

$$g^{-1}[V] \text{ is open}$$

This give us

$$f^{-1}[g^{-1}[V]] \text{ is open}$$

Deduce

$$\begin{aligned} x \in (g \circ f)^{-1}[V] &\iff g(f(x)) \in V \\ &\iff f(x) \in g^{-1}[V] \\ &\iff f^{-1}[g^{-1}[V]] \end{aligned}$$

■

We now introduce connected.

Theorem 2.4.6. (Continuity Map Connected Set to Connected Set) Let D be the domain of continuous function f . Let $A \subseteq D$.

$$A \text{ is connected} \implies f[A] \text{ is connected}$$

Proof. We prove the contrapositive

$$f[A] \text{ is disconnected} \implies A \text{ is disconnected}$$

Suppose

$$f[A] = V \cup W \text{ and } V, W \text{ are separated}$$

Notice

$$\begin{aligned} (f^{-1}[V] \cap A) \cup (f^{-1}[W] \cap A) &= (f^{-1}[V] \cup f^{-1}[W]) \cap A \\ &= f^{-1}[V \cup W] \cap A \\ &= A \end{aligned}$$

WOLG, we only wish to prove

$$\overline{f^{-1}[V] \cap A} \cap (f^{-1}[W] \cap A) = \emptyset$$

Notice

$$\overline{f^{-1}[V] \cap A} \subseteq \overline{f^{-1}[V]} \cap \overline{A} \subseteq \overline{f^{-1}[\overline{V}]} \cap \overline{A} = f^{-1}[\overline{V}] \cap \overline{A}$$

We have reduce our problem into proving

$$f^{-1}[\overline{V}] \cap \overline{A} \cap f^{-1}[W] \cap A = \emptyset$$

Assume $x \in f^{-1}[V] \cap A \cap f^{-1}[W]$. We see $f(x) \in \overline{V}$ and $f(x) \in W$, which is impossible, since V, W are separated. **CaC**

■

Lastly, about compact.

Theorem 2.4.7. (Continuous Function Maps Compact Set to Compact Set) Let D be the domain of f and f be continuous. Let $A \subseteq D$.

$$A \text{ is compact} \implies f[A] \text{ is compact}$$

Proof. Let $\{G_i\}$ be an open cover of $f[A]$. We first wish to show

$$\{f^{-1}[G_i]\} \text{ is an open cover for } A$$

It is clear that $f^{-1}[G_i]$ are all open.

Let $x \in A$. We know $f(x) \in f[A]$. Then we know there exists some G_j such that

$$f(x) \in G_j$$

Then we see $x \in f^{-1}[G_j]$ (done)

Because A is compact, we know there exists a finite sub-cover $\{f^{-1}[G_\lambda] : \lambda \in \Lambda\}$.

We wish to prove

$$\{G_\lambda : \lambda \in \Lambda\} \text{ is an open cover for } f[A]$$

Let $y \in f[A]$. We know there exists $x \in A$ such that

$$y = f(x)$$

Then we know there exists λ_j such that

$$x \in f^{-1}[G_{\lambda_j}]$$

Then we see

$$y = f(x) \in G_{\lambda_j} \text{ (done)}$$

■

Theorem 2.4.8. (Closed Subset of Compact Set is Compact) Let X be compact, and $E \subseteq X$

$$E \text{ is closed} \implies E \text{ is compact}$$

Proof. Let $\{G_\lambda : \lambda\}$ be an open cover of E . Consider

$$U := \{G_\lambda\} \cup \{E^c\}$$

Because E is closed, clearly U is an open cover of X . Then U has a finite sub-cover U' that cover X . If U' contain no E^c , we see that U' is also a finite sub-cover of $\{G_\lambda\}$ that cover E . If U' contain E^c , we see that

$U' \setminus \{E^c\}$ is a finite sub-cover of $\{G_\lambda\}$ that cover E

Since

$$e \in E \subseteq X \implies e \in \bigcup U' \text{ and } e \notin E^c \implies \exists S \in U' \setminus \{E^c\}, e \in S$$

■

2.5 Homeomorphism

Theorem 2.5.1. (Inverse Function of One-to-One Continuous Function is Again Continuous) Let $f : X \rightarrow Z$ be continuous and one-to-one. Let Y be the range of f . Consider the sub-space topology (Y, τ) . The inverse function

$$f^{-1} : (Y, \tau) \rightarrow X \text{ is also continuous}$$

Proof. **Wrong Conjecture** ■

2.6 Three Non-Equivalent Compactness and Separable

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

Definition 2.6.1. (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

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

We call open sets in a base basic open sets.

Theorem 2.6.2. (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

$$\forall O \in \tau, \exists \mathcal{U} \subseteq \mathcal{O}, O = \bigcup \mathcal{U}$$

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

$$X \text{ has a countable dense subset}$$

Theorem 2.6.4. (Base and Dense Set)

A set intersecting with all basic open set is dense

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 2.6.5. (Countable Base Implies Separable)

$$X \text{ has a countable base} \implies X \text{ is separable}$$

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 2.6.4, 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.

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 2.6.6. (Compact Implies Countably Compact)

$$X \text{ is compact} \implies X \text{ is countably compact}$$

Theorem 2.6.7. (Countably Compact Implies Limit Point Compact)

$$X \text{ is countably compact} \implies X \text{ is limit point compact}$$

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

$$A' = \emptyset \implies \overline{A} = A \implies A^c \text{ is open}$$

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

Lastly, we give a theorem involving countable base.

Theorem 2.6.9. (Existence of Countable Base Implies Existence of Countable Sub-Cover)

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

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

$$p \in O_p \subseteq G_p$$

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 Basic Fact of Sequence

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

$$d(x, x) = 0$$

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

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

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

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

$$\lim_{n \rightarrow \infty} a_n = a$$

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

$$n > N_\epsilon \implies |a_n - a| < \epsilon$$

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

$$\lim_{n \rightarrow \infty} a_n = a$$

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

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

Above is definition.

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

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n = a &\iff \lim_{n \rightarrow \infty} d(a_n, a) = 0 \\ &\iff (\forall \epsilon, \exists N_\epsilon, n \geq N_\epsilon \iff d(a_n, a) < \epsilon) \\ &\iff (\forall \epsilon, \exists N_\epsilon, a_n \in B_\epsilon(a) \iff n \geq N_\epsilon) \end{aligned}$$

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

$$\lim_{n \rightarrow \infty} a_n = a \text{ and } \lim_{n \rightarrow \infty} a_n = a' \implies a' = a$$

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

$$\{a_{n_k}\}$$

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

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

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

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

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)

$$\{a_n\} \text{ converge to } a$$

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

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

$\{a_n\}$ is a Cauchy sequence

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

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

Theorem 3.1.11. (Every Convergent Sequence is Cauchy)

$$\{a_n\} \text{ converge} \implies \{a_n\} \text{ is Cauchy}$$

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)

$$\{a_n\} \text{ is Cauchy NOT} \implies \{a_n\} \text{ converge}$$

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

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

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

$$\frac{1}{k} < \frac{|p-q|}{2} \implies \frac{q-p}{2} < \frac{1}{k} < \frac{p-q}{2}$$

And observe

$$\begin{aligned} d(a_k, q) &= \left| p + \frac{1}{k} - q \right| > p - q + \frac{1}{k} > \frac{p-q}{2} \\ &> q - p - \frac{1}{k} > \frac{q-p}{2} \end{aligned}$$

So we have

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

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:

$$\overline{E} = E$$

Theorem 3.2.2. (Another Definition of Closure)

\overline{E} is the smallest closed set containing E

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

$$\lim_{n \rightarrow \infty} x_n = x$$

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 3.1.8. 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}

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

This also give us

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

Now, we wish to prove

$$\lim_{k \rightarrow \infty} x_{m_k}^k = x$$

Because

$$\lim_{k \rightarrow \infty} x_{n_k} = x$$

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

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

Then we have

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

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

$$\begin{aligned} \overline{X} = X \text{ and } \overline{\emptyset} = \emptyset \\ \bigcup_{n=1}^{\infty} \overline{A_n} \subseteq \overline{\bigcup_{n=1}^{\infty} A_n} \text{ and } \overline{\bigcap_{n=1}^{\infty} A_n} \subseteq \bigcap_{n=1}^{\infty} \overline{A_n} \end{aligned}$$

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

Proof. $\overline{X} = X$ and $\overline{\emptyset} = \emptyset$ 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

$$\bigcup_{n=1}^{\infty} \overline{A_n} = \mathbb{Q}$$

We also have

$$\overline{\bigcup_{n=1}^{\infty} A_n} = \overline{\mathbb{Q}} = \mathbb{R}$$

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 3.2.4. ■

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

$$\overline{E \cup F} = \overline{E} \cup \overline{F} \text{ and } \overline{E \cap F} \subseteq \overline{E} \cap \overline{F}$$

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

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

We have

$$\overline{E \cap F} = \overline{\emptyset} = \emptyset \subset \{1\} = [0, 1] \cap [1, 2] = \overline{E} \cap \overline{F}$$
■

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

\emptyset and X are closed.

Intersection of closed sets are closed.

Union of finite closed sets are closed, but union of infinite closed set may not be closed.

Proof. In Theorem 3.2.3, 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 3.2.3, we have

$$\bigcap_{\lambda \in \Lambda} A_\lambda \subseteq \bigcap_{\lambda \in \Lambda} \overline{A_\lambda} = \bigcap_{\lambda \in \Lambda} A_\lambda$$

And notice the other side is trivial.

Finite union of closed sets is closed is an immediate consequence of Theorem 3.2.4 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

$$\overline{\bigcup_{\lambda \in \Lambda} A_\lambda} = \bigcup_{\lambda \in \Lambda} A_\lambda = \bigcup_{\lambda \in \Lambda} \overline{A_\lambda}$$

but Theorem 3.2.3 and Theorem 3.2.4 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

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

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

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

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

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

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

$$\overline{E} = E' \cup E$$

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

$$E \text{ is closed} \iff E' \subseteq E$$

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)

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

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

$$\begin{aligned} (E \cap F)' &\subseteq E' \cap F' \\ E' \cup F' &= (E \cup F)' \end{aligned}$$

Where the \subseteq may be proper. We also have

$$E \subseteq F \implies E' \subseteq F'$$

Where the converse may not hold true. We also have

$$E' \text{ is closed.}$$

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

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

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

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

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

$$\lim_{k \rightarrow \infty} x_k^u = x_u$$

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

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

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

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

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

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

Observe

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

■

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 is a limit point of $E \cap Y$ in $Y \implies p$ is a limit point of E in X

The closure of $E \cap Y$ in Y is a subset of the closure of E in X

E is closed in $X \implies E \cap Y$ is closed in Y

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 :

$$\exists N, \forall n > N, x_n \in E$$

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 :

$$E^\circ = E$$

Theorem 3.3.2. (Another Definition of Interior)

E° is the greatest open set in E

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 3.1.9, we know there exists a sub-sequence $\{x_r\}$ of $\{x_u\}$ such that

$$\forall r \in \mathbb{N}, d(x_r, x) < \frac{1}{r}$$

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 3.1.9, we know there exists a sub-sequence $\{x_k^r\} \subseteq \{x_m^r\}$ such that

$$\forall k \in \mathbb{N}, d(x_k^r, x_r) < \frac{1}{k}$$

Now for each r , notice that $k = r$ implies

$$d(x_k^r, x) < d(x_k^r, x_r) + d(x_r, x) < \frac{1}{k} + \frac{1}{r} = \frac{2}{r}$$

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

$$\forall \epsilon, r > \frac{2}{\epsilon} \implies d(x_r^r, x) < \frac{2}{r} < \epsilon$$

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

$$\begin{aligned} X &= X^\circ \text{ and } \emptyset^\circ = \emptyset \\ \bigcup_{n=1}^{\infty} A_n^\circ &\subseteq \left(\bigcup_{n=1}^{\infty} A_n\right)^\circ \\ \left(\bigcap_{n=1}^{\infty} A_n\right)^\circ &\subseteq \bigcap_{n=1}^{\infty} A_n^\circ \end{aligned}$$

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

$$\bigcup_{n=1}^{\infty} A_n^\circ = \bigcup_{n=1}^{\infty} \emptyset = \emptyset \subset (0, 1) = \left(\bigcup_{n=1}^{\infty} A_n\right)^\circ$$

For an example of the third statement may be proper, let $A_n := (1 - \frac{1}{n}, 1 + \frac{1}{n})$. Then

$$\left(\bigcap_{n=1}^{\infty} A_n\right)^\circ = \{1\}^\circ = \emptyset \subset \{1\} = \bigcap_{n=1}^{\infty} A_n^\circ$$

■

Theorem 3.3.4. (Manipulation of Interior) We have

$$\begin{aligned} E^\circ \cup F^\circ &\subseteq (E \cup F)^\circ \\ (E \cap F)^\circ &= E^\circ \cap F^\circ \end{aligned}$$

And if $E \subseteq F$, we have

$$E^\circ \subseteq F^\circ$$

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

$$E^\circ \cup F^\circ = (0, 1) \cup (1, 2) \neq (0, 2) = [0, 2]^\circ = (E \cup F)^\circ$$

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

$$n > N_E \implies x_n \in E$$

and let N_F satisfy

$$n > N_F \implies x_n \in F$$

Then we see

$$n > \max\{N_E, N_F\} \implies x_n \in E \cap F$$

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 and \emptyset are open.

Union of open sets are open.

Finite intersection of open sets are open

Infinite intersection of open sets may not be open.

Proof. In Theorem 3.3.3, we have proved X and \emptyset are open. Let $\{A_\lambda : \lambda \in \Lambda\}$ be a collection of open sets. By Theorem 3.3.3, we know

$$\bigcup_{\lambda \in \Lambda} A_\lambda = \bigcup_{\lambda \in \Lambda} A_\lambda^\circ \subseteq \left(\bigcup_{\lambda \in \Lambda} A_\lambda \right)^\circ$$

The other side follows from Theorem 3.3.2.

Finite intersection of open sets are open follows from Theorem 3.3.4:

$$(E \cap F)^\circ = E^\circ \cap F^\circ = E \cap F$$

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

$$(E^\circ)^c = \overline{E^c}$$

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)

$$\overline{E^\circ} \subseteq \overline{E} \text{ and } E^\circ \subseteq (\overline{E})^\circ$$

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

$$\overline{E^\circ} = \overline{\emptyset} = \emptyset \subset \mathbb{R} = \overline{\mathbb{Q}} = \overline{E}$$

And we have

$$E^\circ = \emptyset \subset \mathbb{R} = (\mathbb{R})^\circ = (\overline{\mathbb{Q}})^\circ = (\overline{E})^\circ$$

For $\overline{E^\circ} \not\subseteq (\overline{E})^\circ$, let $E = [0, 1]$. We have

$$\overline{E^\circ} = [0, 1] \not\subseteq (0, 1) = (\overline{E})^\circ$$

For $(\overline{E})^\circ \not\subseteq \overline{E^\circ}$, let $E = \mathbb{Q}$. We have

$$(\overline{E})^\circ = \mathbb{R} \not\subseteq \emptyset = \overline{E^\circ}$$

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

$$\begin{aligned} p \text{ is an interior point of } E \text{ in } X &\implies p \text{ is an interior point of } E \cap Y \text{ in } Y \\ Y \cap E^\circ \text{ in } X &\text{ is a subset of the interior of } E \cap Y \text{ in } Y \\ E \text{ is open in } X &\implies E \cap Y \text{ is open in } Y \end{aligned}$$

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

$$n > N \implies x_n \in E$$

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)

$$B_r(p) \text{ is open}$$

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)

$$\begin{aligned} p \text{ is an interior point of } E \\ \iff \text{there exists an small enough open ball } B_r(p) \text{ contained by } E \end{aligned}$$

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:

$$\forall \epsilon, B_\epsilon(p) \setminus E \neq \emptyset$$

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

$$n > N \implies x_n \in B_r(p) \subseteq E$$

This finish our proof. ■

Corollary 3.3.11. (Metric Spaces are Sequential)

E is open

\iff for each point p in E there exists an small enough open ball $B_r(p)$ contained by E

Lastly, the most important.

Theorem 3.3.12. (Topology Axiom)

$$E \text{ is open} \iff E^c \text{ is closed}$$

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

$$n > N \implies x_n \notin E^c \implies x_n \in E$$

■

3.4 Connected and Convex

Theorem 3.4.1. (Interval is Connected)

$[0, 1], (0, 1), [0, 1), (0, 1]$ are all connected

Proof. Arbitrarily partition $[0, 1]$ into $A \cup B$. We wish to prove

A, B are not separated

If $\sup A = 1 = \sup B$, notice that 1 is, WOLG, in A , and we can construct sequence in B that converge to 1. This give us $1 \in A \cap \overline{B}$. WOLG, let $\sup A \neq 1$.

If $\sup A \in A$. We see the tail of

$$\sup A + 0.1, \sup A + 0.01, \sup A + 0.001, \dots$$

is a sequence in B that converge to A . This give us $\sup A \in A \cap \overline{B}$

If $\sup A \notin A$. Then $\sup A \in B$, and we can construct a sequence in A that converge to $\sup A$. This give us $\sup A = \overline{A} \cap B$ (done)

Arbitrarily partition $(0, 1)$ into $A \cup B$. We wish to prove

A, B are not separated

Let $a \in A$. There are two possibilities:

$$\begin{cases} \{b \in B : b < a\} \text{ is empty} \\ \{b \in B : b < a\} \text{ is non-empty} \end{cases}$$

If $\{b \in B : b < a\}$ is empty, we see

$$0 < a \leq \inf B$$

Then there are two possibility: $\inf B \in B$ or $\inf B \notin B$. The former give us an sequence in A converge to $\inf B$ and the latter give us a sequence in B converge to $\inf B \in A$.

If $\{b \in B : b < a\}$ is non-empty. Let

$$b' = \sup\{b \in B : b < a\}$$

If $b' \in B$ we construct a sequence in A that converge to b' . If $b' \in A$, we construct a sequence in B that converge to b' (done)

For $[0, 1)$ and $(0, 1]$ use the same argument for violet part with infimum when dealing with $[0, 1)$. ■

Theorem 3.4.2. (Non-Singleton Connected Set In \mathbb{R}^k Must be Uncountable) In \mathbb{R}^k

S is connected and $|S| > 1 \implies S$ is uncountable

Proof. ■

Theorem 3.4.3. (Cantor Set is Uncountable and totally disconnected)

We now discuss convex set.

3.5 Bounded

Definition 3.5.1. (Definition of Bounded and Totally Bounded) We say E is bounded if there exists $(x \in X, r \in \mathbb{R}^+)$ such that

$$E \subseteq B_r(x)$$

We say E is totally bounded if for all ϵ there exists a finite set

$$\{x_1, x_2, \dots, x_m\}$$

such that

$$E \subseteq \bigcup_{i=1}^m B_\epsilon(x_i)$$

Theorem 3.5.2. (Totally Bounded Implies Bounded)

$$E \text{ is totally bounded} \implies E \text{ is bounded}$$

Proof. Because E is totally bounded, we know there exists a finite set $\{x_1, \dots, x_m\}$ such that

$$E \subseteq \bigcup_{i=1}^m B_1(x_i)$$

Let

$$r := \max_{2 \leq i \leq m} d(a, x_i) + 1$$

We prove such (x_1, r) works. That is

$$E \subseteq B_r(x_1)$$

Let $a \in E$. Because E is covered by $\{B_1(x_i) : 1 \leq i \leq m\}$, we know there exists $k \in \{1, \dots, m\}$ such that

$$a \in B_1(x_k)$$

If $k = 1$, the proof is trivial. If $k > 1$, observe

$$d(a, x_1) \leq d(a, x_k) + d(x_k, x_1) < 1 + d(x_k, x_1) \leq r$$

■

Theorem 3.5.3. (Closure of Bounded Set is Bounded)

$$E \text{ is bounded} \implies \overline{E} \text{ is bounded}$$

Proof. Suppose

$$E \subseteq B_r(y)$$

We prove

$$\overline{E} \subseteq B_{r+1}(y)$$

Let $x \in \overline{E} \setminus E$. Because $x \in \overline{E} \setminus E \implies x \in E'$, we know there exists $z \in E$ such that $d(x, z) < 1$. Then we see

$$d(x, y) \leq d(x, z) + d(y, z) < 1 + r \text{ (done)}$$

■

Theorem 3.5.4. (In \mathbb{R}^n , Bounded implies Totally Bounded) Given $X \subseteq \mathbb{R}^n$

$$X \text{ is bounded} \implies X \text{ is totally bounded}$$

Proof. Suppose

$$X \subseteq B_r(x)$$

Fix ϵ . By Theorem 3.5.3, we know \overline{X} is close and bounded. Then we know \overline{X} is compact. Then we know \overline{X} is totally bounded. Then we see

$$X \subseteq \overline{X} \subseteq \bigcup_{i=1}^m B_\epsilon(x_i)$$

■

Corollary 3.5.5. (In \mathbb{R}^n , Totally Bounded is Equivalent to Bounded) Given $X \subseteq \mathbb{R}^n$

$$X \text{ is bounded} \iff X \text{ is totally bounded}$$

Notice that bounded doesn't implies totally bounded. See Discrete Metric.

3.6 Compactness

There are tools we need to prove the MEGA TFAE.

Theorem 3.6.1. (In Metric Space, Existence of Countable Base is Equivalent to Separable) Let (X, d) be a metric space.

$$X \text{ has a countable base} \iff X \text{ is separable}$$

Proof. Because by Corollary 2.6.5, 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

$$\mathcal{O} := \{B_r(p) : p \in E \text{ and } r \in \mathbb{Q}\}$$

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

$$x \in B_n(x) \subseteq B_r(x) \subseteq O \text{ and } B_n(x) \in \mathcal{O}$$

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

$$x \in B_n(q) \subseteq B_r(x) \subseteq O \text{ and } B_n(q) \in \mathcal{O}$$

and we are done. (done) ■

Lemma 3.6.2. (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

$$d(x, q) \leq d(x, p) + d(p, q) \leq r' + \frac{r}{2} \leq r$$

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

$$2d(p, q) \geq r \implies d(p, q) \leq r - d(p, q) = r'$$

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 3.6.3. (Limit Point Compact Implies Totally Bounded) Let X be a metric space, and let $K \subseteq X$.

$$K \text{ is limit point compact} \implies K \text{ is totally bounded}$$

Proof. Fix ϵ . Let $a_1 \in K$. We only have to construct a finite set of open balls of radius ϵ and centering some points in K

$$\{B_\epsilon(a_1), B_\epsilon(a_2), \dots, B_\epsilon(a_n)\} \text{ where } \{a_1, \dots, a_n\} \subseteq K$$

That cover whole K .

We construct such set by adding $B_\epsilon(a_j)$ such that

$$\forall i < j, a_i \notin B_\epsilon(a_j)$$

Assume **this procedure never ends**. In other words, we assume our method of adding open balls can go on infinitely. Then we have sets

$$A = \{a_i\}_{i \in \mathbb{N}} \subseteq K \text{ and } B_\epsilon(A) := \{B_\epsilon(a_i)\}_{i \in \mathbb{N}}$$

Such that

$$\forall j \in \mathbb{N}, \forall i < j \in \mathbb{N}, a_i \notin B_\epsilon(a_j)$$

Now, we wish to cause a contradiction to the premise that K is limit point compact by showing no point in K is limit point for A .

It is clear that no $\{a_i\}$ is a limit point of A . Let $x \in K \setminus A$. We wish to find an open ball $B_r(x)$ small enough so that $B_r(x)$ contain no a_i .

If x is not contained by no $B_\epsilon(a_i)$, then we see ϵ is small enough to be our r .

If x is contained by some $B_\epsilon(a_i)$, use Lemma 3.9.2, and we are done. **CaC** ■

Theorem 3.6.4. (Limit Point Compact Implies Separable) Let K be a metric space.

$$K \text{ is limit point compact} \implies K \text{ is separable}$$

Proof. By Theorem 3.9.3 (limit point compact implies totally bounded), we know K is totally bounded. This tell us, for each $n \in \mathbb{N}$, there exists a finite set $A_{\frac{1}{n}}$

$$A_{\frac{1}{n}} = \{a_1, a_2, \dots, a_{m_n}\} \text{ s.t. } \forall k \in K, \exists i \leq m_n \in \mathbb{N}, k \in B_{\frac{1}{n}}(a_i)$$

Define

$$A := \bigcup_{n \in \mathbb{N}} A_{\frac{1}{n}}$$

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 3.6.5. (Limit Point Compact Implies Compact)

$$K \text{ is limit point compact} \implies K \text{ is compact}$$

Proof. Let \mathcal{G} be an open cover for K . By Theorem 3.8.4, we know K is separable. Then by Theorem 3.9.1, we know K has a countable base. Then by Theorem 2.6.9, we know \mathcal{G} contain a countable sub-cover

$$\{G_i\}_{i \in \mathbb{N}}$$

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

$$E_n = \bigcup \{G_i\}_{i=1}^n$$

Obviously we have

$$E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$$

We know E_n converge to K , that is

$$\bigcup \{E_i\}_{i=1}^\infty = K$$

Assume $\forall m \in \mathbb{N}, E_m \neq K$. Then we can construct an infinite sequence

$$\{x_i : x_i \in E_i^c\}_{i=1}^\infty$$

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 (done) ■

Theorem 3.6.6. (Limit Point Compact is Equivalent to Sequentially Compact)

$$K \text{ is limit point compact} \iff K \text{ is sequentially compact}$$

Proof. Let $\{x_n\}$ be a sequence in K .

(\longrightarrow)

We wish to prove the existence of a convergent sub-sequence for $\{x_n\}$. If some point appears infinite times in $\{x_n\}$, then the proof is trivial. If not, we see the range $\{x_n : n \in \mathbb{N}\}$ must be infinite. Then because K is limit point compact, we know there exists a limit point x of $\{x_n\}$.

By definition of limit point, we know there exists some x_{n_1} in $B_1(x)$. Because $B_{\frac{1}{2}}(x)$ contain infinite point of $\{x_n\}$, again, we know there exists some n_2 greater than n_1 and $x_{n_2} \in B_{\frac{1}{2}}(x)$.

Again because $B_{\frac{1}{3}}(x)$ contain infinite points of $\{x_n\}$, we know there exists some n_3 greater than n_2 and $x_{n_3} \in B_{\frac{1}{3}}(x)$. Proceeding the procedure infinitely, we get a sub-sequence $\{x_{n_k}\}$ converge to x .

(\longleftarrow)

Let $Y \subseteq K$ be infinite. We wish to prove the existence of some limit point of Y .

Let $\{x_n\}$ be a sequence in Y such that no term repeats. Because K is sequentially compact, we know there exists a sub-sequence $\{x_{n_k}\}$ converge to some point x in K .

If x is in $\{x_{n_k}\}$, we can delete that term, and see that this sequence still converge to x . ■

We now prove the MEGA TFAE.

Theorem 3.6.7. (In Metric Space, Limit Point Compact is Equivalent to Compact) For metric space K , the followings three are equivalent

- (a) K is limit point compact
- (b) K is compact
- (c) K is countably compact
- (d) K is sequentially compact

Above three implies the following two equivalent statement (\mathbb{R} is an example the below 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 3.9.5, 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 3.9.1, we construct a countable base by collecting all open balls of rational radius around each point in the countable dense subset.

In Theorem 3.9.3, 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 3.9.3, we also construct a countable base for K .

In Theorem 2.4.7, 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 3.7.7, 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) by Corollary 3.9.5
- (1) \implies (2) by Theorem 3.9.1
- (a) \implies (1) and (2) by Theorem 3.9.3
- (2) \implies (i) by Theorem 2.4.7
- (a) \implies (b) by Theorem 3.7.7, which use (i)
- (b) \implies (c) by Theorem ??, which is trivial
- (c) \implies (a) by Theorem ??

■

Above is a very important result. One must know how to prove such.
--

3.7 Metric Space: Some Properties Implied by Compact

Notice there is a difference between compact space and compact set

Definition 3.7.1. (Definition of Compact Subset) Let X be a metric space, we say $K \subseteq X$ is a compact subset if

every open cover for K has a finite sub-cover

Theorem 3.7.2. (Compact Subset is itself a Compact Space)

A compact subset regarded as a metric space is compact

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

$$\{G_\lambda : \lambda \in \Lambda\} = \{G_i \cap K : i \in I\}$$

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 3.7.3. (Compact Sets Are Closed) Let X be a metric space, and let $K \subseteq X$

$$K \text{ is compact} \implies K \text{ is closed}$$

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

$$\mathcal{O}' := \{B_{\frac{d(k_i,p)}{2}}(k_i)\}_{i=1}^m$$

Let $r = \min\{\frac{d(k_i,p)}{2} : 1 \leq i \leq m\}$. We only have to prove

the open neighborhood $B_r(p)$ is disjoint to $\bigcup \mathcal{O}'$.

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

$$d(k_i, q) \leq \frac{d(k_i, p)}{2}$$

Then we have

$$d(p, q) + d(q, k_i) < r + \frac{d(k_i, p)}{2} \leq d(k_i, p) \text{ CaC (done)}$$

Definition 3.7.4. (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 3.7.5. (Compact Sets are Bounded) Let X be a metric space and $K \subseteq X$. We have

$$K \text{ is compact} \implies K \text{ is bounded}$$

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

$$K \setminus B_n(q) \neq \emptyset$$

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 3.7.6. (Compact Sets are Closed and Bounded)

$$K \text{ is compact} \implies K \text{ is closed and bounded}$$

We close this section with a method to generate new compact sets. ■

Theorem 3.7.7. (Closed Subsets of Compact Sets Are Compact)

$$K \text{ is compact and } F \subseteq K \text{ is closed} \implies F \text{ is compact}$$

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 3.7.8. 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 ■

3.8 Complete

We now introduce two equivalent definition of bounded.

Definition 3.8.1. (Definition of Bounded) We say a set E is bounded if there exists some open ball $B_r(p)$ that contain E .

Theorem 3.8.2. (Equivalent Definition of Bounded)

$$E \text{ is bounded} \iff \exists r \in \mathbb{R}^+, \forall p, q \in E, d(p, q) < r$$

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 3.8.3. (Converge Implies Bounded)

$$\lim_{n \rightarrow \infty} a_n = a \implies \{a_n\} \text{ is bounded}$$

Proof. Fix r . Because $\{a_n\}$ converge to a , there exists $N_r \in \mathbb{N}$ 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 definition of totally bounded.

Definition 3.8.4. (Definition of Totally Bounded)

3.9 Euclidean Metric

Definition 3.9.1. (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

$$I = \prod_{n=1}^k [a_n, b_n]$$

Lemma 3.9.2. (Shrinking k-cell Sequence Convergence) Let $\{I_n\}_{n=1}^\infty$ be a sequence of k -cell such that

$$\forall n \in \mathbb{N}, I_{n+1} \subseteq I_n$$

Then we have

$$\bigcap \{I_n\}_{n=1}^\infty \neq \emptyset$$

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,

$$\sup\{a_k\}_{k=1}^\infty \in \bigcap_{n=1}^\infty [a_n, b_n]$$

Use the same argument in every other component to get a point in $\bigcap \{I_n\}_{n=1}^\infty$. ■

Lemma 3.9.3. (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

$$|\mathbf{b} - \mathbf{a}| = \left(\sum_{i=1}^k (b_i - a_i)^2 \right)^{\frac{1}{2}}$$

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

$$|\mathbf{x} - \mathbf{y}| = \left(\sum_{i=1}^k |x_i - y_i|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^k (b_i - a_i)^2 \right)^{\frac{1}{2}} = |\mathbf{b} - \mathbf{a}|$$

■

Theorem 3.9.4. (k-Cell is Compact)

Every k-cell I is compact.

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

$$c_n = \frac{a_n + b_n}{2}$$

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

$$z_n = \frac{x_n + y_n}{2}$$

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

$$\cdots \subseteq I_3 \subseteq I_2 \subseteq I_1 \subseteq I_0$$

Denote this sequence $\{I_n\}_{n=1}^\infty$. By Lemma 3.7.7, we know

$$\bigcap \{I_n\}_{n=0}^\infty \neq \emptyset$$

Then, we can let

$$\mathbf{x} \in \bigcap \{I_n\}_{n=0}^\infty$$

Because for each non-negative integer n ,

$$\mathbf{x} \in I_n \text{ and } \{G_\lambda : \lambda \in \Lambda\} \text{ is an open cover for } I_n$$

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

$$\mathbf{x} \in B_r(\mathbf{x}) \subseteq G_\lambda$$

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

$$\mathbf{x} \in I_m \subseteq B_r(\mathbf{x}) \subseteq G_\lambda$$

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

$$\delta = \left(\sum_{j=1}^k (b_j - a_j)^2 \right)^{\frac{1}{2}} = |\mathbf{b} - \mathbf{a}|$$

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 3.9.3, 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 3.9.5. (Heine-Borel Theorem) In $(\mathbb{R}^k, d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|)$, we have

$$K \text{ is closed and bounded } \iff K \text{ is compact}$$

Proof. We have proved from right to left in Theorem 3.9.3 and Theorem 3.9.5. 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. ■

Definition 3.9.6. (Definition of Sequentially Compact) Let $E \subseteq X$ be a topological space. We say E is sequentially compact if

$$\text{for each sequence } \{a_n\} \text{ there exists a converging sub-sequence } \{a'_n\}$$

Notice that in general topology, a sequence can converge to multiple point, the trivial topology of multiple points is an example.	Notic for M
--	----------------

Theorem 3.9.7. (In Metric Space, Limit Point Compact Implies Sequentially Compact)

$$K \text{ is limit point compact } \implies K \text{ is sequentially compact}$$

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

$$B_{\frac{1}{2}}(p) \cap \{p_n : n \in \mathbb{N}\}$$

is infinite, so it is impossible

$$B_{\frac{1}{2}}(p) \cap \{p_n : n \in \mathbb{N}\} \subseteq \{p_n : p \leq k_1 \in \mathbb{N}\}$$

then we know

$$B_{\frac{1}{2}}(p) \cap \{p_n : n > k_1 \in \mathbb{N}\} \neq \emptyset$$

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 3.9.8. (In Euclidean Space, Bounded Implies Existence of Convergent Sub-sequence) In \mathbb{R}^n , we have

$$\{p_n\} \text{ is bounded} \implies \{p_n\} \text{ has a convergent sub-sequence}$$

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 . ■

Theorem 3.9.9. (Characterization of Connected in \mathbb{R}^1 is Interval) Let $I \subseteq \mathbb{R}$.

$$I \text{ is connected} \implies I \text{ is an interval and its endpoint is } \inf I, \sup I$$

Proof. WOLG, suppose $\inf I, \sup I \notin I$. We need to prove

$$I = (\inf I, \sup I)$$

Because $\inf I, \sup I \notin I$, it is clear that

$$I \subseteq (\inf I, \sup I)$$

At this point, one can tell why we can use WOLG. Assume $(\inf I, \sup I) \setminus I \neq \emptyset$. Define $R := (\inf I, \sup I) \setminus I$. Let $x \in R$. Consider

$$\alpha := \sup\{y \in I : y < x\} \text{ and } \beta := \inf\{y \in I : y > x\}$$

We see

$$\alpha \leq x \leq \beta$$

This tell us

$\{y \in I : y < x\}$ and $\{y \in I : y > x\}$ are separated **CaC**



3.10 Cauchy Sequence

Definition 3.10.1. (Definition of Cauchy Sequence) Let X be a metric space. We say $\{p_n\}$ is a Cauchy sequence if

$$\forall \epsilon \in \mathbb{R}^+, \exists N \in \mathbb{N}, \forall k, m > N, d(p_m, p_k) < \epsilon$$

Definition 3.10.2. (Definition of Diameter) Let $E \subseteq X$. We say the diameter of E is

$$\text{diam } E = \sup\{d(p, q) : p, q \in E\}$$

Theorem 3.10.3. (Convergent Implies Cauchy)

$$\{p_n\} \text{ converge} \implies \{p_n\} \text{ is a Cauchy sequence}$$

Proof. Let $\{p_n\}$ converge to p . Fix ϵ , we know there exists $N_{\frac{\epsilon}{2}}$ such that

$$k, m > N_{\frac{\epsilon}{2}} \implies d(p_k, p), d(p_m, p) < \frac{\epsilon}{2} \implies d(p_k, p_m) < d(p_k, p) + d(p_m, p) = \epsilon$$

■

Theorem 3.10.4. (Diameter of Closure Equals to Original Diameter) Let $E \subseteq X$. We have

$$\text{diam } \overline{E} = \text{diam } E$$

Proof. We wish to prove

$$\sup\{d(p, q) : p, q \in \overline{E}\} = \sup\{d(p, q) : p, q \in E\}$$

Because $E \subseteq \overline{E}$, we have

$$\sup\{d(p, q) : p, q \in \overline{E}\} \geq \sup\{d(p, q) : p, q \in E\}$$

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

$$d(p', q') > \text{diam } E$$

Clearly, p', q' can not both be in E . Let $p' \in \overline{E} \setminus E$. We know

$$p' \in E'$$

If $q' \in E$, denote q' by q . We see

$$\forall p \in E, d(p', q) \leq d(p, p') + d(p, q)$$

where $d(p, p') > 0$ can be arbitrarily small. So if

$$d(p, q) < \text{diam } E$$

We have

$$d(p', q) \leq d(p, p') + d(p, q) < \text{diam } E$$

Then we have

$$\forall p \in E, d(p, q) = \text{diam } E$$

Which give us

$$E = \{q\} \text{ CaC}$$

Let $q' \notin E$. We know

$$q' \in E'$$

We see

$$\forall p, q \in E, d(p', q') \leq d(p', p) + d(p, q) + d(q, q')$$

where $d(p', p)$ and $d(q, q')$ can be arbitrarily small.

Let $d(p, q)$ satisfy

$$d(p, q) < \text{diam } E$$

Let $d(p', p)$ and $d(q, q')$ small enough to have

$$d(p', q') \leq d(p', p) + d(p, q) + d(q, q') < \text{diam } E$$

■

Theorem 3.10.5. (Sufficient Condition for Infinite Intersection of Compact Sets being Non-empty) Let $\{K_\lambda : \lambda \in \Lambda\}$ be an infinite collection of compact sets.

every finite sub-collection of $\{K_\lambda : \lambda \in \Lambda\}$ is non-empty $\implies \bigcap_{\lambda \in \Lambda} K_\lambda$ is non-empty

Proof. Assume $\bigcap_{\lambda \in \Lambda} K_\lambda = \emptyset$. Fix K_α . Because

$$K_\alpha \cap \left(\bigcap_{\lambda \in \Lambda, \lambda \neq \alpha} K_\lambda \right) = \emptyset$$

We know no point of K_α belong to every K_λ , otherwise that point is in $K_\alpha \cap \left(\bigcap_{\lambda \in \Lambda, \lambda \neq \alpha} K_\lambda \right)$.

Then we see

$$\{K_\lambda^c : \lambda \in \Lambda\}$$

Is an open cover for K_α .

Because K_α is compact, there exists a finite sub-cover

$$\{K_1^c, \dots, K_m^c\}$$

Notice that for all $x \in K_\alpha$, we have

$$\exists n : 1 \leq n \leq m, x \in K_n^c$$

Where $x \in K_n^c \implies x \notin K_n$. This tell us

no points in K_α belong to all K_1, \dots, K_m

Meaning

$$K_\alpha \cap \bigcap \{K_1, \dots, K_m\} = \emptyset \text{ CaC}$$

■

Corollary 3.10.6. (Single Point) Let $\{K_n\}$ be a sequence of compact sets such that

$$\forall n, K_{n+1} \subseteq K_n \text{ and } \lim_{n \rightarrow \infty} \text{diam } K_n = 0$$

We have

$$\bigcap_{n=1}^{\infty} K_n \text{ consists of exact one point}$$

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

$$\text{diam } \bigcap_{n=1}^{\infty} K_n > 0$$

However, because for all m , we have

$$\bigcap_{n=1}^{\infty} K_n \subseteq K_m$$

We know

$$\forall m, 0 < \text{diam } \bigcap_{n=1}^{\infty} K_n \leq \text{diam } K_m$$

CaC to

$$\lim_{n \rightarrow \infty} \text{diam } K_n = 0$$

■

Definition 3.10.7. (Definition of Completed) We say a subset X of a metric space is complete if

every Cauchy sequence in X converge to some point in X

Theorem 3.10.8. (Compact Implies Complete)

$$K \text{ is compact} \implies K \text{ is complete}$$

Proof. Let $\{p_n\}$ be a Cauchy sequence in K . For each n let

$$E_n := \{p_k : k > n\}$$

We first prove

$$\lim_{n \rightarrow \infty} \text{diam } \overline{E_n} = 0$$

Fix ϵ . We know there exists N such that

$$m, k > N \implies d(p_m, p_k) < \epsilon$$

This implies

$$\epsilon \geq \text{diam } E_N = \text{diam } \overline{E_N} \text{ (done)}$$

Notice that

$$\overline{E_n} \subseteq K \text{ and } \overline{E_n} \text{ is closed} \implies \overline{E_n} \text{ is compact}$$

Then by Corollary 3.10.6, we know there exists a unique p such that

$$\forall n, p \in \overline{E_n}$$

Now, we wish to prove

$$\lim_{n \rightarrow \infty} p_n = p$$

Fix ϵ . Because

$$\lim_{n \rightarrow \infty} \text{diam } \overline{E_n} = 0$$

We know there exists N such that

$$\forall n > N, \text{diam } \overline{E_n} < \epsilon$$

Let $N_1 > N$. We have

$$\text{diam } \overline{E_{N_1}} < \epsilon$$

Notice

$$E_{N_1} = \{p_{N_1+1}, p_{N_1+2}, \dots\} \subseteq \overline{E_{N_1}}$$

This implies

$$\forall n > N_1, d(p_n, p) \leq \text{diam } \overline{E_n} < \epsilon \text{ (done)}$$

■

Corollary 3.10.9. (\mathbb{R}^n is complete)

\mathbb{R}^n is complete

Proof. Let $\{p_n\}$ be a Cauchy sequence in \mathbb{R}^n . We first prove $\{p_n\}$ is bounded.

Again, let

$$E_n := \{p_k : k > n\}$$

Because

$$\lim_{n \rightarrow \infty} \text{diam } \overline{E_n} = 0$$

We know there exists n such that

$$\text{diam } E_n = \text{diam } \overline{E_n} \in \mathbb{R}^+$$

This implies $\{p_n\}$ is bounded. (done)

Pick a k -cell that contain $\{p_n\}$. We know k -cell is compact. By Theorem 3.10.8, we see $\{p_n\}$ must converge to some point in that k -cell. ■

Corollary 3.10.10. (Cauchy Criterion for Series)

$$\sum a_n \text{ converge} \iff \forall \epsilon, \exists N, \forall m > u > N, \left| \sum_{n=u}^m a_n \right| \leq \epsilon$$

Proof. Let

$$s_n = \sum_{k=1}^n a_k$$

We see

$$|s_m - s_u| = \left| \sum_{n=u}^m a_n \right|$$

Because \mathbb{R} is complete, we know

$$s_n \text{ converge} \iff s_n \text{ is Cauchy}$$

Notice s_n is Cauchy means

$$\forall \epsilon, \exists N, \forall m, u > N, |s_m - s_u| < \epsilon$$

■

3.11 Equivalent Definition of Continuity and Two Kinds of Discontinuity for $\mathbb{R} \rightarrow X$

Definition 3.11.1. (Definition of Limit of Function) Let $f : X \rightarrow Y$. We write

$$\lim_{x \rightarrow p} f(x) = q$$

If for all ϵ , there exists δ such that

$$x \in B_\delta(p) \implies f(x) \in B_\epsilon(q)$$

Let $f : \mathbb{R} \rightarrow \mathbb{R}$. The definition is extended to

$$\lim_{x \rightarrow \infty} f(x) = \infty \iff \forall \alpha, \exists M, x > M \implies f(x) > \alpha$$

Let $f : \mathbb{R} \rightarrow Y$. The definition is extended to

$$\lim_{x \rightarrow \infty} f(x) = q \iff \forall \epsilon, \exists M, x > M \implies f(x) \in B_\epsilon(q)$$

Let $f : X \rightarrow \mathbb{R}$. The definition is extended to

$$\lim_{x \rightarrow p} f(x) = \infty \iff \forall \alpha, \exists \delta, x \in B_\delta(p) \implies f(x) > \alpha$$

Theorem 3.11.2. (Sequential Definition of Limit of Function) The following are equivalent

- (a) $\lim_{x \rightarrow p} f(x) = q$
- (b) $p_n \rightarrow p \implies f(p_n) \rightarrow q$

Proof. We first prove

$$\lim_{x \rightarrow p} f(x) = q \implies (b)$$

Arbitrarily pick a sequence $\{p_n\}$ in X that converge to p . Fix ϵ . Because $\lim_{x \rightarrow p} f(x) = q$, we know there exists δ such that

$$p_n \in B_\delta(p) \implies f(p_n) \in B_\epsilon(q)$$

Then because p_n converge to p , we know there exists N such that

$$\forall n > N, p_n \in B_\delta(p)$$

Then we have

$$\forall n > N, f(p_n) \in B_\epsilon(q) \text{ (done)}$$

We now prove

$$(b) \implies \lim_{x \rightarrow p} f(x) = q$$

Assume $\exists \epsilon, \forall \delta, \exists x_\delta \in E \cap B_\delta(p), f(x_\delta) \notin B_\epsilon(q)$. Consider the sequence

$$\{x_{\frac{1}{n}}\}_{n \in \mathbb{N}}$$

This sequence converge to p but we don't have

$$\lim_{n \rightarrow \infty} f(x_{\frac{1}{n}}) = q$$

since $\forall n, f(x_{\frac{1}{n}}) \notin B_\epsilon(q)$ **CaC** (done) ■

Corollary 3.11.3. (Function is Continuous at Limit Point) If $p \in X'$, we have

$$\lim_{x \rightarrow p} f(x) = f(p)$$

Theorem 3.11.4. (Limit of Function is Unique)

$$\lim_{x \rightarrow p} f(x) = q \text{ and } \lim_{x \rightarrow p} f(x) = q' \implies q = q'$$

Proof. Assume $q \neq q'$. Let $\alpha = d(q, q')$. We know $\alpha > 0$. Then we know there exists δ_q and $\delta_{q'}$ such that

$$\begin{aligned} x \in B_{\delta_q}(p) &\implies f(x) \in B_{\frac{\alpha}{2}}(q) \\ \text{and } x \in B_{\delta_{q'}}(p) &\implies f(x) \in B_{\frac{\alpha}{2}}(q') \end{aligned}$$

Because p is an limit point of X , we know there must exists $x \neq p$ such that $x \in B_{\min\{\delta_q, \delta_{q'}\}}(p)$, and that x is in both $B_{\frac{\alpha}{2}}(q)$ and $B_{\frac{\alpha}{2}}(q')$, which is impossible. **CaC** ■

Theorem 3.11.5. (Arithmetic of Limit of \mathbb{R}^n Valued Function) Suppose $f : X \rightarrow \mathbb{R}^n$ and $g : X \rightarrow \mathbb{R}^n$ satisfy

$$\lim_{x \rightarrow p} f(x) = A \text{ and } \lim_{x \rightarrow p} g(x) = B$$

We have

$$(a) \lim_{x \rightarrow p} (f + g)(x) = A + B$$

$$(b) \lim_{x \rightarrow p} (fg)(x) = AB$$

$$(c) \lim_{x \rightarrow p} (f \cdot g)(x) = A \cdot B$$

Proof. Arbitrarily pick a sequence $\{p_n\}$ in X that converge to p . We have

$$\lim_{n \rightarrow \infty} f(p_n) = A \text{ and } \lim_{n \rightarrow \infty} g(p_n) = B$$

Then the result follows from showing each component of the L.H.S converge to R.H.S, which is proved in Theorem 4.1.6 ■

We now introduce continuous.

Definition 3.11.6. (Definition of Function being Continuous at point p in Metric Space) We say a function $f : X \rightarrow Y$ is continuous at p

$$\text{If } \lim_{x \rightarrow p} f(x) = f(p)$$

Definition 3.11.7. (Definition of Continuous) We say a function $f : X \rightarrow Y$ is continuous

$$\text{If } f \text{ is continuous at all } x \in X$$

Theorem 3.11.8. (Equivalent Definition) The following are equivalent

- (a) f is continuous
- (b) For all open $V \subseteq Y$, the set $f^{-1}[V] \subseteq X$ is open
- (c) For all closed $V \subseteq Y$, the set $f^{-1}[V] \subseteq X$ is closed
- (d) For all $A \subseteq X$, we have $f[\overline{A}] \subseteq \overline{f[A]}$

Proof. Theorem 3.11.8 is the special case of Theorem 2.4.4. The only thing we have to prove is

$$f \text{ is continuous} \iff f \text{ is continuous at all } x \in X$$

where the term "continuous" at R.H.S. is in the sense of neighborhood.

(\longrightarrow)

Let W be a neighborhood around $f(x)$. By definition of neighborhood and open in metric space, there exists an open ball $B_\epsilon(f(x))$ centering $f(x)$ small enough to be contained by W . Because f is continuous at x in the sense of metric space, we know either x is an isolated point or we have

$$\lim_{p \rightarrow x} f(p) = f(x)$$

If x is an isolated point, then we can pick a neighborhood V (open ball if you want) around x small enough so that $V \cap X = \{x\}$. Then we see

$$f[V] \subseteq W$$

If we have

$$\lim_{p \rightarrow x} f(p) = f(x)$$

Then we know there exists δ such that

$$p \in B_\delta(x) \implies f(p) \in B_\epsilon(f(x)) \subseteq W$$

Then we see $B_\delta(x) \cap X$ is the neighborhood we want.

(\longleftarrow)

Let $p \in X$. We wish to prove f is continuous at p in the sense of metric space. We don't have to consider when p is an isolated point, so we only need to prove

$$\lim_{x \rightarrow p} f(x) = f(p)$$

Notice that $B_\epsilon(f(p))$ is an neighborhood around p , so we know there exists an neighborhood V around p such that

$$f[V] \subseteq B_\epsilon(f(p))$$

By definition of neighborhood and open in metric space, we know there exists δ small enough so that

$$B_\delta(p) \subseteq V$$

which give us

$$f[B_\delta(p)] \subseteq f[V] \subseteq B_\epsilon(f(p))$$

Then

$$\forall x \in B_\delta(p), f(x) \in f[B_\delta(p)] \subseteq B_\epsilon(f(p))$$

■

An immediate usage: Intermediate Value Theorem.

Theorem 3.11.9. (Intermediate Value Theorem) Suppose $f : X \rightarrow \mathbb{R}$ is continuous and X is connected.

$$y \in (\inf f[\partial X], \sup f[\partial X]) \implies \exists x \in X, f(x) = y$$

Proof. By Theorem 2.4.6 (Continuity map connected set to connected set), and Theorem 3.9.9, we know

$$f[X] \text{ is an interval}$$

Then because

$$\inf f[X] \leq \inf f[\partial X] \leq \sup f[\partial X] \leq \sup f[X]$$

from $y \in (\inf f[\partial X], \sup f[\partial X])$, we can deduce

$$\inf f[X] < y < \sup f[X] \implies y \in f[X]$$

and we are done. ■

We now introduce discontinuity.

Definition 3.11.10. (Definition of Removable Discontinuity) Given $f : I \subseteq \mathbb{R} \rightarrow Y$. We say f has a removable discontinuity at p if there exists $L \neq f(p) \in Y$ such that

$$\lim_{x \rightarrow p} f(x) = L$$

Definition 3.11.11. (Definition of Jump Discontinuity) Suppose $p \in I^\circ$. We say f has a jump discontinuity at p if

$$\lim_{x \rightarrow p^-} f(x) \text{ and } \lim_{x \rightarrow p^+} f(x) \text{ both exists and are not equal}$$

Theorem 3.11.12. (Real Function is Not Jump Discontinuous at Boundary)

$$p \in I \setminus I^\circ \implies f \text{ has no jump discontinuity at } p$$

Definition 3.11.13. (Definition of Essential Discontinuity) Suppose $p \in I^\circ$. We say f has an essential discontinuity at p if

$$\text{one or both of } \lim_{x \rightarrow p^-} f(x), \lim_{x \rightarrow p^+} f(x) \text{ does not exists}$$

Suppose $p \in I \setminus I^\circ$ and, WOLG, $\exists \delta, (p - \delta, p) \subseteq I$. We say f has an essential discontinuity at p if

$$\lim_{x \rightarrow p^-} f(x) \text{ does not exists}$$

Theorem 3.11.14. (Classification of Discontinuity) Let $p \in I$. Exactly one of the following is true

- (a) f is continuous at p
- (b) f has a removable discontinuity at p
- (c) f has a jump discontinuity at p
- (d) f has an essential discontinuity at p

Proof. Recall

$$f \text{ is continuous at } p \iff \lim_{x \rightarrow p} f(x) = f(p)$$

It is clear that f is discontinuous implies the threes below. Suppose f is discontinuous at p . If $p \in I \setminus I^\circ$, because f is discontinuous at p , we know p is not an isolated point, then there must be one side of limit we can discuss. The limit either exists or not, and clearly not both. If $p \in I^\circ$. Either both side of limit exists or not. If the latter, essential. If the former, either they are equal or not. If not equal, jump. If they are equal, notice they are not $f(p)$ because f is not continuous at p , then removable. ■

We now do some exercises.

Theorem 3.11.15. (Example of Discontinuity)

- (a) Define $f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$. Then f is essentially discontinuous every where
- (b) Define $f(x) = \begin{cases} x & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$. Then f is essentially discontinuous every where but $x = 0$.
- (c) Define $f(x) = \begin{cases} x + 2 & \text{if } -3 < x < -2 \\ -x - 2 & \text{if } -2 \leq x < 0 \\ x + 2 & \text{if } 0 \leq x < 1 \end{cases}$. Then f is continuous every where except when $x = 0$, which has a jump discontinuity
- (d) Define $f(x) = \begin{cases} \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$. Then f is continuous every where except when $x = 0$, which has an essential discontinuity.

Proof. For (a), arbitrarily pick $x \in \mathbb{R}$. Let

$$x_n \in \mathbb{Q} \cap (x, x + \frac{1}{n}) \text{ and } y_n \in (\mathbb{R} \setminus \mathbb{Q}) \cap (x, x + \frac{1}{n})$$

It is easy to see $x_n \rightarrow x^+$ and $y_n \rightarrow x^+$ and $f(x_n) \rightarrow 1$ and $f(y_n) \rightarrow 0$. This show

$$\lim_{z \rightarrow x^+} f(z) \text{ D.N.E.}$$

For (b), we first prove f is continuous at $x = 0$. Let $x_n \rightarrow 0$. Fix ϵ . Let N satisfy

$$\forall n > N, |x_n - 0| < \epsilon$$

We see

$$\forall n > N, |f(x_n)| = \begin{cases} |x_n| & \text{if } x_n \text{ is rational} \\ 0 & \text{if } x_n \text{ is irrational} \end{cases} < \epsilon$$

Let $x \neq 0$. We now prove f is discontinuous at x . Fix $\epsilon = |x|$. For all δ , clearly there exists $y \in B_\delta(x)$ such that y is an irrational. Then we see

$$y \in B_\delta(x) \text{ and } f(y) = 0 \notin B_{|x|}(x) = B_\epsilon(x) = B_\epsilon(f(x))$$

For (c), we first prove f has a jump discontinuity at $x = 0$. For all ϵ , we see

$$\begin{aligned} y \in B_{\min\{2, \epsilon\}}(0) \cap (-\infty, 0) &\implies f(y) = -y - 2 \text{ and } 0 < -y < |\min\{2, \epsilon\}| \leq \epsilon \\ &\implies f(y) \in B_\epsilon(-2) \end{aligned}$$

By using $\delta = \min\{2, \epsilon\}$, we have proved

$$\lim_{y \rightarrow 0^-} f(y) = -2$$

With similar method $\delta = \min\{\epsilon, 1\}$, we can prove

$$\lim_{y \rightarrow 0^+} f(y) = 2 \neq -2$$

Notice that at $x = -2$, we can use similar method (both sides use $\min\epsilon, 1$) to show $\lim_{y \rightarrow -2} f(y) = 0 = f(-2)$. At other points, just observe the restrict $(-3, -2), [-2, 0), [0, 1)$ is continuous.

For (d), first notice that the composition of continuous functions is continuous (Theorem 2.4.5). We only have to prove f has an essential discontinuous at $x = 0$.

Observe that $x_n = \frac{1}{\frac{\pi}{2} + n\pi}$ and $y_n = \frac{1}{\frac{3\pi}{2} + n\pi}$ satisfy

$$x_n \rightarrow 0 \text{ and } y_n \rightarrow 0 \text{ and } \forall n, x_n > y_n > 0$$

Yet

$$f(x_n) \rightarrow 1 \text{ and } f(y_n) \rightarrow -1$$

■

Lastly, we do some characterization of monotone function.

Theorem 3.11.16. (Monotone Function Has Only Jump Discontinuity) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be monotone.

$$f \text{ is discontinuous at } x \implies f \text{ has jump discontinuity at } x$$

Proof. WOLG, let f be increasing. We prove

$$\lim_{p \rightarrow x^-} f(p) = \sup_{p < x} f(p) \leq f(x) \leq \inf_{p > x} f(p) = \lim_{p \rightarrow x^+} f(p)$$

Because $p < x < q \implies f(p) \leq f(x) \leq f(q)$, the inequality follows from definition of supremum and infimum. Suppose

$$p_n \rightarrow x \text{ and } \forall n, p_n < x$$

We wish to prove

$$f(p_n) \rightarrow \sup_{p < x} f(p)$$

Fix ϵ . By definition of supremum, we know there exists $q < x$ such that

$$f(q) \in (\sup_{p < x} f(p) - \epsilon, \sup_{p < x} f(p)]$$

Because $p_n \rightarrow x$, we know there exists N such that

$$\forall n > N, q < p_n < x$$

Then we have

$$\forall n > N, \left| f(p_n) - \sup_{p < x} f(p) \right| < \epsilon \text{ (done)}$$

■

Theorem 3.11.17. (Monotone Function Has Countable Discontinuity) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be monotone.

f has countable discontinuity

Proof. Let E be the set of points that f has a discontinuity. Suppose $p < q \in E$. We see

$$\lim_{x \rightarrow p^+} f(x) \leq \lim_{x \rightarrow q^-} f(x)$$

Then we see the set of interval

$$\{(\lim_{x \rightarrow p^-} f(x), \lim_{x \rightarrow p^+} f(x)) : p \in E\} \text{ is pairwise disjoint}$$

Because \mathbb{Q} is dense in \mathbb{R} . For each $p \in E$, we can associate it with a rational number q_p such that

$$q_p \in (\lim_{x \rightarrow p^-} f(x), \lim_{x \rightarrow p^+} f(x))$$

Because the set of intervals is pairwise disjoint, we know the function q_p is one-to-one. The inverse function of q_p show that E is countable. ■

Theorem 3.11.18. (Inverse Function of a Continuous Function on Compact Domain is Continuous) Let X be compact. Given bijective continuous $f : X \rightarrow Y$

$$f^{-1} : Y \rightarrow X \text{ is continuous}$$

Proof. Let V be open in X . We have to show

$$(f^{-1})^{-1}[V] \text{ is open}$$

We know V^c is closed. Then because $V^c \subseteq X$, we know V^c is compact (Theorem 2.4.8). Then by Theorem 2.4.7, we know $f[V^c]$ is compact. Then by Theorem 3.9.3, we know $f[V^c]$ is closed. This tell us

$$(f^{-1})^{-1}[V] = f[V] = (f[V^c])^c \text{ is open}$$

■

Chapter 4

Numerical Sequence and Series

4.1 Special Real Sequence and Series

Definition 4.1.1. (Definition of Extended Real Numbers) We define

$$\begin{aligned}\forall x \in \mathbb{R}, -\infty < x < \infty \\ \forall x \in \mathbb{R}, x + \infty = \infty \text{ and } x - \infty = -\infty \\ \infty + \infty = \infty \text{ and } -\infty - \infty = -\infty \\ \forall x \in \mathbb{R}^+, x(\infty) = \infty \text{ and } x(-\infty) = -\infty \\ 0(\infty) = 0 = 0(-\infty) \\ \forall x \in \mathbb{R}^-, x(\infty) = -\infty \text{ and } x(-\infty) = \infty \\ \forall x \in \mathbb{R}, \frac{x}{\infty} = 0 = \frac{x}{-\infty} \\ (\infty)(\infty) = \infty = (-\infty)(-\infty) \text{ and } (\infty)(-\infty) = -\infty \\ \forall r > 0, \sqrt[r]{\infty} = \infty = |-\infty| = |\infty|\end{aligned}$$

Notice that we don't define $\frac{\infty}{\infty}$ and $\infty - \infty$ and $\sqrt[r]{-\infty}$

Definition 4.1.2. (Definition of Converge to Infinity) We say a sequence $\{x_n\}$ converge to positive infinity ∞ :

$$\lim_{n \rightarrow \infty} x_n = \infty$$

If the sequence $\{x_n\}$ get sufficiently big after some point:

$$\forall M, \exists N, \forall n > N, x_n > M$$

Similarly, we define

$$\lim_{n \rightarrow \infty} x_n = -\infty \iff \forall M, \exists N, \forall n > N, x_n < M$$

Theorem 4.1.3. (Convergent Implies Bounded)

$\{x_n\}$ converge to a real number $x \implies \{x_n\}$ is bounded

Corollary 4.1.4. (Unbounded Implies Existence of a sub-sequence Converge to Infinity)

$\{x_n\}$ is unbounded above $\implies \exists \{x_{f(n)}\}, \lim_{n \rightarrow \infty} x_{f(n)} = \infty$

Proof. Assume $\{x_n\}$ is unbounded. WOLG, let $\sup\{x_n\} = \infty$. Observe for all R , there exists infinite n such that $x_n > R$. Then for each ϵ , it is impossible $x_n < x + \epsilon$ after some finite N . **CaC** ■

Notice that divergent doesn't imply unbounded.

Notice that $\lim_{n \rightarrow \infty} x_n = \infty \implies \{x_n\}$ is unbounded above,

An example for sequence converge to infinity is $x_n = n - \frac{1}{n}$. A non-example is

$$x_n = \begin{cases} n & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases}$$

Notice that we seldom say a series converge to infinity.

Now we introduce some special sequence.

Definition 4.1.5. (Definition of Little-o and Equivalent) We say $b_n = o(a_n)$ if

$$\lim_{n \rightarrow \infty} \frac{b_n}{a_n} = 0$$

And say $a_n \sim b_n$ if

$$\lim_{n \rightarrow \infty} \frac{b_n}{a_n} = 1$$

Notice that interestingly,

$$a_n \sim b_n \iff \lim_{n \rightarrow \infty} b_n - a_n = 0$$

Theorem 4.1.6. (Sequence Arithmetic) Suppose

$$x_n \rightarrow x \in \mathbb{R} \cup \{\infty, -\infty\} \text{ and } y \in \mathbb{R}$$

We have

(a) $-x_n \rightarrow -x$

$$(b) \ x_n + y_n \rightarrow x + y$$

$$(c) \ cx_n \rightarrow cx$$

$$(d) \ c + x_n \rightarrow c + x$$

$$(e) \ x_n y_n \rightarrow xy$$

$$(f) \ x_n^{-1} \rightarrow x^{-1}$$

If the arithmetic on right hand side is possible.

Proof. Fix ϵ . Respectively use

$$(a) \ N(\epsilon)$$

$$(b) \ N(\frac{\epsilon}{2})$$

$$(c) \ N(\frac{\epsilon}{|c|})$$

$$(d) \ N(\epsilon)$$

$$(e) \ \max\{N(\sqrt{\frac{\epsilon}{3}}), N(\frac{\epsilon}{3|x|}), N(\frac{\epsilon}{3|y|})\}$$

$$(f) \ \max\{N(\frac{|x|}{2}), N(\frac{\epsilon|x|^2}{2})\}$$

For $\lim_{n \rightarrow \infty} x_n y_n = xy$, notice

$$\begin{aligned} |x_n y_n - xy| &= |(x_n - x)(y_n - y) + x(y_n - y) + y(x_n - x)| \\ &\leq |x_n - x| \times |y_n - y| + |x| \times |y_n - y| + |y| \times |x_n - x| \end{aligned}$$

For $\lim_{n \rightarrow \infty} x_n^{-1} = x^{-1}$, notice

$$|x - x_n| < \frac{|x|}{2} \implies |x_n| \geq \frac{|x|}{2}$$

Then notice

$$\left| \frac{1}{x_n} - \frac{1}{x} \right| = \left| \frac{x - x_n}{x \times x_n} \right| = \left| \frac{1}{x} \right| \times \left| \frac{1}{x_n} \right| \times |x - x_n| \leq \frac{1}{|x|} \times \frac{2}{|x|} \times |x - x_n|$$

■

Corollary 4.1.7. (Equivalent)

\sim is an equivalent relation.

Corollary 4.1.8. (Sequence Arithmetic) Let $p \in \mathbb{Z}, q \in \mathbb{N}$

$$\begin{aligned}\lim_{n \rightarrow \infty} x_n = x &\implies \lim_{n \rightarrow \infty} x_n^p = x^p \\ \lim_{n \rightarrow \infty} x_n = x \text{ and } x_n \text{ is positive} &\implies \lim_{n \rightarrow \infty} x_n^{\frac{p}{q}} = x^{\frac{p}{q}}\end{aligned}$$

Proof. The first statement follows from Theorem 4.1.6.

To prove the second statement, we only have to prove

$$\lim_{n \rightarrow \infty} \sqrt[q]{x_n} = \sqrt[q]{x}$$

Notice that we have

$$\left| \sqrt[q]{x_n} - \sqrt[q]{x} \right| = \frac{|x_n - x|}{\sum_{i=0}^{q-1} x_n^{\frac{q-1-i}{q}} \times x^{\frac{i}{q}}} = \frac{|x_n - x|}{\sum_{i=0}^{q-1} (x_n^{q-1-i} \times x^i)^{\frac{1}{q}}}$$

We wish to find a lower bound for

$$\sum_{i=0}^{q-1} (x_n^{q-1-i} \times x^i)^{\frac{1}{q}}$$

Notice that for sufficiently large n , we have $x_n > \frac{x}{2}$. Then we know

$$\sum_{i=0}^{q-1} (x_n^{q-1-i} \times x^i)^{\frac{1}{q}} > \sum_{i=0}^{q-1} \left(\left(\frac{1}{2} \right)^{q-1-i} \times x^{q-1} \right)^{\frac{1}{q}}$$

Then we know for sufficiently large n , we have

$$\left| \sqrt[q]{x_n} - \sqrt[q]{x} \right| = \frac{|x_n - x|}{\sum_{i=0}^{q-1} (x_n^{q-1-i} \times x^i)^{\frac{1}{q}}} < \frac{|x_n - x|}{\sum_{i=0}^{q-1} \left(\left(\frac{1}{2} \right)^{q-1-i} \times x^{q-1} \right)^{\frac{1}{q}}}$$

■

Theorem 4.1.9. (Special Sequence) Let $p > 0$ and $q > 1$ and $\alpha \in \mathbb{R}$. We have

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{1}{n^p} &= 0 \\ \lim_{n \rightarrow \infty} \sqrt[n]{p} &= 1 \\ \lim_{n \rightarrow \infty} \sqrt[n]{n} &= 1 \\ \lim_{n \rightarrow \infty} \frac{1}{q^n} &= 0 \\ \lim_{n \rightarrow \infty} \frac{n^\alpha}{q^n} &= 0\end{aligned}$$

Proof. For $\lim_{n \rightarrow \infty} \frac{1}{n^p} = 0$, observe

$$\frac{1}{n^p} < \epsilon \iff n > \sqrt[p]{\frac{1}{\epsilon}}$$

For $\lim_{n \rightarrow \infty} \sqrt[n]{p} = 1$, observe

$$\sqrt[n]{p} - 1 < \epsilon \iff p < (1 + \epsilon)^n \text{ and } 1 - \sqrt[n]{p} < \epsilon \iff (1 - \epsilon)^n < p$$

For $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$, observe

$$(\sqrt[n]{n} - 1 + 1)^n = n \implies \binom{n}{2} (\sqrt[n]{n} - 1)^2 < n \implies \sqrt[n]{n} - 1 < \sqrt{\frac{2}{n-1}}$$

And observe

$$\sqrt{\frac{2}{n-1}} < \epsilon \iff n > \frac{2}{\epsilon^2} + 1$$

For $\lim_{n \rightarrow \infty} \frac{1}{q^n} = 0$, observe

$$\frac{1}{q^n} < \epsilon \iff q^n > \frac{1}{\epsilon}$$

For $\lim_{n \rightarrow \infty} \frac{n^\alpha}{q^n} = 0$. Let $q = 1 + p$. We have

$$\forall k \leq n \in \mathbb{N}, q^n > \binom{n}{k} p^k = n(n-1) \cdots (n-k+1) \frac{p^k}{k!}$$

Notice that if $k < \frac{n}{2}$, we have

$$n - k + 1 > \frac{n}{2} \text{ and } n(n-1) \cdots (n-k+1) \frac{p^k}{k!} > \left(\frac{n}{2}\right)^k \frac{p^k}{k!}$$

Then if we pick $k > \alpha$, for large n ($n > 2k$), we have

$$\frac{n^\alpha}{q^n} < \frac{n^\alpha}{\left(\frac{n}{2}\right)^k \frac{p^k}{k!}} = n^{\alpha-k} \frac{2^k k!}{p^k}$$

We see the right hand side approach to 0. ■

Remark: to prove $\lim_{n \rightarrow \infty} (f(n))^{g(n)}$ converge to sth. One trick is to use binomial theorem so that $\binom{\alpha}{k} (f(n)) < h(n)$, then we get an upper bound $h(n)$ that may be easier to handle.

4.2 Sequence in Normed Space

Theorem 4.2.1. (Limit Arithmetic In Normed Space) Let x, y be in an normed space. Suppose

$$x_n \rightarrow x \text{ and } y_n \rightarrow y$$

We have

(a) $x_n + y_n \rightarrow x + y$

(b) $cx_n \rightarrow cx$

Proof. Use $N(\frac{\epsilon}{2})$ and $N(\frac{\epsilon}{|c|})$. ■

4.3 Comparison Test, Geometric Series and p-Series

We first give an easier way to determine if a series converge

Theorem 4.3.1. (Cauchy Criterion For Series)

$$\sum_{k=1}^{\infty} a_k \text{ converge} \iff \forall \epsilon, \exists N, \forall n > m > N, \left| \sum_{k=m}^n a_k \right| < \epsilon$$

Proof. Notice

$$\sum_{k=1}^{\infty} a_k = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n a_k \right)$$

Because \mathbb{R} is complete, we know

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sum_{k=1}^n a_k \in \mathbb{R} \\ & \iff \left\{ \sum_{k=1}^n a_k \right\}_{n \in \mathbb{N}} \text{ is Cauchy} \\ & \iff \forall \epsilon, \exists N, \forall n > m > N, \left| \sum_{k=m}^n a_k \right| = \left| \sum_{k=1}^n a_k - \sum_{k=1}^m a_k \right| < \epsilon \end{aligned}$$

■

Corollary 4.3.2. (Cauchy Criterion For Series)

$$\sum_{k=1}^{\infty} a_k \text{ diverge} \iff \exists \epsilon, \forall N, \exists n > m \in \mathbb{N}, \left| \sum_{k=m}^n a_k \right| > \epsilon$$

We now introduce absolutely converge.

Definition 4.3.3. (Definition of Absolutely Convergent) We say $\sum_{k=1}^{\infty} a_k$ absolutely converge if

$$\sum_{k=1}^{\infty} |a_k| \text{ converge to a real number}$$

Theorem 4.3.4. (Non-negative Series Converge Implies Absolutely Converge)

Let $\{a_n\}$ be a sequence that eventually become non-negative. We have

$$\sum_{k=1}^{\infty} a_k \text{ converge} \implies \sum_{k=1}^{\infty} a_k \text{ absolutely converge}$$

Theorem 4.3.5. (Absolutely Convergent Implies Convergent)

$$\sum_{k=1}^{\infty} a_k \text{ absolutely converge} \implies \sum_{k=1}^{\infty} a_k \text{ converge}$$

Proof. Notice

$$\forall n > m \in \mathbb{N}, \left| \sum_{k=m}^n a_k \right| \leq \sum_{k=m}^n |a_k| = \left| \sum_{k=m}^n |a_k| \right|$$

Then from $\sum_{k=1}^{\infty} |a_k|$ satisfy Cauchy criterion, we can deduce $\sum_{k=1}^{\infty} a_k$ satisfy Cauchy Criterion. ■

We now introduce comparison test. Notice that in comparison test, the $\{b_n\}, \{c_n\}$ must eventually be non-negative.

Theorem 4.3.6. (Squeeze Theorem) If for large n , we have $x_n < y_n < z_n$, then we have

$$\lim_{n \rightarrow \infty} x_n = \alpha = \lim_{n \rightarrow \infty} z_n \implies \lim_{n \rightarrow \infty} y_n = \alpha$$

Proof. For large n ,

$$\alpha - \epsilon < x_n < y_n < z_n < \alpha + \epsilon$$

Theorem 4.3.7. (Comparison Test) Let $\{c_n\}$ be eventually non-negative. We have

$$\sum_{k=1}^{\infty} c_k \text{ converge and for large } n, |b_n| \leq c_n \implies \sum_{k=1}^{\infty} b_k \text{ absolutely converge}$$

$$\sum_{k=1}^{\infty} a_k \text{ diverge and for large } n, b_n > |a_n| \implies \sum_{k=1}^{\infty} b_k \text{ diverge}$$

Proof. Observe for all large n and m ,

$$\left| \sum_{k=m}^n |b_k| \right| = \sum_{k=m}^n |b_k| \leq \sum_{k=m}^n c_k \leq \left| \sum_{k=m}^n c_k \right|$$

Observe for all n, m

$$\left| \sum_{k=m}^n a_k \right| < \sum_{k=m}^n |a_k| < \sum_{k=m}^n b_k = \left| \sum_{k=m}^n b_k \right|$$

We now introduce geometric series.

Theorem 4.3.8. (Geometric Series)

$$x \in (-1, 1) \implies \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$
$$x \notin (-1, 1) \implies \sum_{n=0}^{\infty} x^n \text{ diverge}$$

Proof. Let

$$s_n := \sum_{k=0}^n x^k$$

Notice

$$(1-x)s_n = \sum_{k=0}^n x^k - \sum_{k=0}^n x^{k+1} = 1 - x^{n+1}$$

This give us

$$s_n = \frac{1 - x^{n+1}}{1 - x}$$

Then we have

$$x \in (-1, 1) \implies \lim_{n \rightarrow \infty} x^{n+1} = 0 \implies \lim_{n \rightarrow \infty} s_n = \frac{1}{1-x}$$

And have

$$x \notin (-1, 1) \implies \lim_{n \rightarrow \infty} x^{n+1} = \infty \text{ or } -\infty \implies \lim_{n \rightarrow \infty} s_n = \frac{-\infty}{1-x} \text{ or } \frac{\infty}{1-x}$$

■

We now introduce monotonic series and important Theorem 4.3.10

Definition 4.3.9. (Monotonic) We say a real sequence $\{x_n\}$ is monotonic if $\{x_n\}$ monotonically increase or monotonically decrease.

$\{x_n\}$ monotonically increase if $\forall n, x_{n+1} \geq x_n$

$\{x_n\}$ monotonically decrease if $\forall n, x_{n+1} \leq x_n$

Theorem 4.3.10. (Monotonic Sequence Converge if and only if Bounded) Let $\{x_n\}$ be a monotonic sequence. We have

$$\{x_n\} \text{ converge} \iff \{x_n\} \text{ is bounded}$$

Proof. Let $\{x_n\}$ monotonically increase. We know $\{x_n\}$ is bounded below by x_1 . We first prove

$$\{x_n\} \text{ is bounded above} \implies \{x_n\} \text{ converge}$$

Notice that

$$\{x_n\} \text{ is bounded above} \implies \sup\{x_n\} \in \mathbb{R}$$

Let $x := \sup\{x_n\}$. We wish to prove

$$\lim_{n \rightarrow \infty} x_n = x$$

Fix ϵ . We know there exists some N such that

$$x_N > x - \epsilon$$

Notice that

$$\forall m > u > N, x_m, x_u \in (x - \epsilon, x) \text{ and } |x_m - x_u| < \epsilon$$

We have proved $\{x_n\}$ is Cauchy. (done)

For $\{x_n\}$ converge $\implies \{x_n\}$ is bounded, look at Theorem 2.1.5.

Let $\{x_n\}$ monotonically decrease. We know $\{-x_n\}$ monotonically 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. ■

Lastly, we introduce p-series.

Lemma 4.3.11. Let $\{a_n\}$ be a positive monotonically decreasing sequence. We have

$$\sum_{n=1}^{\infty} a_n \text{ converge} \iff \sum_{n=0}^{\infty} 2^n a_{2^n} \text{ converge}$$

Proof. Notice that $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=0}^{\infty} 2^n a_{2^n}$ both monotonically increase. We have introduced the question into proving

$$\sum_{n=1}^{\infty} a_n \text{ is bounded above} \iff \sum_{n=0}^{\infty} 2^n a_{2^n} \text{ is bounded above}$$

We first prove

$$\forall n, \sum_{k=0}^n 2^k a_{2^k} \leq M \implies \forall n, \sum_{k=1}^n a_k \leq M$$

Fix n . Let u be the smallest natural such that $2^{u+1} - 1 \geq n$.

Observe

$$M \geq \sum_{k=0}^u 2^k a_{2^k} \geq \sum_{k=0}^u \sum_{r=1}^{2^k} a_{2^k+r-1} = \sum_{k=1}^{2^{u+1}-1} a_k \geq \sum_{k=1}^n a_k \text{ (done)}$$

We now prove

$$\forall n, \sum_{k=1}^n a_k < M \implies \forall n, \sum_{k=0}^n 2^k a_{2^k} < 2M$$

Fix n . Observe

$$\begin{aligned} 2M &\geq 2 \sum_{k=1}^{2^{n+2}-1} a_k = 2 \sum_{k=0}^{n+1} \sum_{r=1}^{2^k} a_{2^k+r-1} \geq 2 \sum_{k=0}^{n+1} \sum_{r=1}^{2^k} a_{2^{k+1}} \\ &= 2 \sum_{k=0}^{n+1} 2^k a_{2^{k+1}} \\ &= \sum_{k=0}^{n+1} 2^{k+1} a_{2^{k+1}} \\ &\geq \sum_{k=0}^n 2^k a_{2^k} \text{ (done)} \end{aligned}$$

■

Corollary 4.3.12. (p-series)

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ converge} \iff p > 1$$

Proof. If $p \leq 0$, observe

$$\lim_{n \rightarrow \infty} \frac{1}{n^p} = \infty \text{ or } 1 \implies \sum_{n=1}^{\infty} \frac{1}{n^p} = \infty$$

If $p > 0$, observe

$$\sum_{n=0}^{\infty} 2^n \frac{1}{(2^n)^p} = \sum_{n=0}^{\infty} 2^{n(1-p)} = \sum_{n=0}^{\infty} (2^{1-p})^n \text{ converge} \iff 2^{1-p} < 1 \iff p > 1$$

■

4.4 Sub-Sequential Limit

Definition 4.4.1. (Definition of Limit Superior)

$$\limsup_{n \rightarrow \infty} x_n := \lim_{n \rightarrow \infty} \sup\{x_k : k \geq n\}$$

Notice that as extended real numbers, $\limsup_{n \rightarrow \infty} x_n$ always exists, since each term exists and the sequence decrease.

The following suggest that if the first term $\sup\{x_k : k \geq 1\}$ equals to infinity, then every term equals to infinity, in spite of the fact the sequence decrease.

Theorem 4.4.2. (Investigation of the Behavior of The Sequence)

$$\sup_{n \in \mathbb{N}} x_n = \infty \iff \forall n, \sup\{x_k : k \geq n\} = \infty \iff \limsup_{n \rightarrow \infty} x_n = \infty$$

Proof. (\longrightarrow)

Assume $\exists n, \sup\{x_k : k \geq n\} < \infty$. We see

$$\sup_{k \in \mathbb{N}} x_k = \max\{x_1, \dots, x_{n-1}, \sup\{x_k : k \geq n\}\} < \infty \text{ CaC}$$

(\longleftarrow)

Notice that

$$\{\sup_{k \geq n} x_k\}_{n \in \mathbb{N}} \text{ decrease}$$

■

Theorem 4.4.3. (Common Sense)

$$x_n \rightarrow \alpha \iff \alpha = \limsup_{n \rightarrow \infty} x_n = \liminf_{n \rightarrow \infty} x_n$$

Proof. (\longrightarrow)

Assume $\limsup_{n \rightarrow \infty} x_n < \alpha$. Because $\{\sup\{x_k : k \geq n\}\}_{n \in \mathbb{N}}$ decrease, there exists m such that

$$\sup_{n > m} x_n < \alpha$$

Because $x_n \rightarrow \alpha$, we can let N satisfy

$$\forall n > N, x_n > \sup_{k \geq m} x_k$$

Let $n > \max\{N, m\}$. We see

$$\sup_{k \geq m} x_k \geq x_n > \sup_{k \geq m} x_k \text{ CaC}$$

(\leftarrow)

Observe that

$$\inf_{k \geq n} x_k \leq x_n \leq \sup_{k \geq n} x_k$$

The result follows from Squeeze Theorem. ■

Corollary 4.4.4. (Common Sense)

$$\limsup_{n \rightarrow \infty} x_n = -\infty \implies x_n \rightarrow -\infty$$

Theorem 4.4.3 is a very basic fact that we shall give another proof later on.

Theorem 4.4.5. (Main Property of Limit Superior) Suppose

$$\liminf_{n \rightarrow \infty} x_n = \alpha < \infty \text{ and } \limsup_{n \rightarrow \infty} x_n = \beta > -\infty$$

For all ϵ , there exists N such that

$$\forall n > N, x_n \in (\alpha - \epsilon, \beta + \epsilon)$$

Proof. We first consider when $\alpha, \beta \in \mathbb{R}$. Because of monotone of $\{\inf_{k \geq n} x_k\}_{n \in \mathbb{N}}$ and $\{\sup_{k \geq n} x_k\}_{n \in \mathbb{N}}$, we know there exists m such that

$$\alpha - \epsilon < \inf_{k \geq m} x_k \leq \sup_{k \geq m} x_k < \beta + \epsilon$$

We now see m work as N , as

$$n > m \implies \alpha - \epsilon < x_n < \beta + \epsilon$$

If $\alpha = -\infty$ or $\beta = \infty$, the theorem hold true by same argument. ■

Notice in the proof for Theorem 4.4.5, the constraint allow us to exclude the situation $x_n \rightarrow \infty$ or $x_n \rightarrow -\infty$. One can observe that if $\alpha = \infty$, Theorem 4.4.5 become false.

Interestingly, Theorem 4.4.5 can be restated in the form

$$x_n \in (\alpha - \epsilon, \beta + \epsilon) \text{ for large } n$$

Theorem 4.4.6. (Main Property of Limit Superior) Suppose

$$\limsup_{n \rightarrow \infty} x_n = \beta \in \mathbb{R}$$

For all ϵ , there exists sub-sequences x_{n_u} such that

$$\forall u, x_{n_u} > \beta - \epsilon$$

Proof. Because $\{\sup_{k \geq n} x_k\}_{n \in \mathbb{N}} \searrow \beta$, for all $m \in \mathbb{N}$, we have

$$\sup_{k \geq m} x_k \geq \beta > \beta - \epsilon$$

Such sequence x_{n_u} can then be recursively constructed by picking

$$n_u = \min\{k \in \mathbb{N} : k \geq n_{u-1} \text{ and } x_k > \beta - \epsilon\}$$

■

Theorem 4.4.7. (Main Property of Limit Superior) Suppose

$$\limsup_{n \rightarrow \infty} x_n = \infty$$

Then there exists a sub-sequence x_{n_k} such that

$$x_{n_k} \nearrow \infty$$

Proof. Let

$$n_k = \min\{n \in \mathbb{N} : n_k > n_{k-1} \text{ and } x_{n_k} > x_{n_{k-1}} + 1\}$$

This is possible because

$$\sup_{u \geq n_{k-1}} x_u = \infty$$

from Theorem 4.4.2.

■

Notice that like Theorem 4.4.5, both Theorem 4.4.6 and Theorem 4.4.15 avoid the null discussion of $\limsup_{n \rightarrow \infty} x_n = -\infty$.

Also, notice that in Theorem 4.4.6, we can not restate it in the form of "for large n ", as we are getting a sub-sequence instead of a tail.

Theorem 4.4.8. (Useful Property of Limit Superior) Given

$$x_n \leq y_n \text{ for large } n$$

We have

$$\limsup_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} y_n \text{ and } \liminf_{n \rightarrow \infty} x_n \leq \liminf_{n \rightarrow \infty} y_n$$

Proof. Let N be large enough such that

$$\forall n > N, x_n \leq y_n$$

We then see

$$\forall n > N, \sup_{k \geq n} x_k \leq \sup_{k \geq n} y_k$$

This give us

$$\limsup_{n \rightarrow \infty} x_k \leq \limsup_{n \rightarrow \infty} y_k$$

The other side is done with similar method. ■

We now introduce another equivalent definition of limit superior.

Theorem 4.4.9. (Set of sub-Sequential Limit is Closed) Let $x_n \in X$. Suppose

E is the set of sub-sequential limit of x_n

Then

E is closed

Proof. Let y_n be a sequence in E and $y_n \rightarrow y$. We have to prove

$$y \in E$$

by finding a sub-sequence x_{n_k} such that

$$x_{n_k} \rightarrow y$$

By Theorem 3.1.9, we can, WOLG, suppose

$$\forall n, d(y_n, y) < \frac{1}{n}$$

Because $y_n \in E$, for each $m \in \mathbb{N}$, we can define a sub-sequence

$$x_{n_{m,k}} \rightarrow y_m \text{ as } k \rightarrow \infty$$

Notice that in our definition, for each fixed m , we have specified $n_{m,k} \nearrow \infty$.

Also, again by Theorem 3.1.9, we can, WOLG, suppose

$$\forall k, d(x_{n_{m,k}}, y_m) < \frac{1}{k}$$

Now, for each $u \in \mathbb{N}$, we recursively define k_u by

$$k_u := \min\{k \in \mathbb{N} : k > u \text{ and } n_{u,k} > n_{u-1, k_{u-1}}\}$$

This is possible because for each fixed u , we have

$$n_{u,k} \nearrow \infty \text{ and } x_{n_{u,k}} \rightarrow y_u \text{ as } k \rightarrow \infty$$

It is easy to see $n_{u,k_u} \nearrow \infty$. We see that $x_{n_{u,k_u}}$ is a sub-sequence. We now prove

$$x_{n_{u,k_u}} \rightarrow y \text{ as } u \rightarrow \infty$$

Fix ϵ . Let $N > 2\epsilon$. Notice that by definition of k_u , we have $k_u > u$. Now we see

$$u > N \implies d(x_{n_{u,k_u}}, y) \leq d(x_{n_{u,k_u}}, y_u) + d(y_u, y) < \frac{1}{k_u} + \frac{1}{u} < \frac{2}{u} < \frac{2}{N} = \epsilon \text{ (done)}$$

■

Theorem 4.4.10. (Equivalent Definition of Limit Superior) Suppose

E is the set of sub-sequential limit of x_n

We have

$$\limsup_{n \rightarrow \infty} x_n = \sup E$$

Proof. Let $\alpha = \limsup_{n \rightarrow \infty} x_n$. If $\alpha \in \mathbb{R}$, because E is closed, we have to first

find a sub-sequence x_{n_k} that converge to α

and show that

every convergent sub-sequence $x_{n_u} \rightarrow x$ satisfy $x \leq \alpha$

Let

$$n_k = \min\{n \in \mathbb{N} : n > n_{k-1} \text{ and } x_n < \alpha + \frac{1}{k}\}$$

This is possible because of the main property of limit superior Theorem 4.4.5. Then we see $x_{n_k} \searrow \alpha$ (done)

Assume $x_{n_u} \rightarrow x$ and $x > \alpha$. This CaC to Theorem 4.4.5. (done)

If $\alpha = -\infty$, we see $x_n \rightarrow -\infty$, then the result follows from Theorem 3.1.8, as the only sub-sequential limit is $-\infty$. If $\alpha = \infty$, result follows from main property of limit superior Theorem 4.4.15. ■

Theorem 4.4.11. (Limit Points Set of Range Are Exactly Set of sub-Sequential Limits) Let E be the set of sub-sequence limits of $\{x_n\}$. We have

$$\{x_n : n \in \mathbb{N}\}' \subseteq E$$

Proof. Let $y \in \{x_n : n \in \mathbb{N}\}'$. We know there exists a sequence $\{x_{f(n)}\}_{n \in \mathbb{N}}$ converge to y , where we are not guarantee $a < b \implies f(a) < f(b)$.

We wish to construct a sub-sequence of $\{x_{f(n)}\}$ that is also a sub-sequence of $\{x_n\}$.

Notice $\forall n, \exists m > n, f(m) > f(n)$ and we are done. ■

Theorem 4.4.12. (Topological Properties of sub-Sequential Limits) Let E be the set of sub-sequential limits of $\{x_n\}$, and let Y be the set of points that appears infinite times in $\{x_n\}$, i.e.

$$Y = \{y : \exists \{x_{n_k}\}, \forall k, y = x_{n_k}\}$$

We have

$$E = \{x_n : n \in \mathbb{N}\}' \cup Y$$

Proof. By Theorem 4.4.15, it is clear

$$\{x_n : n \in \mathbb{N}\}' \cup Y \subseteq E$$

Arbitrarily pick $z \in E$. We wish to prove

$$z \in \{x_n : n \in \mathbb{N}\}' \cup Y$$

In other words, we wish to prove

$$z \notin Y \implies z \in \{x_n : n \in \mathbb{N}\}'$$

We know there exist a sub-sequence $\{x_{f(n)}\}$ such that

$$\lim_{n \rightarrow \infty} x_{f(n)} = z$$

Because $z \notin Y$. We know there a sub-sequence $\{x_{f(g(n))}\}$ such that

$$\forall n, x_{f(g(n))} \neq z$$

Notice that

$$z = \lim_{n \rightarrow \infty} x_{f(n)} = \lim_{n \rightarrow \infty} x_{f(g(n))}$$

And we are done. ■

□

Theorem 4.4.13. (Useful Lemma) Suppose $x \in \mathbb{R}$

$$x_n \rightarrow x \implies \text{existence of } x_{n_k} \nearrow x \text{ or } x_{n_k} \searrow x$$

Proof. If it happens infinite times that $x_n = x$. The proof is trivial. WOLOG, suppose there exists a sub-sequence x_{n_j} such that $\forall j, x_{n_j} < x$. Clearly we have $x_{n_j} \rightarrow x$. Then we can define

$$j_1 = 1 \text{ and } \forall k > 1, j_k := \min\{s \in \mathbb{N} : x_{n_s} > \max_{u < k} x_{n_{j_u}}\}$$

We now prove

$$x_{n_{j_k}} \nearrow x$$

It is clear that $x_{n_{j_k}} \rightarrow x$. We first prove well definiteness.

Suppose j_k exists, we need to prove j_{k+1} exists. In other words, we need to prove there exists $s \in \mathbb{N}$ such that

$$x_{n_s} > \max_{u < k} x_{n_{j_u}}$$

This is true because $x_{n_{j_k}} \rightarrow x$ and $\max_{u < k} x_{n_{j_u}} < x$.

We now prove

$$x_{n_{j_{k-1}}} < x_{n_{j_k}}$$

Notice that j_k satisfy

$$x_{n_{j_k}} > \max_{u < k} x_{n_{j_u}} \geq x_{n_{j_{k-1}}} \text{ (done)}$$

■

Theorem 4.4.14. (Useful Lemma) Given

$$(a) |a_n| \leq |b_n| + c$$

$$(b) b_n \rightarrow 0$$

We have

$$\limsup_{n \rightarrow \infty} |a_n| \leq c$$

Proof. Assume $\limsup_{n \rightarrow \infty} |a_n| > c$. Let

$$\alpha = \frac{\limsup_{n \rightarrow \infty} |a_n| + c}{2}$$

We have

$$c < \alpha < \limsup_{n \rightarrow \infty} |a_n|$$

We then know there exists a sub-sequence a_{n_k} such that

$$\forall k, |a_{n_k}| > \alpha$$

Yet if we let $\epsilon = \alpha - c$. Because $b_{n_k} \rightarrow 0$, we know there exists N such that

$$\forall k > N, |b_{n_k}| < \epsilon$$

We now see

$$\forall k, \alpha < |a_{n_k}| \leq |b_{n_k}| + c < \epsilon + c = \alpha \text{ CaC}$$

■

Theorem 4.4.15. (Useful Lemma) Let $c \in \mathbb{R}^+$. If for all ϵ , we have

$$\limsup_{n \rightarrow \infty} |a_n| \leq \epsilon c$$

Then

$$a_n \rightarrow 0$$

Proof. It is clear that

$$\limsup_{n \rightarrow \infty} |a_n| = 0$$

Notice that

$$\forall n, |a_n| \geq 0 \implies \liminf_{n \rightarrow \infty} |a_n| \geq 0$$

We now have

$$\liminf_{n \rightarrow \infty} |a_n| = 0 = \limsup_{n \rightarrow \infty} |a_n|$$

■

4.5 Euler Number

Lemma 4.5.1. (Bernoulli's Inequality) Given

(a) $x \geq -1$

(b) $r \in \mathbb{N}$

Then

$$(1+x)^r \geq 1+rx$$

Proof. We prove by induction. Notice that base case is trivial.

Suppose

$$(1+x)^r \geq 1+rx$$

Observe

$$\frac{(1+x)^{r+1}}{1+rx+x} = \frac{(1+x)^r}{1+rx} \times \frac{(1+x)(1+rx)}{1+rx+x} = \frac{(1+x)^r}{1+rx} \times \left(1 + \frac{rx^2}{1+rx+x}\right)$$

Notice that $1+x \geq 0$. If $1+rx+x > 0$, then the result follows from seeing

$$\frac{(1+x)^{r+1}}{1+rx+x} \geq 1$$

If $1+rx+x < 0$. Then the proof is trivial, since $(1+x)^r \geq 0$ ■

Theorem 4.5.2. (Existence of Euler Number)

$$\sum_{n=0}^{\infty} \frac{1}{n!} \text{ converge}$$

Proof. Notice

$$\forall n, n! \geq 2^{n-1} \implies \forall n, \frac{1}{n!} \leq \frac{1}{2^{n-1}}$$

Observe

$$\sum_{n=0}^{\infty} \frac{1}{2^{n-1}} = \frac{2}{\frac{1}{2}} = 4$$

Then by comparison test, we are done. ■

Definition 4.5.3. (Definition of Euler Number)

$$e := \sum_{n=0}^{\infty} \frac{1}{n!}$$

Theorem 4.5.4. (Existence of Euler Number)

$$\lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n \text{ converge}$$

Proof. We first prove

$$(1 + \frac{1}{n})^n \text{ monotonically increase}$$

We wish to prove

$$\frac{(1 + \frac{1}{n+1})^{n+1}}{(1 + \frac{1}{n})^n} \geq 1$$

Observe

$$\begin{aligned} \frac{(1 + \frac{1}{n+1})^{n+1}}{(1 + \frac{1}{n})^n} &= \frac{(\frac{n+2}{n+1})^{n+1}}{(\frac{n+1}{n})^n} \\ &= (\frac{n(n+2)}{(n+1)^2})^n (1 + \frac{1}{n+1}) \\ &= (1 - \frac{1}{(n+1)^2})^n (1 + \frac{1}{n+1}) \\ &\geq (1 - \frac{n}{(n+1)^2}) (1 + \frac{1}{n+1}) \\ &= \frac{(n^2 + n + 1)(n + 2)}{(n + 1)^3} \\ &= \frac{n^3 + 3n^2 + 3n + 3}{n^3 + 3n^2 + 3n + 1} > 1 \text{ (done)} \end{aligned}$$

We now prove

$$\forall n, (1 + \frac{1}{n})^n < e$$

Observe

$$\begin{aligned} (1 + \frac{1}{n})^n &= \sum_{k=0}^n \binom{n}{k} \frac{1}{n^k} \\ &= \sum_{k=0}^n \frac{n(n-1) \cdots (n-k+1)}{n^k k!} \\ &= \sum_{k=0}^n \frac{1}{k!} (\frac{n}{n}) (\frac{n-1}{n}) (\frac{n-2}{n}) \cdots (\frac{n-k+1}{n}) \leq \sum_{k=0}^n \frac{1}{k!} < \sum_{n=0}^{\infty} \frac{1}{n!} = e \text{ (done)} \end{aligned}$$



Theorem 4.5.5. (Existence of Euler Number)

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

Proof. Define

$$t_n = \left(1 + \frac{1}{n}\right)^n$$

In Theorem 4.5.4, we have proved

$$t_n \text{ is increasing and } \forall n, t_n \leq e$$

Fix ϵ . Now, we only have to find n such that

$$t_n > e - \epsilon$$

Let

$$m = \min\{n \in \mathbb{N} : \sum_{k=0}^n \frac{1}{k!} > e - \epsilon\}$$

Let $n \geq m$. Observe

$$\begin{aligned} \left(1 + \frac{1}{n}\right)^n &= \sum_{k=0}^n \binom{n}{k} \frac{1}{n^k} \\ &\geq \sum_{k=0}^m \binom{n}{k} \frac{1}{n^k} \\ &= \sum_{k=0}^m \frac{n(n-1) \cdots (n-k+1)}{n^k k!} \end{aligned}$$

This give us

$$\left(1 + \frac{1}{n}\right)^n \geq \sum_{k=0}^m \frac{n(n-1) \cdots (n-k+1)}{n^k k!} \rightarrow \sum_{k=0}^m \frac{1}{k!} > e - \epsilon \text{ as } n \rightarrow \infty \text{ (done)}$$

■

Theorem 4.5.6. (Property of Exponential Function) Given $r \in \mathbb{R}$

$$e^r = \lim_{n \rightarrow \infty} \left(1 + \frac{r}{n}\right)^n$$

Proof. Observe

$$\left(1 + \frac{r}{n}\right)^n = \left(\left(1 + \frac{1}{\frac{n}{r}}\right)^{\frac{n}{r}}\right)^r$$

■

Then

$$\begin{aligned}\lim_{n \rightarrow \infty} \left(1 + \frac{r}{n}\right)^n &= \lim_{n \rightarrow \infty} \left(1 + \frac{1}{\frac{n}{r}}\right)^{\frac{n}{r}r} \\ &= \lim_{u \rightarrow \infty} \left(1 + \frac{1}{u}\right)^u r \\ &= e^r\end{aligned}$$

4.6 Tests for Convergence of Series

Theorem 4.6.1. (Root Test) Given $\sum a_n$. Let

$$\alpha := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$$

Then

(a) $\alpha < 1 \implies \sum a_n$ absolutely converge

(b) $\alpha > 1 \implies \sum a_n$ diverge

Proof. Let

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \alpha < 1$$

We know there exists N such that

$$\forall n > N, \sqrt[n]{|a_n|} \leq \alpha < 1$$

Then we have

$$\forall n > N, a_n \leq |a_n| < \alpha^n$$

Because $\alpha < 1$, we know

$$\sum_{n=1}^{\infty} \alpha^n = \frac{\alpha}{1 - \alpha}$$

Then by Comparison Test, we know $\sum a_n$ converge. If

$$\alpha = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1$$

Let $\{a_{f(n)}\}$ be a sub-sequence such that

$$\forall n, \sqrt[f(n)]{|a_{f(n)}|} > 1$$

Which means

$$\forall n, |a_{f(n)}| > 1^{f(n)} = 1$$

Then it is impossible $\lim_{n \rightarrow \infty} a_n = 0$, let alone $\sum_{n=1}^{\infty} a_n$ converge. ■

Theorem 4.6.2. (Ratio Test) Given series $\sum a_n$. We have

(a) $\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1 \implies \sum a_n$ converge

(b) $\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1 \implies \sum a_n$ diverge

Proof. We first prove

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1 \implies \sum a_n \text{ converge}$$

Let

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \alpha < 1$$

We know there exists N that satisfy

$$\forall n \geq N, \left| \frac{a_{n+1}}{a_n} \right| \leq \alpha < 1$$

This give us

$$\forall n > N, |a_n| \leq \alpha^{n-N} |a_N|$$

In other words,

$$\forall k, |a_{N+k}| \leq \alpha^k |a_N|$$

Then because

$$\alpha < 1 \implies \sum_{k=1}^{\infty} \beta^k |a_{N+k}| = \frac{|a_N| \alpha}{1 - \alpha}$$

By comparison test we know the series

$$\sum_{k=1}^{\infty} a_{N+k} \text{ converges}$$

This tell us

$$\sum a_n = \sum_{k=1}^N a_k + \sum_{k=1}^{\infty} a_{N+k} \text{ converge (done)}$$

We now prove

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1 \implies \sum a_n \text{ diverge}$$

Let

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \beta > 1$$

Then we know there exists N that satisfy

$$\forall n \geq N, \left| \frac{a_{n+1}}{a_n} \right| \geq \beta > 1$$

This give us

$$\forall k, |a_{N+k}| \geq \beta^k |a_N|$$

Then as

$$\beta > 1 \implies \lim_{k \rightarrow \infty} \beta^k |a_N| = \infty$$

We can see it is impossible $\lim_{n \rightarrow \infty} a_n = 0$, let alone $\sum a_n$ converge. (done)

Theorem 4.6.3. (Root Test is Stronger than Ratio Test) For positive sequence $\{a_n\}$, we have

$$\liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \limsup_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

Proof. Clearly,

$$\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \limsup_{n \rightarrow \infty} \sqrt[n]{a_n}$$

We first prove

$$\liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{a_n}$$

Let

$$\alpha = \liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

If $\alpha = -\infty$, the theorem hold true trivially. We only have to consider when $\alpha > -\infty$. Arbitrarily pick positive β smaller than α :

$$\beta < \alpha = \liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

Then we know there exists N such that

$$\forall n \geq N, \frac{a_{n+1}}{a_n} > \beta$$

This implies

$$\forall k, a_{N+k} > \beta^k a_N$$

Then for all $n > N$, we have

$$\sqrt[n]{a_n} > \sqrt[n]{\beta^{n-N} a_N} = \beta \sqrt[n]{\beta^{-N} a_N}$$

Because

$$\lim_{n \rightarrow \infty} \beta \sqrt[n]{\beta^{-N} a_N} = \beta$$

We see

$$\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \geq \beta$$

Then because $\beta < \alpha$ can be arbitrarily close to α , we see

$$\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \geq \alpha \text{ (done)}$$

■

We now introduce Dirichlet's Test

Lemma 4.6.4. (Abel's Formula) Let

$$A_n = \sum_{k=0}^n a_k \text{ and } q \geq p \geq 0$$

We have

$$\sum_{n=p}^q a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$$

Proof.

$$\begin{aligned} \sum_{n=p}^q a_n b_n &= \sum_{n=p}^q (A_n - A_{n-1}) b_n \\ &= \sum_{n=p}^q A_n b_n - \sum_{n=p-1}^{q-1} A_n b_{n+1} \\ &= \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p \end{aligned}$$

■

Theorem 4.6.5. (Dirichlet's Test) Let $A_n = \sum_{k=0}^n a_k$. Given

$$\begin{aligned} b_1 &\geq b_2 \geq b_3 \cdots \\ \forall n, b_n &\geq 0 \\ \{A_n\} &\text{ is bounded} \end{aligned}$$

We have

$$\sum_{n=0}^{\infty} a_n b_n \text{ converge}$$

Proof. Given ϵ , there exists N such that for all $n > N$, we have $b_n < \frac{\epsilon}{2M}$. Let

$$q \geq p \geq N$$

We have

$$\begin{aligned}
\left| \sum_{n=p}^q a_n b_n \right| &= \left| \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p \right| \\
&\leq \sum_{n=p}^{q-1} |A_n (b_n - b_{n+1})| + |A_q b_q| + |A_{p-1} b_p| \\
&= \sum_{n=p}^{q-1} |A_n| |b_n - b_{n+1}| + |A_q b_q| + |A_{p-1} b_p| \\
&= \sum_{n=p}^{q-1} |A_n| (b_n - b_{n+1}) + |A_q| |b_q| + |A_{p-1}| |b_p| \\
&\leq M \left(\sum_{n=p}^{q-1} (b_n - b_{n+1}) + b_q + b_p \right) \\
&= M(b_p - b_q + b_q + b_p) = 2M b_p < 2M \frac{\epsilon}{2M} = \epsilon
\end{aligned}$$

■

Theorem 4.6.6. (Abel's Test) Given

$$\begin{aligned}
&\sum_{n=1}^{\infty} a_n \text{ converges} \\
&\{b_n\} \text{ is monotone and bounded}
\end{aligned}$$

We have

$$\sum_{k=1}^{\infty} a_k b_k \text{ converges}$$

Proof. Define $b'_n := b_n - b$. We have

$$\begin{aligned}
\sum_{k=1}^{\infty} a_k &= \sum_{k=1}^{\infty} a_k (b + b'_k) \\
&= b \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} a_k b'_k
\end{aligned}$$

Where by Theorem 4.6.5 (Dirichlet's test), the series converge.

■

4.7 Cauchy Product and Mertens' Theorem

Definition 4.7.1. (Definition of Cauchy Product) Given two power series $\sum_{n=0}^{\infty} a_n z^n$ and $\sum_{n=0}^{\infty} b_n z^n$. We say the Cauchy product of these two power series is

$$\sum_{n=0}^{\infty} c_n z^n \text{ where } c_n := \sum_{k=0}^n a_k b_{n-k}$$

Theorem 4.7.2. (Example of $\sum_{n=0}^{\infty} c_n$ Diverge When $\sum_{n=0}^{\infty} a_n, \sum_{n=0}^{\infty} b_n$ Both Converge) Let

$$a_n := \frac{(-1)^n}{\sqrt{n+1}} \text{ and } b_n := \frac{(-1)^n}{\sqrt{n+1}}$$

We see

$$\sum_{n=0}^{\infty} a_n \text{ and } \sum_{n=0}^{\infty} b_n \text{ converge}$$

Yet

$$\sum_{n=0}^{\infty} c_n \text{ diverge}$$

Proof. It is easy to see $\sum_{n=0}^{\infty} a_n$ converge by Alternating Series Test. We prove

$$|c_n| \not\rightarrow 0$$

We now compute c_n

$$\begin{aligned} c_n &= \sum_{k=0}^n a_k b_{n-k} \\ &= \sum_{k=0}^n \frac{(-1)^k}{\sqrt{k+1}} \times \frac{(-1)^{n-k}}{\sqrt{n-k+1}} \\ &= (-1)^n \sum_{k=0}^n \frac{1}{\sqrt{(k+1)(n-k+1)}} \end{aligned}$$

Observe

$$(n-k+1)(k+1) = \left(\frac{n}{2} + 1\right)^2 - \left(\frac{n}{2} - k\right)^2 \leq \left(\frac{n}{2} + 1\right)^2$$

Notice that one can verify $\left(\frac{n}{2} + 1\right)^2 - \left(\frac{n}{2} - k\right)^2$ is always positive. The observation give us

$$\sqrt{(n-k+1)(k+1)} < \left(\frac{n}{2} + 1\right)$$

We now have

$$|c_n| = \sum_{k=0}^n \frac{1}{\sqrt{(k+1)(n-k+1)}} \geq \sum_{k=0}^n \frac{1}{\frac{n}{2} + 1} = \frac{2n+2}{n+2} \text{ (done)}$$

■

Theorem 4.7.3. (Merten's Theorem) Given

(a) $\sum_{n=0}^{\infty} a_n$ converges absolutely.

(b) $\sum_{n=0}^{\infty} a_n = A$

(c) $\sum_{n=0}^{\infty} b_n = B$

When $z = 1$, the Cauchy product of the two power series $\sum_{n=0}^{\infty} a_n z^n$ and $\sum_{n=0}^{\infty} b_n z^n$ converge to AB

$$\sum_{n=0}^{\infty} c_n = AB$$

Proof. Define

(a) $A_n := \sum_{k=0}^n a_k$

(b) $B_n := \sum_{k=0}^n b_k$

(c) $C_n := \sum_{k=0}^n c_k$

(d) $\gamma_n := C_n - A_n B$

(e) $\beta_n := B_n - B$

In this setting, because $A_n \rightarrow A$, to show $C_n \rightarrow AB$, we only have to show

$$\gamma_n \rightarrow 0$$

We first compute

$$\begin{aligned}
C_n &= \sum_{u=0}^n c_u \\
&= \sum_{u=0}^n \sum_{k=0}^u a_k b_{u-k} \\
&= \sum_{0 \leq k \leq u \leq n} a_k b_{u-k} \\
&= \sum_{k=0}^n a_k \sum_{u=k}^n b_{u-k} \\
&= \sum_{k=0}^n a_k \sum_{s=0}^{n-k} b_s \\
&= \sum_{k=0}^n a_k B_{n-k} \\
&= \sum_{k=0}^n a_k (B + \beta_{n-k}) \\
&= A_n B + \sum_{k=0}^n a_k \beta_{n-k}
\end{aligned}$$

Then we have

$$\gamma_n = \sum_{k=0}^n \beta_k a_{n-k}$$

Let

$$\alpha = \sum_{n=0}^{\infty} |a_n|$$

By Theorem 4.4.15, Fix ϵ . It remains to show

$$\limsup_{n \rightarrow \infty} |\gamma_n| \leq \epsilon \alpha$$

Notice that $\beta_k \rightarrow 0$. Then we can choose N such that $\forall k > N, |\beta_k| \leq \epsilon$. Let $n > N$. We

now have

$$\begin{aligned}
|\gamma_n| &:= \left| \sum_{k=0}^n \beta_k a_{n-k} \right| \\
&\leq \left| \sum_{k=0}^N \beta_k a_{n-k} \right| + \left| \sum_{k=N+1}^n \beta_k a_{n-k} \right| \\
&\leq \left| \sum_{k=0}^N \beta_k a_{n-k} \right| + \sum_{k=N+1}^n |\beta_k| |a_{n-k}| \\
&\leq \left| \sum_{k=0}^N \beta_k a_{n-k} \right| + \sum_{k=N+1}^n \epsilon |a_{n-k}| \\
&\leq \left| \sum_{k=0}^N \beta_k a_{n-k} \right| + \epsilon \alpha
\end{aligned}$$

We only have to prove

$$\sum_{k=0}^N \beta_k a_{n-k} \rightarrow 0 \text{ as } n \rightarrow \infty$$

This is obvious, since there are only N terms and we have $a_n \rightarrow 0$ (done) . ■

4.8 Reimann Rearrangement Theorem

Definition 4.8.1. (Permutation) We say σ is a permutation of X if $\sigma : X \rightarrow X$ is bijective.

Definition 4.8.2. (Converge Conditionally) We say a series $\sum_{n=0}^{\infty} a_n$ converge conditionally if

$$\sum_{n=0}^{\infty} a_n \text{ converge and } \sum_{n=0}^{\infty} |a_n| \text{ diverge}$$

Theorem 4.8.3. (Reimann Rearrangement Theorem) If $\sum_{n=0}^{\infty} a_n$ converge conditionally, then for all $M \in \mathbb{R}$, there exists a permutation σ_M such that

$$\sum_{n=0}^{\infty} a_{\sigma_M(n)} = M$$

There also exists permutation $\sigma_{\infty}, \sigma_{-\infty}, \sigma'$ such that

$$\sum_{n=0}^{\infty} a_{\sigma_{\infty}(n)} = \infty \text{ and } \sum_{n=0}^{\infty} a_{\sigma_{-\infty}(n)} = -\infty \text{ and } \sum_{n=0}^{\infty} a_{\sigma'(n)} \text{ diverge}$$

Proof. Define

$$a_n^+ := \max\{a_n, 0\} \text{ and } a_n^- := \min\{a_n, 0\}$$

Notice that

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} a_n^+ + a_n^- \text{ and } \sum_{n=0}^{\infty} |a_n| = \sum_{n=0}^{\infty} a_n^+ - a_n^-$$

Then because $\sum_{n=0}^{\infty} a_n$ converge conditionally, we know

$$\sum_{n=0}^{\infty} a_n^+ \text{ diverge and } \sum_{n=0}^{\infty} a_n^- \text{ diverge}$$

Then because their positiveness, we know

$$\sum_{n=0}^{\infty} a_n^+ = \infty \text{ and } \sum_{n=0}^{\infty} a_n^- = -\infty$$

Let p_1 be the smallest natural such that

$$\sum_{n=0}^{p_1} a_n^+ > M$$

Let p_2 be the smallest natural such that

$$\sum_{n=0}^{p_1} a_n^+ + \sum_{n=0}^{q_1} a_n^- < M$$

Let p_2 be the smallest natural such that

$$\sum_{n=0}^{p_2} a_n^+ + \sum_{n=0}^{q_1} a_n^- = \sum_{n=0}^{p_1} a_n^+ + \sum_{n=0}^{q_1} a_n^- + \sum_{n=p_1+1}^{p-2} a_n^+ > M$$

This is possible because $\sum_{n=p_1+1}^{\infty} a_n^+ = \infty$. Recursively do such. Define $p_0 = -1 = q_0$. Define

$$\alpha_n := |\{k \in \mathbb{N} : p_{n-1} < k \leq p_n, a_k > 0\}|$$

and define

$$\beta_n := |\{k \in \mathbb{N} : q_{n-1} < k \leq q_n, a_k < 0\}|$$

Notice that we don't have to consider when $a_k = 0$, since we can add finite term or a sub-sequence of 0 to make σ actually a permutation.

We now recursively define σ by

$$a_0^+, \dots, a_{p_1}^+, a_0^-, \dots, a_{q_1}^-$$

■

Chapter 5

Calculus and Relevant Notions

5.1 Derivative

Definition 5.1.1. (Definition of Fréchet derivative) Given two normed vector space V, W , an open set $S \subseteq V$ and a function $f : S \subseteq V \rightarrow W$, we say f is differentiable at $\mathbf{x} \in S$ if there exists a bounded linear operator $A : V \rightarrow W$ such that

$$\frac{\|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - A(\mathbf{h})\|_W}{\|\mathbf{h}\|_V} \rightarrow 0 \text{ as } \|\mathbf{h}\|_V \rightarrow 0$$

We say A is the differential or derivative of f at \mathbf{x}

Normally, we write the derivative of f at x by $\partial f(x)$

Theorem 5.1.2. (Equivalent Definition of Fréchet derivative) The following are equivalent

- (a) A is the derivative of f at x
- (b) $\frac{f(x+h)-f(x)}{h} = f'(x) + o(h)$

Proof. Notice

$$f(x+h) - f(x) - Ah = \|h\|_V o(h)$$

■

Theorem 5.1.3. (Uniqueness of Fréchet derivative)

$$A, B \text{ are both derivative of } f \text{ at } \mathbf{x} \implies A = B$$

Proof. Notice that the definition can be re written into

$$\frac{f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - A(\mathbf{h})}{\|\mathbf{h}\|_V} \rightarrow 0$$

Then by Theorem 4.2.1, we have

$$\frac{(A - B)\mathbf{h}}{\|\mathbf{h}\|_V} \rightarrow 0$$

Assume $\exists \mathbf{v}, A\mathbf{v} \neq B\mathbf{v}$. Notice

$$\frac{(A - B)\frac{\mathbf{v}}{c}}{\|\frac{\mathbf{v}}{c}\|_V} \rightarrow \frac{(A - B)(\mathbf{v})}{\|\mathbf{v}\|_V} \neq 0 \text{ as } \frac{\mathbf{v}}{c} \rightarrow 0 \quad \text{CaC}$$

■

Theorem 5.1.4. (Addition of Fréchet derivative) Given two functions $f, g : S \subseteq V \rightarrow W$ that is differentiable at \mathbf{x} with derivative A and B

The derivative of $f + g$ at \mathbf{x} is $A + B$

Proof. Observe

$$\begin{aligned} & \frac{\|(f + g)(\mathbf{x} + \mathbf{h}) - (f + g)\mathbf{x} - (A + B)\mathbf{h}\|_W}{\|\mathbf{h}\|_V} \\ & \leq \frac{\|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - A\mathbf{h}\|_W}{\|\mathbf{h}\|_V} + \frac{\|g(\mathbf{x} + \mathbf{h}) - g(\mathbf{x}) - B\mathbf{h}\|_W}{\|\mathbf{h}\|_V} \rightarrow 0 \end{aligned}$$

■

Theorem 5.1.5. (Chain Rule in Normed Space Space)

$$\partial(g \circ f)(x_0) = \partial g(f(x_0)) \circ \partial f(x_0)$$

Proof. Denote

- (a) $y = f(x)$
- (b) $\tilde{f}(x) = f(x + x_0)$
- (c) $\hat{f}(x) = \tilde{f}(x) - \tilde{f}(0)$
- (d) $\tilde{g}(y) = g(y + \tilde{f}(0))$

We first prove that if the chain rule hold for $x = 0$, it hold true for all $x \in X$.

Because

$$\frac{g \circ f(x_0 + h) - g \circ f(x_0) - \partial(g \circ f)x_0}{\|h\|_V} \rightarrow 0$$

and

$$\frac{g \circ \tilde{f}(h) - g \circ \tilde{f}(0) - \partial(g \circ \tilde{f})(0)}{\|h\|_V} \rightarrow 0$$

and

$$\tilde{f}(h) = f(x_0 + h) \text{ and } \tilde{f}(0) = f(x_0)$$

we know

$$\partial(g \circ f)(x_0) = \partial(g \circ \tilde{f})(0)$$

Then we can compute

$$\partial(g \circ f)(x_0) = \partial(g \circ \tilde{f})(0)$$

■

Theorem 5.1.6. (Differentiable Implies Continuous) Given $f : \mathbb{R} \rightarrow \mathbb{R}$

$$f \text{ is differentiable at } x \implies f \text{ is continuous at } x$$

Proof. Assume $x_n \rightarrow x$ and $f(x_n) \rightarrow y \neq x$.

$$\frac{f(x_n) - f(x)}{x_n - x} \rightarrow \pm\infty \text{ CaC}$$

■

5.2 Discontinuity and Convex

In this section, λ^c is defined to be $1 - \lambda$.

Definition 5.2.1. (Definition of Line Segment) Suppose $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. We define the line segment \overline{AB} by

$$\overline{AB} = \{\lambda \mathbf{a} + \lambda^c \mathbf{b} \in \mathbb{R}^n : \lambda \in [0, 1]\}$$

Definition 5.2.2. (Definition of Convex Set) We say a set X in \mathbb{R}^n is convex if

$$\forall \mathbf{a}, \mathbf{b} \in X, \overline{AB} \subseteq X$$

Theorem 5.2.3. (Convex Set is Path Connected)

$$X \text{ is convex} \implies X \text{ is path-connected}$$

Proof. The function $f : [0, 1] \rightarrow X$ defined by

$$f(\lambda) = \lambda \mathbf{a} + \lambda^c \mathbf{b}$$

is clearly a path from a to b , as we can prove it is continuous at each component. ■

Definition 5.2.4. (Definition of Convex Function) Let $X \subseteq \mathbb{R}^n$ be convex. We say a function $f : X \rightarrow \mathbb{R}$ is convex if

$$\forall \mathbf{a}, \mathbf{b} \in X, \forall \lambda \in [0, 1], f(\lambda \mathbf{a} + \lambda^c \mathbf{b}) \leq \lambda f(\mathbf{a}) + \lambda^c f(\mathbf{b})$$

Definition 5.2.5. (Definition of Mid-Point Convex) Let $X \subseteq \mathbb{R}^n$ be convex. We say $f : X \rightarrow \mathbb{R}$ is mid-point convex if

$$\forall \mathbf{a}, \mathbf{b} \in X, f\left(\frac{\mathbf{a} + \mathbf{b}}{2}\right) \leq \frac{f(\mathbf{a}) + f(\mathbf{b})}{2}$$

Theorem 5.2.6. (Equivalent Definition of Convex Function) Given a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$. The following are equivalent

- (a) f is convex
- (b) The epigraph $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in X \text{ and } y \geq f(\mathbf{x})\}$ is convex

Proof. We now prove (a) \implies (b).

Denote

$$Y = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in X \text{ and } y \geq f(\mathbf{x})\}$$

Assume Y is not convex. That is, there exists two points $(\mathbf{x}, y_x), (\mathbf{z}, y_z) \in Y$ such that

$$\{\lambda(\mathbf{x}, y_x) + \lambda^c(\mathbf{z}, y_z) \in \mathbb{R}^{n+1} : \lambda \in [0, 1]\} \not\subseteq Y$$

Fix the $\lambda + \lambda^c = 1$ such that

$$(\lambda\mathbf{x} + \lambda^c\mathbf{z}, \lambda y_x + \lambda^c y_z) = \lambda(\mathbf{x}, y_x) + \lambda^c(\mathbf{z}, y_z) \notin Y$$

Because X is convex, we know

$$\lambda\mathbf{x} + \lambda^c\mathbf{z} \in X$$

Then because $(\lambda\mathbf{x} + \lambda^c\mathbf{z}, \lambda y_x + \lambda^c y_z) \notin Y$, by definition of Y , we can deduce

$$\lambda f(\mathbf{x}) + \lambda^c f(y_z) \leq \lambda y_x + \lambda^c y_z < f(\lambda\mathbf{x} + \lambda^c\mathbf{z}) \quad \text{CaC } f \text{ is convex (done)}$$

We now prove $(b) \implies (a)$. We first prove

■

Look at the following figure.

Theorem 5.2.7. (Mid Point Convex and Continuity Implies Convex)

$$f \text{ is mid point convex and continuous} \implies f \text{ is convex}$$

Proof. Assume f is not convex. Then there exists \mathbf{x}, \mathbf{y} such that if we define $g : [0, 1] \rightarrow \mathbb{R}$ by

$$g(\lambda) := f(\lambda\mathbf{x} + \lambda^c\mathbf{y}) - (\lambda f(\mathbf{x}) + \lambda^c f(\mathbf{y}))$$

there exists some λ such that $g(\lambda) > 0$. Notice that g is continuous and $[0, 1]$ is compact, so we know there exists maximum

$$\max_{\lambda \in [0, 1]} g(\lambda) = M$$

Again, because g is continuous, we then know there exists a minimum λ_0 such that

$$\lambda_0 = \min\{\lambda \in [0, 1] : g(\lambda) = M\}$$

Let δ be small enough such that

$$\lambda_0 \in (\lambda_0 - \delta, \lambda_0 + \delta) \subseteq (0, 1)$$

Define

$$\bar{\mathbf{x}} := (\lambda_0 - \delta)\mathbf{x} + (\lambda_0^c + \delta)\mathbf{y}$$

$$\bar{\mathbf{y}} := (\lambda_0 + \delta)\mathbf{x} + (\lambda_0^c - \delta)\mathbf{y}$$

It is easy to check $\bar{\mathbf{x}}, \bar{\mathbf{y}}$ are in the geodesic from \mathbf{x} to \mathbf{y} . Then because f is mid point convex, we can deduce

$$f(\lambda_0^c \mathbf{x} + \lambda_0 \mathbf{y}) = f\left(\frac{\bar{\mathbf{x}} + \bar{\mathbf{y}}}{2}\right) \leq \frac{f(\bar{\mathbf{x}}) + f(\bar{\mathbf{y}})}{2}$$

Observe

$$\begin{aligned} f(\bar{\mathbf{x}}) &= f((\lambda_0 - \delta)\mathbf{x} + (\lambda_0^c + \delta)\mathbf{y}) \\ &= g(\lambda_0 - \delta) + (\lambda_0 - \delta)f(\mathbf{x}) + (\lambda_0^c + \delta)f(\mathbf{y}) \end{aligned}$$

and

$$\begin{aligned} f(\bar{\mathbf{y}}) &= f((\lambda_0 + \delta)\mathbf{x} + (\lambda_0^c - \delta)\mathbf{y}) \\ &= g(\lambda_0 + \delta) + (\lambda_0 + \delta)f(\mathbf{x}) + (\lambda_0^c - \delta)f(\mathbf{y}) \end{aligned}$$

This tell us

$$\frac{f(\bar{\mathbf{x}}) + f(\bar{\mathbf{y}})}{2} = \frac{g(\lambda_0 - \delta) + g(\lambda_0 + \delta)}{2} + \lambda_0 f(\mathbf{x}) + \lambda_0^c f(\mathbf{y})$$

Also we compute

$$\begin{aligned} f\left(\frac{\bar{\mathbf{x}} + \bar{\mathbf{y}}}{2}\right) &= f(\lambda_0 \mathbf{x} + \lambda_0^c \mathbf{y}) \\ &= g(\lambda_0) + \lambda_0 f(\mathbf{x}) + \lambda_0^c f(\mathbf{y}) \end{aligned}$$

Then because f is mid point convex, we can deduce

$$\begin{aligned} g(\lambda_0) + \lambda_0 f(\mathbf{x}) + \lambda_0^c f(\mathbf{y}) &= f\left(\frac{\bar{\mathbf{x}} + \bar{\mathbf{y}}}{2}\right) \\ &\leq \frac{f(\bar{\mathbf{x}}) + f(\bar{\mathbf{y}})}{2} = \frac{g(\lambda_0 - \delta) + g(\lambda_0 + \delta)}{2} + \lambda_0 f(\mathbf{x}) + \lambda_0^c f(\mathbf{y}) \end{aligned}$$

This give us

$$M = g(\lambda_0) \leq \frac{g(\lambda_0 - \delta) + g(\lambda_0 + \delta)}{2}$$

which is impossible **CaC**, since M is the greatest of $g(\lambda)$, where $g(\lambda_0 - \delta) < M$ and g is continuous. ■

Lemma 5.2.8. Given

- (a) f is convex on (a, b)
- (b) $a < s < t < u < b$

We have

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t}$$

Proof. We first prove

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s}$$

Observe

$$\begin{aligned} \frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} &\iff f(t) - f(s) \leq \frac{t - s}{u - s} f(u) - \frac{t - s}{u - s} f(s) \\ &\iff f(t) \leq \frac{t - s}{u - s} f(u) + \frac{u - t}{u - s} f(s) \end{aligned}$$

Notice that

$$\frac{t - s}{u - s} + \frac{u - t}{u - s} = \frac{u - s}{u - s} = 1 \text{ and } \frac{t - s}{u - s} \in (0, 1)$$

The result then follow from definition of convex (done)

We now prove

$$\frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t}$$

Observe

$$\begin{aligned} \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t} &\iff \frac{u - t}{u - s} f(u) - \frac{u - t}{u - s} f(s) \leq f(u) - f(t) \\ &\iff f(t) \leq \left(1 - \frac{u - t}{u - s}\right) f(u) + \frac{u - t}{u - s} f(s) \end{aligned}$$

Then because $\frac{u-t}{u-s} \in (0, 1)$, the result again follows from definition of convex (done) ■

Theorem 5.2.9. (Convex Function is Continuous on Open Interval) Given $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$

$$f \text{ is convex on } (a, b) \implies f \text{ is continuous on } (a, b)$$

Proof. Let $y \in (a, b)$. We wish to prove

$$\lim_{x \rightarrow y^-} f(x) = f(y) = \lim_{x \rightarrow y^+} f(x)$$

Let

$$0 < x < y < z < 1$$

For small δ , by Lemma 5.2.8, we have

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(y + \delta) - f(y)}{\delta} \leq \frac{f(z) - f(y)}{z - y}$$

Then

$$f(y + \delta) - f(y) \in \left[\frac{f(y) - f(x)}{y - x} \delta, \frac{f(z) - f(y)}{z - y} \delta \right] \rightarrow [0, 0] \text{ as } \delta \rightarrow 0$$

This shows that $\lim_{x \rightarrow y^+} f(x) = f(y)$. Again, for small δ , by Lemma 5.2.8, we have

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(y - \delta)}{\delta} \leq \frac{f(z) - f(y)}{z - y}$$

Then again for the same reason, $f(y) - f(y - \delta) \rightarrow 0$ (done) . ■

5.3 Uniformly Continuous

Definition 5.3.1. (Uniformly Continuous) We say a function $f : X \rightarrow Y$ is uniformly continuous if

$$\forall \epsilon, \exists \delta, d(p, q) < \delta \implies d(f(p), f(q)) < \epsilon$$

Theorem 5.3.2. (Sequential Definition of Uniform Continuous) The following are equivalent

(a) f is uniformly continuous

(b) $\lim_{n \rightarrow \infty} d(p_n, q_n) = 0 \implies \lim_{n \rightarrow \infty} d(f(p_n), f(q_n)) = 0$

Proof. We first prove

$$f \text{ is uniformly continuous} \implies \left(\lim_{n \rightarrow \infty} d(p_n, q_n) = 0 \implies \lim_{n \rightarrow \infty} d(f(p_n), f(q_n)) = 0 \right)$$

Fix ϵ . Let δ be from uniform continuity of f . Because $\lim_{n \rightarrow \infty} d(p_n, q_n) = 0$, we know there exists N such that

$$\forall n > N, d(p_n, q_n) < \delta$$

This implies

$$\forall n > N, d(f(p_n), f(q_n)) < \epsilon \text{ (done)}$$

We now prove

$$\left(\lim_{n \rightarrow \infty} d(p_n, q_n) = 0 \implies \lim_{n \rightarrow \infty} d(f(p_n), f(q_n)) = 0 \right) \implies f \text{ is uniformly continuous}$$

Assume f is not uniformly continuous. Fix ϵ . We know there exists a sequence

$$\lim_{n \rightarrow \infty} d(p_n, q_n)$$

such that

$$\forall n, d(f(p_n), f(q_n)) > \epsilon \text{ CaC (done)}$$

■

Theorem 5.3.3. (Uniformly Continuous is Stronger than Continuous) Given $f : X \rightarrow Y$

$$f \text{ is uniformly continuous} \implies f \text{ is continuous}$$

Proof. Arbitrarily pick $x \in X$ and $\epsilon \in \mathbb{R}^+$. We wish to prove there exists δ such that

$$p \in B_\delta(x) \implies f(p) \in B_\epsilon(f(x))$$

Because f is uniformly continuous, we know there exists δ' such that

$$d(p, q) < \delta' \implies d(f(p), f(q)) < \epsilon$$

Then we can deduce

$$p \in B_{\delta'}(x) \implies d(p, x) < \delta' \implies d(f(p), f(x)) < \epsilon \implies f(p) \in B_\epsilon(f(x))$$

We have proved δ' is the desired δ . ■

Theorem 5.3.4. (Uniformly Continuous is Strictly Stronger than Continuous)

Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f(x) = x^2$$

We have

f is continuous but not uniformly continuous

Proof. Let $a \in \mathbb{R}$. We wish to prove f is continuous at a . Fix ϵ . Observe

$$\begin{aligned} |x^2 - a^2| &= |x^2 - ax + ax - a^2| \\ &\leq |x^2 - ax| + |ax - a^2| \\ &= |x| \times |x - a| + |a| \times |x - a| \\ &= (|x| + |a|) |x - a| \\ &= (|x - a + a| + |a|) |x - a| \\ &\leq (|x - a| + 2|a|) |x - a| \\ &= |x - a|^2 + 2|a| \times |x - a| \end{aligned}$$

Then if we let

$$\delta < \min\left\{\frac{\epsilon}{4|a|}, \sqrt{\frac{\epsilon}{2}}\right\}$$

we see

$$|x - a| < \delta \implies |x^2 - a^2| \leq |x - a|^2 + 2|a| \times |x - a| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

We now prove f is not uniformly continuous. Arbitrarily pick $\delta \in \mathbb{R}^+$, we wish to find a pair of real number b, c such that $d(b, c) < \delta$ but $d(f(b), f(c)) > 1$.

Observe

$$d(f(b), f(c)) = |f(c) - f(b)| = |c^2 - b^2| = |c - b| \times |c + b|$$

Then we can let $b = \frac{2}{\delta}$ and $c = \frac{2}{\delta} + \frac{\delta}{2}$, and we have

$$d(b, c) = \frac{\delta}{2} < \delta$$

and

$$d(f(b), f(c)) = |c - b| \times |c + b| = \frac{\delta}{2} \times \left(\frac{4}{\delta} + \frac{\delta}{2}\right) = 2 + \frac{\delta^2}{4} > 1$$

■

We now prove Heine-Cantor Theorem.

Lemma 5.3.5. (Lebesgue Number Covering Lemma)

Theorem 5.3.6. (Linear Function is Uniformly Continuous) Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f(x) = ax + b$$

Then

f is uniformly continuous on every subset of \mathbb{R}

Proof. It is cleat that $\delta = \frac{\epsilon}{a}$ works.

■

Theorem 5.3.7. (Uniformly Continuous Function Need not be Differentiable)

Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f(x) = |x|$$

Then

f is uniformly continuous but not differentaible everywhere

Proof. Clearly, f is not differentiable at $x = 0$. We now prove $\delta = \epsilon$ works. Let $0 < h < \delta$. If $a \geq 0$, we see

$$|f(a + h) - f(a)| = ||a + h| - |a|| = h < \delta = \epsilon$$

If $a \leq -\delta$, we see

$$|f(a + h) - f(a)| = |(-a - h) - (-a)| = h < \delta = \epsilon$$

If $-\delta < a < 0$, there are two possibilities $a + h > 0$ or $a + h \leq 0$. It the former

$$|f(a + h) - f(a)| = ||a + h| + a| = |2a + h| = 2a + h < h < \delta = \epsilon$$

If the latter

$$|f(a + h) - f(a)| = ||a + h| - |a|| = |-h| = h < \delta = \epsilon$$

■

Theorem 5.3.8. (Addition of Uniform Continuous Functions is Uniform Continuous) Given $f : X \rightarrow \mathbb{R}$ and $g : X \rightarrow \mathbb{R}$

f, g are both uniform continuous $\implies f + g$ is uniform continuous

Proof. Fix ϵ . Let

$$\delta = \min\{\delta_f(\frac{\epsilon}{2}), \delta_g(\frac{\epsilon}{2})\}$$

We now prove such δ works. Suppose $d(p, q) < \delta$. We have

$$d(p, q) < \delta_f(\frac{\epsilon}{2}) \text{ and } \delta_g(\frac{\epsilon}{2})$$

which give us

$$|f(p) - f(q)| < \frac{\epsilon}{2} \text{ and } |g(p) - g(q)| < \frac{\epsilon}{2}$$

Then we see

$$\begin{aligned} |(f + g)(p) - (f + g)(q)| &= |f(p) - f(q) + g(p) - g(q)| \\ &\leq |f(p) - f(q)| + |g(p) - g(q)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

■

Lastly, we introduce Cauchy continuous.

Theorem 5.3.9. (Cauchy Continuous) Given $f : A \subseteq X \rightarrow Y$. Suppose

- (a) f is uniformly continuous
- (b) A is dense in X
- (c) Y is complete

Then

There exists a unique continuous extension \bar{f} on X

Proof. Define

$$\bar{f}(x) = \lim_{n \rightarrow \infty} f(x_n) \text{ where } x_n \rightarrow x$$

We first prove \bar{f} is well defined.

Assume $x_n \rightarrow x$ and $y_n \rightarrow x$ and $\lim_{n \rightarrow \infty} f(x_n) \neq \lim_{n \rightarrow \infty} f(y_n)$. Let

$$p = \lim_{n \rightarrow \infty} f(x_n) \text{ and } q = \lim_{n \rightarrow \infty} f(y_n) \text{ and } \alpha = d(p, q)$$

Let N satisfy

$$\forall n > N, f(x_n) \in B_{\frac{\alpha}{3}}(p) \text{ and } f(y_n) \in B_{\frac{\alpha}{3}}(q)$$

Let $\delta = \delta(\frac{\alpha}{3})$, and let N_1 satisfy

$$n > N_1 \implies x_n, y_n \in B_{\frac{\delta}{2}}(x)$$

Let $m > \max\{N, N_1\}$. We now have

$$(a) \ d(f(x_m), p) < \frac{\alpha}{3}$$

$$(b) \ d(f(y_m), q) < \frac{\alpha}{3}$$

$$(c) \ d(f(x_m), f(y_m)) < \frac{\alpha}{3}$$

Then

$$\alpha = d(p, q) < \alpha \text{ CaC (done)}$$

We now prove \bar{f} is defined every where on X .

It is clear that \bar{f} is defined every where on A , since if $a \in A$, we can let $a_n = a$ and see

$$\bar{f}(a) = \lim_{n \rightarrow \infty} f(a_n) = f(a)$$

This also show that \bar{f} is indeed an extension of f .

Suppose $x \in X \setminus A$. We know $x \in A'$. Let $a_n \rightarrow x$. Because Y is complete, we only have to prove

$$\{f(a_n)\} \text{ is Cauchy}$$

Fix ϵ . Let $\delta = \delta(\epsilon)$. Let N satisfy

$$\forall n > N, a_n \in B_{\frac{\delta}{2}}(x)$$

We see for all $p, q > N$

$$a_p, a_q \in B_{\frac{\delta}{2}}(x) \implies d(a_p, a_q) < \delta \implies d(f(a_p), f(a_q)) < \epsilon \text{ (done)}$$

We now show \bar{f} is continuous on X .

Because f is uniformly continuous on A . We know \bar{f} is continuous on A . Then we see by Theorem 3.11.2, \bar{f} is continuous on X . (done)

Lastly, we prove such completion is unique.

Let g be a completion of f . It is clear that for $a \in A$, we have $g(a) = \bar{f}(a)$. Suppose $x \in X \setminus A$. Assume $g(x) \neq \bar{f}(x)$. Let

$$a_n \rightarrow x$$

We now have

$$\lim_{n \rightarrow \infty} f(a_n) = \bar{f}(x) \neq g(x)$$

By Theorem 3.11.2, we see g is not continuous CaC (done) ■

5.4 Tests for Uniform Continuity

Theorem 5.4.1. (Heine-Cantor Theorem) Given $f : X \rightarrow Y$

X is compact and f is continuous $\implies f$ is uniformly continuous

Proof. Fix ϵ . We wish to find δ such that

$$d(x, y) < \delta \implies d(f(x), f(y)) < \epsilon$$

Because f is continuous, for each $x \in X$ we can associate it with a positive number δ_x such that

$$\forall y \in B_{\delta_x}(x), f(y) \in B_{\frac{\epsilon}{2}}(f(x))$$

The set $\{B_{\frac{\delta_x}{2}}(x) : x \in X\}$ is an open cover of X . Because X is compact, we know there exists a finite sub-cover

$$\{B_{\frac{\delta_{x_1}}{2}}(x_1), \dots, B_{\frac{\delta_{x_m}}{2}}(x_m)\}$$

Let

$$\delta = \min\left\{\frac{\delta_{x_1}}{2}, \dots, \frac{\delta_{x_m}}{2}\right\}$$

Suppose $d(x, y) < \delta$. We have to prove

$$d(f(x), f(y)) < \epsilon$$

Because

$$\{B_{\frac{\delta_{x_1}}{2}}(x_1), \dots, B_{\frac{\delta_{x_m}}{2}}(x_m)\} \text{ is an sub-cover}$$

We know there exists $r \in \{1, \dots, m\}$ such that

$$x \in B_{\frac{\delta_{x_r}}{2}}(x_r)$$

Which means

$$d(x, x_r) < \frac{\delta_{x_r}}{2}$$

Then because

$$d(x, y) < \delta \leq \frac{\delta_{x_r}}{2}$$

We have

$$d(y, x_r) \leq d(y, x) + d(x, x_r) < \frac{\delta_{x_r}}{2} + \frac{\delta_{x_r}}{2} = \delta_{x_r}$$

Then because

$$d(x, x_r) < \frac{\delta_{x_r}}{2} < \delta_{x_r}$$

We now have

$$x \in B_{\delta_{x_r}}(x_r) \text{ and } y \in B_{\delta_{x_r}}(x_r)$$

By triangular inequality, we now have

$$d(f(x), f(y)) \leq d(f(x), f(x_r)) + d(f(x_r), f(y)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

■

Theorem 5.4.2. (Bounded Derivative and Open Interval Domain Implies Uniformly Continuous) Let I be an open interval. Given differentiable $f : I \rightarrow \mathbb{R}$

$$f'[I] \text{ is bounded} \implies f \text{ is uniformly continuous}$$

Proof. Let M satisfy

$$\forall x, |f'(x)| < M$$

Fix ϵ . We prove $\delta := \frac{\epsilon}{M}$ works. Let h satisfy

$$0 < h < \delta$$

We see

$$|f(a+h) - f(a)| = \left| \int_a^{a+h} f'(x) dx \right| \leq hM < \delta M = \epsilon$$

■

Theorem 5.4.3. (Continuous and Horizontal Asymptote Implies Uniformly Continuity) Given $f : [0, \infty) \rightarrow \mathbb{R}$. Suppose

- (a) f is continuous.
- (b) $f(x) \rightarrow L$ as $x \rightarrow \infty$

Then

$$f \text{ is uniformly continuous}$$

Proof. Fix ϵ . Because $f(x) \rightarrow L$ as $x \rightarrow \infty$, we know there exists $R \in \mathbb{R}^+$ such that

$$\forall x > R, |f(x) - L| < \frac{\epsilon}{2}$$

Also, because $[0, R+1]$ is compact and f is continuous, we know f is uniformly continuous on $[0, R+1]$, then we know there exists δ' such that

$$\forall y', z' \in [0, R+1], |y' - z'| < \delta' \implies |f(y') - f(z')| < \epsilon$$

We now prove

$$\delta := \min\{\delta', 1\} \text{ works}$$

WOLG, let $z + \delta > y > z \geq 0$. If $z < R$, then $y < z + \delta < R + \delta \leq R + 1$. This tell us $z, y \in [0, R + 1]$. Then because $y - z < \delta < \delta'$, we have $|f(y') - f(z')| < \epsilon$. If $z \geq R$. We have $y > z > R$. Then we have

$$|f(y) - f(z)| \leq |f(y) - L| + |f(z) - L| < \epsilon$$

■

Theorem 5.5.1, Theorem 5.5.2 and Theorem 5.5.3 are three easy positive criteria to check for elementary function. We now proceed to give other three negative criteria.

Theorem 5.4.4. (Derivative Converge to Infinity Implies Not Uniformly Continuous) Given differentiable $f : \mathbb{R} \rightarrow \mathbb{R}$

$$\lim_{x \rightarrow \infty} f'(x) = \infty \implies f \text{ is not uniformly continuous}$$

Proof. Fix ϵ and δ . We need to find a pair (a, h) such that

$$h < \delta \text{ and } |f(a + h) - f(a)| > \epsilon$$

Let $h < \delta$. Because $f'(x) \rightarrow \infty$ as $x \rightarrow \infty$, we know there exists r such that

$$\forall x > r, f'(x) > \frac{\epsilon + 1}{h}$$

Let $a > r$. By Fundamental Theorem of Calculus, we know

$$f(a + h) = \int_a^{a+h} f'(x)dx + f(a)$$

Then

$$|f(a + h) - f(a)| = \int_a^{a+h} f'(x)dx \geq h \times \frac{\epsilon + 1}{h} > \epsilon$$

■

Theorem 5.4.5. (Uniformly Continuous Map Totally Bounded Set to Totally Bounded Set) Let X be a totally bounded set.

$$f : X \rightarrow Y \text{ is uniformly continuous} \implies f[X] \text{ is totally bounded}$$

Proof. Fix ϵ . Let δ be from uniform continuity of f . Because X is totally bounded, we can cover X with

$$X \subseteq \bigcup_{i=1}^m B_\delta(x_i)$$

We now prove

$$f[X] \subseteq \bigcup_{i=1}^m B_\epsilon(f(x_i))$$

Arbitrarily pick $y \in f[X]$. We wish to prove it is in some ball $B_\epsilon(f(x_i))$. We know there exists x such that

$$y = f(x)$$

Because X is covered by the finite m balls, we know there exists $n \in \{1, \dots, m\}$ such that

$$x \in B_\delta(x_n)$$

Then because $d(x, x_n) < \delta$, we see $d(f(x), f(x_n)) < \epsilon$. Then we have

$$y = f(x) \in B_\epsilon(f(x_n)) \text{ (done)}$$

■

Theorem 5.4.6. (Uniform Continuous has No Vertical Asymptote) Given $f : X \rightarrow \mathbb{R}$

$$\lim_{x \rightarrow p} f(x) = \infty \implies f \text{ is not uniformly continuous}$$

Proof. We know there exists a strict sequence $p_n \rightarrow p$ such that

$$f(p_n) \rightarrow \infty$$

Fix $\epsilon = 1$ and δ . We wish to find (p_m, p_u) such that

$$d(p_m, p_u) < \delta \text{ and } d(f(p_m), f(p_u)) > 1$$

Let $p_m \in B_{\frac{\delta}{2}}(p)$. We know there exists N such that

$$n > N \implies f(p_n) > f(p_m) + 1$$

Let $u > \min\{N, m\}$. Because $p_n \rightarrow p$ is strict, $u > m$ and $p_m \in B_{\frac{\delta}{2}}(p)$. We see

$$p_u \in B_{\frac{\delta}{2}}(p)$$

This tell us

$$d(p_m, p_u) < \delta$$

Then because $u > N$, we see

$$f(p_n) > f(p_m) + 1$$

■

Theorem 5.4.7. (Tests for Uniform Continuity) Given $f : X \rightarrow Y$. The following suggest f is uniform continuous

- (a) f is continuous and X is compact
- (b) $X \subseteq \mathbb{R}$ is an open interval, $Y \subseteq \mathbb{R}$, f is differentiable and $f'[X]$ is bounded
- (c) X is of the form $[a, \infty)$, $Y \subseteq \mathbb{R}$, f is continuous and $f(x) \rightarrow L$ as $x \rightarrow \infty$

The following suggest f is NOT uniform continuous

- (a) $\lim_{x \rightarrow \infty} f'(x) = \infty$
- (b) X is totally bounded while $f[X]$ isn't
- (c) $\lim_{x \rightarrow p} f(x) = \infty$

Proof. For the three positive suggestion, see Theorem 5.5.1, Theorem 5.5.2 and Theorem 5.5.3. For the three negative suggestion, see Theorem 5.5.4, Theorem 5.5.5 and Theorem 5.5.6 ■

Notice that it is possible $\lim_{x \rightarrow p} f'(x) = \infty$, but f still be uniform continuous. See \sqrt{x} .

Theorem 5.4.8. (Example and Non-Example)

- (a) $\sin x$ is uniformly continuous on \mathbb{R}
- (b) $\frac{\cos x^3}{x}$ is uniformly continuous on $[1, \infty)$
- (c) $\frac{\cos x^3}{x}$ is Not uniformly continuous on $(0, \infty)$
- (d) $x \sin x$ is Not uniformly continuous on \mathbb{R}
- (e) $\tan x$ is not uniform continuous on $(\frac{-\pi}{2}, \frac{\pi}{2})$

Proof. $\sin x$ is differentiable and the derivative is bounded by 2.

$\frac{\cos x^3}{x}$ is uniformly continuous on $[1, \infty)$ because $\frac{\cos x^3}{x} \rightarrow 0$ as $x \rightarrow \infty$.

$\frac{\cos x^3}{x}$ is not uniformly continuous on $(0, \infty)$ because $\frac{\cos x^3}{x} \rightarrow \infty$ as $x \rightarrow 0$.

Observe

$$\begin{aligned} (a+h)\sin(a+h) - a\sin a &= (a+h)(\cos h \sin a + \sin h \cos a) - a\sin a \\ &= (a+h)\sin h \cos a + (\sin a)(a(\cos h - 1) + h \cos h) \end{aligned}$$

If we let a be the of form $a = 2n\pi$. We have

$$\begin{aligned}(a + h) \sin(a + h) - a \sin a &= (h + 2n\pi) \sin h + 2n\pi(\cos h - 1) + h \cos h \\ &= 2n\pi(\sin h + \cos h - 1) + h(\sin h + \cos h)\end{aligned}$$

Fix $0 < h < \delta$. It is now easy to see

$$|(a + h) \sin(a + h) - a \sin a| \rightarrow \infty \text{ as } n \rightarrow \infty$$

$\tan x$ is not uniformly continuous on $(\frac{-\pi}{2}, \frac{\pi}{2})$ because $\tan x \rightarrow \infty$ as $x \rightarrow \frac{\pi}{2}$ ■

5.5 Absolute Continuous and Lipschitz Continuous

Absolutely continuous is about induced measure and Lebesgue measure. I should keep that in note and study this years after.

Definition 5.5.1. (Absolute Continuous) Let I be an interval in \mathbb{R} . We say a function $f : I \rightarrow \mathbb{R}$ is absolute continuous if for all ϵ there exists δ such that whenever a finite sequence of pairwise disjoint sub-intervals $(x_k, y_k) \subseteq I$ satisfy

$$\sum_{k=1}^m (y_k - x_k) < \delta$$

we have

$$\sum_{k=1}^m |f(y_k) - f(x_k)| < \epsilon$$

Theorem 5.5.2. (Absolute Continuous is Stronger than Uniformly Continuous)

Let I be an interval of \mathbb{R} . Given $f : I \rightarrow \mathbb{R}$, we have

$$f \text{ is absolute continuous} \implies f \text{ is uniformly continuous}$$

Proof. Fix ϵ . Because f is absolutely continuous, we know there exists δ such that for all finite sequence of pairwise disjoint sub-interval $(x_k, y_k) \subseteq I$ that satisfy

$$\sum_{k=1}^m (y_k - x_k) < \delta$$

we have

$$\sum_{k=1}^m |f(y_k) - f(x_k)| < \epsilon$$

Then for all pair of real number p, q that satisfy

$$d(p, q) = |p - q| < \delta$$

we can WOLG, let $a_1 = \min\{p, q\}$ and $b_1 = \max\{p, q\}$, and see

$$\sum_{k=1}^1 b_1 - a_1 = |p - q| < \delta$$

which give us

$$d(f(p), f(q)) = |f(p) - f(q)| = |f(b_1) - f(a_1)| = \sum_{k=1}^1 |f(b_k) - f(a_k)| < \epsilon$$

as we wish

■

Theorem 5.5.3. (Absolutely Continuous is Strictly Stronger Than Uniformly Continuous) Define

$$f(x) = \begin{cases} x \cos \frac{1}{x} & \text{if } x \in (0, 1] \\ 0 & \text{if } x = 0 \end{cases}$$

We have

f is uniformly continuous but not absolutely continuous

Proof. Because f is defined on $[0, 1]$ and f is clearly continuous, by Theorem 5.5.1, we know f is uniformly continuous. We now prove f is not absolutely continuous. Fix $\epsilon = 1$. It suffice to show that for all δ , there exists a finite sequence $\{(a_k, b_k)\}_{k=1}^n$ of pairwise disjoint sub-interval such that

$$\sum_{k=1}^n b_k - a_k < \delta \text{ and } \sum_{k=1}^n f(b_k) - f(a_k) > 1 = \epsilon$$

Let m be an odd natural. For all $k \in \mathbb{N}$, define

$$(a_k, b_k) := \left(\frac{1}{(2k+1)m\pi}, \frac{1}{2km\pi} \right)$$

We have

$$b_k - a_k = \frac{1}{(4k^2 + 2k)m\pi}$$

and because m is odd, we have

$$f(b_k) = \frac{\cos(2km\pi)}{2km\pi} = \frac{1}{2km\pi} \text{ and } f(a_k) = \frac{\cos((2k+1)m\pi)}{(2k+1)m\pi} = \frac{-1}{(2k+1)m\pi}$$

This give us

$$|f(b_k) - f(a_k)| = \frac{4k+1}{(4k^2 + 2k)m\pi}$$

Notice that by limit comparison test, we know $\sum_{k=1}^{\infty} \frac{1}{4k^2+2k}$ converge and $\sum_{k=1}^{\infty} \frac{4k+1}{4k^2+2k}$ diverge. Define

$$\alpha := \sum_{k=1}^{\infty} \frac{1}{4k^2 + 2k} = m\pi \sum_{k=1}^{\infty} b_k - a_k$$

We have

$$\sum_{k=1}^{\infty} b_k - a_k = \frac{\alpha}{m\pi}$$

Then for all δ , we see if we require $m > \frac{\alpha}{\pi\delta}$. No matter how large is n , we have

$$\sum_{k=1}^n b_k - a_k < \sum_{k=1}^{\infty} b_k - a_k = \frac{\alpha}{m\pi} < \delta$$

And because

$$\sum_{k=1}^{\infty} |f(b_k) - f(a_k)| = \frac{1}{m\pi} \sum_{k=1}^{\infty} \frac{4k+1}{4k^2+2k} = \infty$$

We know there exists n such that

$$\sum_{k=1}^n |f(b_k) - f(a_k)| > 1 = \epsilon$$

■

We now introduce Lipschitz continuous.

Definition 5.5.4. (Lipschitz Continuous) We say a function $f : X \rightarrow Y$ is Lipschitz continuous if there exists a non-negative real constant K such that

$$\forall p, q \in X, d(f(p), f(q)) \leq Kd(p, q)$$

Theorem 5.5.5. (Lipschitz Continuous is Stronger than Uniform Continuous)
Given $f : X \rightarrow Y$

$$f \text{ is Lipschitz continuous} \implies f \text{ is uniform continuous}$$

Proof. Fix ϵ . If $K = 0$, we see

$$\forall p, q \in X, d(f(p), f(q)) \leq (0)d(p, q) = 0$$

Then clearly f is uniform continuous as every positive real can be our δ .

If $K > 0$, we see

$$d(p, q) < \frac{\epsilon}{K} \implies d(f(p), f(q)) < Kd(p, q) = \epsilon$$

In this case, $\frac{\epsilon}{K}$ can be our δ .

■

Theorem 5.5.6. (Lipschitz Continuous is Stronger than Absolutely Continuous)
Given $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$

$$f \text{ is Lipschitz continuous} \implies f \text{ is absolutely continuous}$$

Proof. Because f is Lipschitz continuous, we know there exists non-negative K such that

$$\forall x, y \in I, |f(y) - f(x)| \leq K |y - x|$$

If $K = 0$, then f is a constant function. Every positive real can work as δ . We only have to consider when $K > 0$. Let $\delta = \frac{\epsilon}{K}$. We proceed to $\delta = \frac{\epsilon}{K}$ works.

Arbitrarily pick a finite sequence $\{(a_k, b_k)\}_{k=1}^n$ of pairwise disjoint sub-interval of I such that

$$\sum_{k=1}^n b_k - a_k < \delta$$

It remains to show

$$\sum_{k=1}^n |f(b_k) - f(a_k)| < \epsilon$$

Observe

$$\begin{aligned} \sum_{k=1}^n |f(b_k) - f(a_k)| &\leq \sum_{k=1}^n K |b_k - a_k| \\ &= K \sum_{k=1}^n b_k - a_k \\ &< K \delta \\ &= \epsilon \end{aligned}$$

■

Lemma 5.5.7.

$$0 \leq h < a < b \implies \sqrt{b} - \sqrt{a} \leq \sqrt{b-h} - \sqrt{a-h}$$

Proof. Define

$$f(h) = \sqrt{b-h} - \sqrt{a-h} - (\sqrt{b} - \sqrt{a})$$

We have to show

$$\forall h \in [0, a), f(h) > 0$$

Because $b > a$, we can compute

$$f'(h) = \frac{\sqrt{b-h} - \sqrt{a-h}}{2\sqrt{(a-h)(b-h)}} > 0$$

Observe that $f(0) = 0$ and we are done.

■

Theorem 5.5.8. (Lipschitz Continuous is Strictly Stronger Than Absolutely Continuous) Define

$$f : [0, 1] \rightarrow \mathbb{R}, x \mapsto \sqrt{x}$$

Then

f is absolutely continuous but not Lipschitz continuous

Proof. Notice that given $y \neq x$, we have

$$|\sqrt{y} - \sqrt{x}| \leq K |y - x| \iff \frac{1}{\sqrt{y} + \sqrt{x}} \leq K$$

Then for all K we can use $x = (\frac{2}{K})^2$ and $y = (\frac{1}{K})^2$ to have $|\sqrt{y} - \sqrt{x}| > K |y - x|$. We have proved f is not Lipschitz continuous.

We now prove f is absolutely continuous. Fix ϵ . We proceed to show

$$\delta = \epsilon^2 \text{ works}$$

Let $\{(a_k, b_k)\}_{k=1}^n$ be a finite sequence of pairwise disjoint sub-interval of $[0, 1]$ such that

$$\sum_{k=1}^n b_k - a_k < \delta$$

For all $k \in \{1, \dots, n\}$, define

$$h_k = b_k - a_k \text{ and } c_k = \sum_{u=1}^{k-1} h_u \text{ and } d_k = c_k + h_k$$

We have

$$\sum_{k=1}^n d_k - c_k < \sum_{k=1}^n h_k < \delta$$

It is easy to see $c_{k+1} = d_k$, then we have

$$\sum_{k=1}^n \sqrt{d_k} - \sqrt{c_k} = \sqrt{d_n} = \sqrt{\sum_{k=1}^n h_k} = \sqrt{\sum_{k=1}^n b_k - a_k} < \sqrt{\delta} = \epsilon$$

Then, it suffice to show

$$\sum_{k=1}^n \sqrt{b_k} - \sqrt{a_k} \leq \sum_{k=1}^n \sqrt{d_k} - \sqrt{c_k}$$

Notice that $d_k - c_k = b_k - a_k$, so to satisfy the hypothesis of Lemma 5.5.7, we only have to show

$$\forall k, c_k \leq a_k$$

We prove this by induction. Base case is trivial, as $c_1 = 0$. Suppose $c_m \leq a_m$. We have to show $c_{m+1} \leq a_{m+1}$.

Observe

$$c_{m+1} = c_m + b_m - a_m \leq b_m \leq a_{m+1}$$

We can now use Lemma 5.5.7 to see

$$\forall k \in \{1, \dots, n\}, \sqrt{b_k} - \sqrt{a_k} \leq \sqrt{d_k} - \sqrt{c_k}$$

Then

$$\sum_{k=1}^n \sqrt{b_k} - \sqrt{a_k} \leq \sum_{k=1}^n \sqrt{d_k} - \sqrt{c_k} < \epsilon \text{ (done)}$$

■

Corollary 5.5.9. (Lipschitz Continuous is Strictly Stronger than Uniformly Continuous)

Proof. The f in Lemma 5.5.8 is uniformly continuous, as Theorem 5.5.2 suggest. ■

Chapter 6

Rudin: Differentiation

6.1 Derivative

Definition 6.1.1. (Definition of Derivative) Given $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$. Suppose x is a limit point of I , we say f is differentiable at x if

$$\frac{f(x+h) - f(x)}{h} \text{ converge as } h \rightarrow 0$$

and we say

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \text{ is the derivative of } f$$

The derivative function is then defined

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

Notice that if $f'(x) = \infty$, then at that point it is not even continuous?

Theorem 6.1.2. (Equivalent Writing of Derivative)

$$f(x+h) - f(x) = h(f'(x) + o(h))$$

Proof. Deduce

$$\begin{aligned} & \frac{f(x+h) - f(x)}{h} \rightarrow f'(x) \\ \implies & \frac{f(x+h) - f(x)}{h} - f'(x) = o(h) \\ \implies & f(x+h) - f(x) = h(f'(x) + o(h)) \end{aligned}$$



Theorem 6.1.3. (Arithmetic of Derivative of Real to Real Function) Suppose $f, g : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ are both differentiable at x , we have

$$(a) (f + g)'(x) = f'(x) + g'(x)$$

$$(b) (fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

$$(c) \left(\frac{f}{g}\right)'(x) = \frac{g(x)f'(x) - g'(x)f(x)}{g^2(x)}$$

Proof. (a) is a special case of Theorem 5.1.4. We now prove

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

Observe

$$\begin{aligned} fg(x+h) - fg(x) &= f(x+h)g(x+h) - f(x)g(x) \\ &= f(x+h)[g(x+h) - g(x)] + g(x)[f(x+h) - f(x)] \end{aligned}$$

Then

$$\frac{fg(x+h) - fg(x)}{h} = f(x+h) \times \frac{g(x+h) - g(x)}{h} + g(x) \times \frac{f(x+h) - f(x)}{h} \rightarrow 0 \text{ (done)}$$

We now prove

$$\left(\frac{f}{g}\right)'(x) = \frac{g(x)f'(x) - g'(x)f(x)}{g^2(x)}$$

Observe

$$\begin{aligned} \frac{\frac{f}{g}(x+h) - \frac{f}{g}x}{h} &= \frac{f(x+h)g(x) - f(x)g(x+h)}{g(x+h)g(x)h} \\ &= \frac{1}{g(x+h)g(x)} \times \frac{f(x+h)g(x) - f(x)g(x+h)}{h} \\ &= \frac{1}{g(x+h)g(x)} \times \frac{g(x)[f(x+h) - f(x)] + f(x)[g(x) - g(x+h)]}{h} \rightarrow 0 \text{ (done)} \end{aligned}$$

■

Theorem 6.1.4. (Chain Rule of Real to Real Function) Given $f, g : \mathbb{R} \rightarrow \mathbb{R}$. Suppose

(a) f is differentiable at x

(b) g is differentiable at $f(x)$

we have

$$(g \circ f)'(x) = g'(f(x))f'(x)$$

Proof. We are given

$$\begin{aligned}f(h+x) - f(x) &= h[f'(x) + o(h)] \\g(h+x) - g(x) &= h[g'(x) + o(h)]\end{aligned}$$

Then because $f(x+h) \rightarrow f(x)$

$$\begin{aligned}(g \circ f)(x+h) - (g \circ f)(x) &= [f(x+h) - f(x)] \times [g'(f(x)) + o(f(x+h) - f(x))] \\&= h[f'(x) + o(h)] \times [g'(f(x)) + o(f(x+h) - f(x))]\end{aligned}$$

Notice that because $f(x+h) - f(x) \rightarrow 0$ as $h \rightarrow 0$, then

$$\begin{aligned}\frac{g \circ f(x+h) - g \circ f(x)}{h} &= [f'(x) + o(h)] \times [g' \circ f(x) + o(f(x+h) - f(x))] \\&\rightarrow f'(x) \times g' \circ f(x) \text{ as } h \rightarrow 0\end{aligned}$$

■

6.2 Mean Value Theorem

Definition 6.2.1. (Definition of Local Maximum) Given $f : X \rightarrow \mathbb{R}$, we say f has a local maximum at x if there exists δ such that

$$y \in B_\delta(x) \implies f(y) \leq f(x)$$

Theorem 6.2.2. (For $f : \mathbb{R} \rightarrow \mathbb{R}$, Local Maximum is Either Isolated or Has 0 Derivative) Given $f : \mathbb{R} \rightarrow \mathbb{R}$, if f has a local maximum at x , then either

- (a) x is an isolated point of domain of f
- (b) f is not differentiable at x
- (c) $f'(x) = 0$

Proof. Observe left side limit

$$\lim_{h \rightarrow 0^-} \frac{f(x+h) - f(x)}{h} \geq 0$$

Observe right side limit

$$\lim_{h \rightarrow 0^+} \frac{f(x+h) - f(x)}{h} \leq 0$$

Then we see $f'(x)$ must equals to 0 if possible. ■

Theorem 6.2.3. (Generalized Mean Value Theorem) Given

- (a) f, g are continuous on $[a, b]$
- (b) f, g are differentiable on (a, b)

Then there exists a point $x \in (a, b)$ such that

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x)$$

Proof. Define $v : [a, b] \rightarrow \mathbb{R}$ by

$$v(t) = [f(b) - f(a)]g(t) - [g(b) - g(a)]f(t)$$

We know v is continuous on $[a, b]$ and differentiable on (a, b) . Notice

$$v'(t) = [f(b) - f(a)]g'(t) - [g(b) - g(a)]f'(t)$$

Then, we reduce the problem into proving the existence of $x \in (a, b)$ such that

$$v'(x) = 0$$

Observe

$$\begin{aligned} v(a) &= [f(b) - f(a)]g(a) - [g(b) - g(a)]f(a) \\ &= f(b)g(a) - g(b)f(a) \\ &= [f(b) - f(a)]g(b) - [g(b) - g(a)]f(b) = v(b) \end{aligned}$$

Notice that v is continuous, then because the domain is $[a, b]$, we know v has maximum and minimum. If one of maximum and minimum is in (a, b) , the result follows from Theorem 6.2.2. If not, notice that v is a constant so the derivative is 0 every where. ■

Corollary 6.2.4. (Useful Form of Cauchy Mean Value Theorem) Given

- (a) f, g are continuous on $[a, b]$
- (b) f, g are differentiable (a, b)
- (c) g' does not equals to 0 in (a, b)

Then there exists a point $x \in (a, b)$ such that

$$\frac{f'(x)}{g'(x)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

Corollary 6.2.5. (Mean Value Theorem) Given

- (a) f is continuous on $[a, b]$
- (b) f is differentiable on (a, b)

there exists $x \in (a, b)$ such that

$$f(b) - f(a) = f'(x)(b - a)$$

Proof. Define g by

$$g(x) = x$$

Notice that $g'(x) = 1$ and $g(b) - g(a) = b - a$, then by Theorem 6.2.3, we know there exists x such that

$$f(b) - f(a) = (b - a)f'(x)$$

Corollary 6.2.6. (Derivative to Increasing without FTC)

- (a) $f'(x) \geq 0$ on $(a, b) \implies f$ increase on (a, b)
- (b) $f'(x) = 0$ on $(a, b) \implies f$ is constant on (a, b)
- (c) $f'(x) \leq 0$ on $(a, b) \implies f$ decrease on (a, b)

Proof. Assume **the opposite** and use corollary 6.2.5 ■

6.3 The Continuity of Derivative

Theorem 6.3.1. (Derivative Function Satisfy Intermediate Value Property) Given $f : [a, b] \rightarrow \mathbb{R}$

(a) f is differentiable

(b) $f'(a) < \lambda < f'(b)$

then there exists a point $x \in (a, b)$ such that

$$f'(x) = \lambda$$

Proof. Define $g : [a, b] \rightarrow \mathbb{R}$ by

$$g(t) = f(t) - \lambda t$$

We see g is differentiable and

$$g'(t) = f'(t) - \lambda$$

We have reduced the problem into finding $x \in (a, b)$ such that

$$g'(x) = 0$$

Notice that because f is differentiable, we know g is continuous. Then because $[a, b]$ is compact, we know g has a maximum, at which by Theorem 6.2.2, has 0 derivative. ■

Notice that if $f'(a) = \lambda = f'(b)$, such x may not exist in (a, b)

Corollary 6.3.2.

f is differentiable on $[a, b] \implies f'$ has only essential discontinuity on $[a, b]$

Proof. Assume f has a removable or jump discontinuity at $x \in [a, b]$, WOLOG, it must happen

$$f'(a) \neq \lim_{h \rightarrow 0^+}$$

$$f'(a) < \lim_{h \rightarrow 0} f'(a + h) = \lim_{h \rightarrow 0^+} f'(a + h) = \alpha$$

We can find δ small enough such that

$$x \in B_\delta(a) \implies f'(x) > \frac{\alpha + f'(a)}{2} > f'(a)$$

We now see

(a) f is differentiable on $[a, a + \frac{\delta}{2}]$

(b) $f'(a) < \frac{\alpha + f'(a)}{2} < f'(a + \frac{\delta}{2})$

From Theorem 6.3.1, we now obtain a contradiction. CaC

6.4 L'Hospital Rule

Theorem 6.4.1. (L'Hospital Rule) Given

- (a) $c \in \mathbb{R} \cup \{\pm\infty\}$
- (b) $f(x), g(x) \rightarrow 0$ or $|f(x)|, |g(x)| \rightarrow \infty$ as $x \rightarrow c$
- (c) f, g are differentiable at an interval I containing c except possibly at c
- (d) $g'(x) \neq 0$ for all $x \in I \setminus \{c\}$
- (e) $\lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} = \alpha \in \mathbb{R}$

We have

$$\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \alpha$$

Proof. We first prove when

$$f(x), g(x) \rightarrow c \text{ as } x \rightarrow c \text{ and } c \in \mathbb{R}$$

Let x_n be a sequence converge to c that does not contain c . WOLG, We only have to prove

$$\frac{f(x_n)}{g(x_n)} \rightarrow \alpha \text{ as } x \rightarrow c$$

Let $y_{n,k}$ be a double sequence satisfy

$$y_{n,k} \text{ is between } x_n \text{ and } c \text{ and } y_{n,k} \rightarrow c \text{ as } k \rightarrow \infty$$

By Theorem 6.2.3, we can construct a sequence $t_{n,k}$ such that

$$t_{n,k} \text{ is between } x_n \text{ and } y_{n,k} \text{ and } \frac{f'(t_{n,k})}{g'(t_{n,k})} = \frac{f(x_n) - f(y_{n,k})}{g(x_n) - g(y_{n,k})}$$

Observe

$$\begin{aligned} \frac{f'(t_{n,k})}{g'(t_{n,k})} &= \frac{f(x_n) - f(y_{n,k})}{g(x_n) - g(y_{n,k})} \\ &= \frac{\frac{f(x_n)}{g(x_n)} - \frac{f(y_{n,k})}{g(y_{n,k})}}{1 - \frac{g(y_{n,k})}{g(x_n)}} \end{aligned}$$

Then because

$$f(y_{n,k}), g(y_{n,k}) \rightarrow 0 \text{ as } k \rightarrow \infty$$

On right hand side, we have

$$\lim_{k \rightarrow \infty} \frac{\frac{f(x_n)}{g(x_n)} - \frac{f(y_{n,k})}{g(x_n)}}{1 - \frac{g(y_{n,k})}{g(x_n)}} = \frac{f(x_n)}{g(x_n)}$$

Now, we have

$$\lim_{k \rightarrow \infty} \frac{f'(t_{n,k})}{g'(t_{n,k})} = \frac{f(x_n)}{g(x_n)}$$

This give us

$$\lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \frac{f'(t_{n,k})}{g'(t_{n,k})}$$

For each n , we can select a m_n such that

$$d\left(\frac{f'(t_{n,m_n})}{g'(t_{n,m_n})}, \lim_{k \rightarrow \infty} \frac{f'(t_{n,k})}{g'(t_{n,k})}\right) < \frac{1}{n}$$

Then because t_{n,m_n} is bounded by x_n, c , we see

$$\begin{aligned} \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} &= \lim_{n \rightarrow \infty} \frac{f'(t_{n,m_n})}{g'(t_{n,m_n})} \\ &= \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \frac{f'(t_{n,k})}{g'(t_{n,k})} \\ &= \lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} \text{ (done)} \end{aligned}$$

■

6.5 Taylor's Theorem

Theorem 6.5.1. (Taylor's Theorem) Given

(a) f is n time continuously differentiable on I , which contain a

We have

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)(x-a)^k}{k!} + o(x-a)(x-a)^n$$

Proof. Let

$$P(x) = \sum_{k=0}^n \frac{f^{(k)}(a)(x-a)^k}{k!}$$

We wish to prove

$$\frac{f(x) - P(x)}{(x-a)^n} \rightarrow 0 \text{ as } x \rightarrow a$$

Notice that

$$P(x) = \sum_{k=0}^n \frac{f^{(k)}(a)(x-a)^k}{k!} \rightarrow f(a) = \frac{f(a)(x-a)^0}{0!} \text{ as } x \rightarrow a$$

We now can recursively applies L'Hospital Rule

$$\begin{aligned} \lim_{x \rightarrow a} \frac{f(x) - P(x)}{(x-a)^n} &= \lim_{x \rightarrow a} \frac{f'(x) - P'(x)}{n(x-a)^{n-1}} \\ &= \lim_{x \rightarrow a} \frac{f''(x) - P''(x)}{n(n-1)(x-a)^{n-2}} \\ &= \lim_{x \rightarrow a} \frac{f^{(n)}(x) - P^{(n)}(x)}{n!} \\ &= \lim_{x \rightarrow a} \frac{f^{(n)}(x) - f^{(n)}(x)}{n!} = 0 \end{aligned}$$

■

Theorem 6.5.2. (Mean Value Form Remainder of Taylor's Theorem) Given

$f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ is n time continuously differentiable at $a \in I$

Define

$$(a) \ P_n(x) = \sum_{k=0}^n \frac{f^{(k)}(a)(x-a)^k}{k!}$$

$$(b) \ R_n(x) = f(x) - P_n(x)$$

If

$$(a) \ G \text{ is continuous on } [a, x]$$

$$(b) \ G' \text{ exists and not equals to 0 on } (a, x)$$

We have

$$\exists \xi, \in (a, x) R_n(x) = \frac{f^{(n+1)}(\xi)(x-\xi)^n}{n!} \frac{G(x) - G(a)}{G'(\xi)}$$

Proof. Define $F : (a, x) \rightarrow \mathbb{R}$ by

$$F(t) = \sum_{k=0}^n \frac{f^{(k)}(t)(x-t)^k}{k!}$$

By Corollary 6.2.5, we now have

$$\frac{F'(\xi)}{G'(\xi)} = \frac{F(x) - F(a)}{G(x) - G(a)}$$

Compute

$$F(x) = f(x)$$

Compute

$$F(a) = \sum_{k=0}^n \frac{f^{(k)}(a)(x-a)^k}{k!} = P_n(x)$$

Compute

$$\begin{aligned} F'(\xi) &= \sum_{k=0}^n \frac{f^{(k+1)}(\xi)(x-\xi)^k - k f^{(k)}(\xi)(x-\xi)^{k-1}}{k!} \\ &= \frac{f^{(n+1)}(\xi)(x-\xi)^n}{n!} \end{aligned}$$

We now have

$$\frac{\frac{f^{(n+1)}(\xi)(x-\xi)^n}{n!}}{G'(\xi)} = \frac{R_n(x)}{G(x) - G(a)}$$

Then we can deduce

$$R_n(x) = \frac{f^{(n+1)}(\xi)(x - \xi)^n}{n!} \frac{G(x) - G(a)}{G'(\xi)}$$

■

Corollary 6.5.3. (Lagarange Form of Remainders in Taylor's Theorem) Let

$$G(t) = (x - t)^{n+1}$$

We have

$$\exists \xi \in (a, x), R_n(x) = \frac{f^{(n+1)}(\xi)(x - a)^{n+1}}{(n + 1)!}$$

Proof. Compute

$$G'(\xi) = -(n + 1)(x - \xi)^n$$

$$G(x) = 0$$

$$G(a) = (x - a)^{n+1}$$

The result now follows from Theorem 6.5.2.

■

Chapter 7

Rudin: Riemann-Stieltjes Integral

7.1 Definition and Existence of The Integral

In this section, we shall assume f is defined and bounded on $[a, b]$ and α is defined and increasing on $[a, b]$.

Definition 7.1.1. (Definition of Partition) By a Partition P of a compact interval $[a, b]$, we mean a finite set of points

$$P = \{x_1, \dots, x_n\}$$

such that

$$a = x_1 \leq x_2 \leq \dots \leq x_n = b$$

Definition 7.1.2. (Definition of Riemman Upper Sum) Given a partition P of $[a, b]$, the Riemann Upper and Lower Sum are respectively defined by

$$U(P, f) = \sum_{i=1}^n M_i \Delta x_i$$
$$L(P, f) = \sum_{i=1}^n m_i \Delta x_i$$

where

$$M_i = \sup_{x \in [x_{i-1}, x_i]} f(x)$$

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x)$$

$$\Delta x_i = x_i - x_{i-1}$$

Definition 7.1.3. (Definition of Upper Riemann Integral) The Upper and Lower Riemann Integral is defined

$$\overline{\int_a^b} f dx = \inf_P U(P, f)$$

$$\underline{\int_a^b} f dx = \sup_P L(P, f)$$

Theorem 7.1.4. (Existence of Upper and Lower Riemann Integral)

$$\overline{\int_a^b} f dx \text{ and } \underline{\int_a^b} f dx \text{ exist}$$

Proof. Let

$$m = \inf_{[a,b]} f \text{ and } M = \sup_{[a,b]} f$$

Because f is bounded, we know both α, β are finite. We only wish to show

$M(b-a)$ is an upper bound of $\{L(P, f) \in \mathbb{R} : P \text{ is a partition of } [a, b]\}$
 $m(b-a)$ is an lower bound of $\{U(P, f) \in \mathbb{R} : P \text{ is a partition of } [a, b]\}$

Arbitrarily pick a partition P of $[a, b]$. Observe that

$$\forall i \in \mathbb{N}, m \leq m_i \leq M_i \leq M$$

This is true because $[x_{i-1}, x_i] \subseteq [a, b]$. Now, we can deduce

$$m(b-a) = \sum_{i=1}^n m \Delta x_i \leq \sum_{i=1}^n m_i \Delta x_i \leq \sum_{i=1}^n M_i \Delta x_i \leq \sum_{i=1}^n M \Delta x_i = M(b-a)$$

Then we can just write

$$m(b-a) \leq L(P, f) \leq U(P, f) \leq M(b-a)$$

Because P is an arbitrary partition, our proof is done. (done) ■

Definition 7.1.5. (Definition of Riemann Integrable) We say bounded f is integrable on $[a, b]$ and write $f \in \mathcal{R}$ if

$$\overline{\int_a^b} f dx = \underline{\int_a^b} f dx$$

Now, we consider integral with respect to an increasing function.

Definition 7.1.6. (Definition of Riemman-Stieltjes Upper Sum) The Riemann-Stieltjes Upper and Lower Sum are respectively defined by

$$U(P, f, \alpha) = \sum_{i=1}^n M_i \Delta\alpha(x_i)$$

$$L(P, f, \alpha) = \sum_{i=1}^n m_i \Delta\alpha(x_i)$$

where

$$M_i = \sup_{x \in [x_{i-1}, x_i]} f(x)$$

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x)$$

$$\Delta\alpha(x_i) = \alpha(x_i) - \alpha(x_{i-1})$$

Theorem 7.1.7. (Useful Lemma)

$$\sup_{[x,y] \subseteq [x_{i-1}, x_i]} |f(x) - f(y)| = M_i - m_i$$

Proof. Assume $\sup_{[x,y] \subseteq [x_{i-1}, x_i]} |f(x) - f(y)| < M_i - m_i$. Then there exists $z_1 \in [x_{i-1}, x_i]$ such that

$$f(z_1) > \sup_{[x,y] \subseteq [x_{i-1}, x_i]} |f(x) - f(y)| + m_i$$

Also, there exists $z_2 \in [x_{i-1}, x_i]$

$$f(z_1) - \sup_{[x,y] \subseteq [x_{i-1}, x_i]} |f(x) - f(y)| > f(z_2)$$

Then we see

$$|f(z_1) - f(z_2)| > \sup_{[x,y] \subseteq [x_{i-1}, x_i]} |f(x) - f(y)| \quad \text{CaC}$$

Assume $\sup_{[x,y] \subseteq [x_{i-1}, x_i]} |f(x) - f(y)| > M_i - m_i$. Then there exists $[x, y] \subseteq [x_{i-1}, x_i]$ such that

$$|f(x) - f(y)| > M_i - m_i$$

Then we see

$$\forall z_1 \in [x_{i-1}, x_i], |f(x) - f(y)| + m_i > f(z_1)$$

and

$$\forall z_1 \in [x_{i-1}, x_i], \forall z_2 \in [x_{i-1}, x_i], f(z_1) > f(z_2) - |f(x) - f(y)|$$

Then we have

$$\forall z_1, z_2, |f(x) - f(y)| > f(z_2) - f(z_1) \text{ CaC}$$

■

Definition 7.1.8. (Definition of Riemman-Stieltjes Upper Integral) The Riemann-Stieltjes Upper and Lower Integral of f with respect to α is

$$\begin{aligned} \overline{\int_a^b} f d\alpha &= \inf_P U(P, f, \alpha) \\ \underline{\int_a^b} f d\alpha &= \sup_P L(P, f, \alpha) \end{aligned}$$

Theorem 7.1.9. (Existence of Riemman-Stieltjes Upper Integral)

$$\overline{\int_a^b} f d\alpha, \underline{\int_a^b} f d\alpha \text{ both exists}$$

Proof. The proof is similar to that of Theorem 7.1.4

■

Definition 7.1.10. (Definition of Riemman-Stieltjes Integral) We say f is integrable on $[a, b]$ with respect to α if

$$\overline{\int_a^b} f d\alpha = \underline{\int_a^b} f d\alpha$$

We now consider the difference of upper and lower sums when different partition is given.

Definition 7.1.11. (Definition of Refinement of Partition) We say P^* is a refinement of P if

$$P \subseteq P^*$$

We say P^* is a common refinement of P_1, P_2 if

$$P_1 \cup P_2 \subseteq P^*$$

Theorem 7.1.12. (Upper and Lower Sum is Closer to the Value if Partition is more Refined)

$$P^* \text{ is a refinement of } P \implies L(P^*, f, \alpha) \geq L(P, f, \alpha) \text{ and } U(P^*, f, \alpha) \leq U(P, f, \alpha)$$

Proof. Notice that $P^* \setminus P$ is minus. We then complete the proof by induction, proving the result only when P^* is of one point larger than P . Let x^* be the extra point in P^* , and let x_j be the smallest point greater than x^* . We see

$$\begin{aligned} L(P, f, g) &= \sum_{i=1}^n m_i \Delta g(x_i) \\ L(P^*, f, g) &= \sum_{i=1}^{j-1} m_i \Delta g(x_i) + \sum_{i=j+1}^n m_i \Delta g(x_i) \\ &\quad + \left(\inf_{x \in [x_{j-1}, x^*]} f(x) \right) (g(x^*) - g(x_{j-1})) + \left(\inf_{x \in [x^*, x_j]} f(x) \right) (g(x_j) - g(x^*)) \end{aligned}$$

This give us

$$\begin{aligned} L(P^*, f, g) - L(P, f, g) &= m_j \Delta g(x_j) \\ &\quad - \left(\inf_{x \in [x_{j-1}, x^*]} f(x) \right) (g(x^*) - g(x_{j-1})) - \left(\inf_{x \in [x^*, x_j]} f(x) \right) (g(x_j) - g(x^*)) \end{aligned}$$

Because $[x_{j-1}, x^*], [x^*, x_j] \subseteq [x_{j-1}, x_j]$, we know

$$m_j \leq \min \left\{ \inf_{x \in [x_{j-1}, x^*]} f, \inf_{x \in [x^*, x_j]} f \right\}$$

Then we have

$$\begin{aligned} L(P^*, f, g) - L(P, f, g) &\leq m_j \Delta g(x_j) - m_j (g(x^*) - g(x_{j-1})) - m_j (g(x_j) - g(x^*)) \\ &= 0 \text{ (done)} \end{aligned}$$

■

Theorem 7.1.13. (Upper Sum is Greater than Lower Sum) Given a function f defined almost everywhere and bounded on $[a, b]$, a function g defined and increase on $[a, b]$, we have

$$\underline{\int_a^b f dg} \leq \overline{\int_a^b f dg}$$

Proof. Assume $\underline{\int_a^b f dg} > \overline{\int_a^b f dg}$. Then we know there exists P_1 and P_2 such that

$$U(P_2, f, g) < L(P_1, f, g)$$

Let $P^* = P_1 \cup P_2$. We have

$$L(P_1, f, g) \leq L(P^*, f, g) \leq U(P^*, f, g) \leq U(P_2, f, g) \text{ CaC}$$

■

Theorem 7.1.14. (Iff for f being Integrable on $[a, b]$ W.R.T. α)

$$f \text{ is integrable on } [a, b] \text{ W.R.T } \alpha \iff \forall \epsilon, \exists P, U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$$

Proof. (\longrightarrow)

Let P_1, P_2 satisfy

$$L(P_1, f, \alpha) > \int_a^b f d\alpha - \frac{\epsilon}{2} \text{ and } U(P_2, g, \alpha) < \int_a^b f d\alpha + \frac{\epsilon}{2}$$

Let $P^* = P_1 \cup P_2$, we see

$$[L(P^*, f, \alpha), U(P^*, f, \alpha)] \subseteq [\int_a^b f d\alpha - \frac{\epsilon}{2}, \int_a^b f d\alpha + \frac{\epsilon}{2}]$$

(\longleftarrow)

Assume $\int_a^b f d\alpha < \overline{\int_a^b f d\alpha}$. Let $\epsilon = \frac{\overline{\int_a^b f d\alpha} - \int_a^b f d\alpha}{2}$. We see for all P ,

$$U(P, f, \alpha) \geq \overline{\int_a^b f d\alpha} = \int_a^b f d\alpha + 2\epsilon \geq L(P, f, \alpha) + 2\epsilon$$

Then

$$\forall P, U(P, f, \alpha) - L(P, f, \alpha) \geq 2\epsilon \geq \epsilon \text{ CaC}$$

■

We now introduce some theorem regard to more computation.

Theorem 7.1.15. (Continuous on Compact Interval Implies Integrable)

$$f \text{ is continuous on } [a, b] \implies f \in \mathcal{R}(\alpha) \text{ on } [a, b]$$

Proof. Fix ϵ . We wish to

find a partition P such that $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$

We know f is uniformly continuous on $[a, b]$. Let $\delta = \delta(\frac{\epsilon}{2(\alpha(b) - \alpha(a))})$. Let P be a partition such that

$$\forall i, \Delta x_i < \delta$$

This is clearly possible as $b - a$ is finite. Let $[x, y] \subseteq [x_{i-1}, x_i]$. We see

$$d(x, y) < x_i - x_{i-1} = \Delta x_i < \delta$$

Then we know

$$|f(x) - f(y)| \leq \frac{\epsilon}{2(\alpha(b) - \alpha(a))}$$

Because

$$M_i - m_i = \sup_{[x,y] \subseteq [x_{i-1}, x_i]} |f(x) - f(y)|$$

We have

$$M_i - m_i \leq \frac{\epsilon}{2(\alpha(b) - \alpha(a))}$$

Then we see

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^n (M_i - m_i) \Delta \alpha(x_i) \\ &\leq \sum_{i=1}^n \frac{\epsilon}{2(\alpha(b) - \alpha(a))} \Delta \alpha(x_i) = \frac{\epsilon}{2} < \epsilon \text{ (done)} \end{aligned}$$

■

Theorem 7.1.16. (Monotonic Function f is Integrable W.R.T. Continuous α)

$$\begin{cases} f \text{ is monotone on } [a, b] \\ \alpha \text{ is continuous on } [a, b] \end{cases} \implies f \in \mathcal{R}(\alpha) \text{ on } [a, b]$$

Proof. Fix ϵ . We wish

to find P such that $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$

WOLG, suppose f increase on $[a, b]$. Let

$$m > \frac{(\alpha(b) - \alpha(a))(f(b) - f(a))}{\epsilon}$$

Because α is continuous, we can find a partition P such that

$$\alpha(x_j) = \sum_{i=1}^j \frac{i(\alpha(b) - \alpha(a))}{m} + \alpha(a)$$

Because f increase, we see

$$M_i = f(x_i) \text{ and } m_i = f(x_{i-1})$$

Then

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^n (f(x_i) - f(x_{i-1}))(\alpha(x_i) - \alpha(x_{i-1})) \\ &= \frac{\alpha(b) - \alpha(a)}{m} \sum_{i=1}^n f(x_i) - f(x_{i-1}) \\ &= \frac{(\alpha(b) - \alpha(a))(f(b) - f(a))}{m} < \epsilon \text{ (done)} \end{aligned}$$

■

Theorem 7.1.17. (Integrable Functions Are Closed Under Continuous Composition) Given

- (a) $f \in \mathcal{R}(\alpha)$ on $[a, b]$
- (b) $h : \mathbb{R} \rightarrow \mathbb{R}$ is continuous on $f[[a, b]]$

We have

$$h \circ f \in \mathcal{R}(\alpha) \text{ on } [a, b]$$

Proof. Fix ϵ . We wish to find a partition P of $[a, b]$ such that

$$U(P, h \circ f, \alpha) - L(P, h \circ f, \alpha) < \epsilon$$

Because f is bounded, we can Let $f[[a, b]] \subseteq [m, M]$. Also, let $K = \sup_{x \in [m, M]} |h|$.

Notice h is uniform continuous on $[m, M]$. Let $\epsilon_0 = \frac{\epsilon}{\alpha(b) - \alpha(a) + 2K}$. Let $\delta = \frac{\min\{\epsilon_0, \delta(\epsilon_0)\}}{2}$.

Because $f \in \mathcal{R}(\alpha)$ on $[a, b]$, we can find a partition $P = \{x_0, \dots, x_n\}$ of $[a, b]$ such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \delta^2$$

We now prove

such P works

Define

$$\begin{cases} M_i = \sup_{x \in [x_{i-1}, x_i]} f(x) \text{ and } M_i^* = \sup_{x \in [x_{i-1}, x_i]} h \circ f(x) \\ m_i = \inf_{x \in [x_{i-1}, x_i]} f(x) \text{ and } m_i^* = \inf_{x \in [x_{i-1}, x_i]} h \circ f(x) \end{cases}$$

Divide $\{(m_i^*, M_i^*) \in \mathbb{R}^2 : 1 \leq i \leq n\}$ into two classes

$$A = \{i \in \mathbb{N} : i \leq n, M_i - m_i < \delta\} \text{ and } B = \{i \in \mathbb{N} : i \leq n, M_i - m_i \geq \delta\}$$

Because $\delta < \delta(\epsilon_0)$, we have

$$\sum_{i \in A} (M_i^* - m_i^*) \Delta \alpha(x_i) \leq \sum_{i \in A} \epsilon_0 \Delta \alpha(x_i) \leq \epsilon_0 \sum_{i=1}^n \Delta \alpha(x_i) = \epsilon_0 (\alpha(b) - \alpha(a))$$

Notice that because $U(P, f, \alpha) - L(P, f, \alpha) < \delta^2$, we see

$$i \in B \implies M_i - m_i \geq \delta \implies \Delta \alpha(x_i) \leq \delta$$

Then because $\delta < \epsilon_0$, we have

$$\sum_{i \in B} (M_i^* - m_i^*) \Delta \alpha(x_i) \leq \delta \sum_{i \in B} (M_i^* - m_i^*) \leq \delta 2K \leq \epsilon_0 2K$$

We now see

$$\sum_{i=1}^n (M_i^* - m_i^*) \Delta \alpha(x_i) \leq \epsilon_0 (\alpha(b) - \alpha(a) + 2K) < \epsilon \text{ (done)}$$

■

Chapter 8

Vector Analysis

8.1 Important Inequalities

Theorem 8.1.1. (Young's Inequality) If $a, b \geq 0$ and $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Proof.

$$\begin{aligned} ab &= \exp(\ln a + \ln b) \\ &= \exp\left[\left(\frac{1}{p} \ln a^p\right) + \left(\frac{1}{q} \ln b^q\right)\right] \\ &\leq \frac{1}{p} \exp(\ln a^p) + \frac{1}{q} \exp(\ln b^q) \\ &= \frac{a^p}{p} + \frac{b^q}{q} \end{aligned}$$

The inequality above hold true because \exp is concave upward, and $\frac{1}{p} + \frac{1}{q} = 1$. ■

Theorem 8.1.2. (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

$$\|\mathbf{v}\|_p := \left(\sum |v_i|^p\right)^{\frac{1}{p}}$$

almost form a norm.

Proof. Observe

$$\begin{aligned}
\|\lambda \mathbf{v}\|_p &= \left(\sum |\lambda v_i|^p \right)^{\frac{1}{p}} \\
&= \left[\sum (|\lambda| |v_i|)^p \right]^{\frac{1}{p}} \\
&= \left[\sum (|\lambda|)^p (|v_i|)^p \right]^{\frac{1}{p}} \\
&= [(|\lambda|)^p \sum (|v_i|)^p]^{\frac{1}{p}} \\
&= [(|\lambda|)^p]^{\frac{1}{p}} \left(\sum |v_i|^p \right)^{\frac{1}{p}} \\
&= |\lambda| (\|\mathbf{v}\|_p)
\end{aligned}$$

Nonnegative and positive definiteness is trivial. Triangle Inequality *is left to proved*. ■

Theorem 8.1.3. (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

$$\|\mathbf{xy}\|_1 \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q$$

Proof. Normalize

$$\hat{\mathbf{x}} := \frac{\mathbf{x}}{\|\mathbf{x}\|_p} \text{ and } \hat{\mathbf{y}} := \frac{\mathbf{y}}{\|\mathbf{y}\|_q}$$

So we only wish to prove

$$\|\hat{\mathbf{x}}\hat{\mathbf{y}}\|_1 \leq 1$$

Notice because $(\sum |\hat{x}_i|^p)^{\frac{1}{p}} = \|\hat{\mathbf{x}}\|_p = \left\| \frac{\mathbf{x}}{\|\mathbf{x}\|_p} \right\|_p = \frac{\|\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = 1$, and similarly vice versa for \mathbf{y} , we have

$$\sum |\hat{x}_i|^p = 1 = \sum |\hat{y}_i|^q$$

So by Young's inequality, we have

$$\|\hat{\mathbf{x}}\hat{\mathbf{y}}\|_1 = \sum |\hat{x}_i \hat{y}_i| = \sum |\hat{x}_i| |\hat{y}_i| \leq \sum \frac{|\hat{x}_i|^p}{p} + \frac{|\hat{y}_i|^q}{q} = \frac{1}{p} + \frac{1}{q} = 1$$

■

Corollary 8.1.4. (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

$$p > 1 \implies \|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

Proof. Let $q = \frac{p}{p-1}$, so we have

$$\frac{1}{p} + \frac{1}{q} = \frac{p}{p} = 1$$

Then we deduce

$$\begin{aligned}
\|\mathbf{x} + \mathbf{y}\|_p^p &= \sum |x_i + y_i|^p \\
&= \sum |(x_i + y_i)^p| \\
&= \sum |(x_i + y_i)(x_i + y_i)^{p-1}| \\
&\leq \sum |x_i(x_i + y_i)^{p-1}| + |y_i(x_i + y_i)^{p-1}| \\
&= \|\mathbf{x}(\mathbf{x} + \mathbf{y})^{p-1}\|_1 + \|\mathbf{y}(\mathbf{x} + \mathbf{y})^{p-1}\|_1 \quad (\text{Operation is component wise}) \\
&\leq \|\mathbf{x}\|_p \|(\mathbf{x} + \mathbf{y})^{p-1}\|_q + \|\mathbf{y}\|_p \|(\mathbf{x} + \mathbf{y})^{p-1}\|_q \quad (\text{Holder's Inequality}) \\
&= \|\mathbf{x}\|_p \left[\sum (|(x_i + y_i)^{p-1}|)^q \right]^{\frac{1}{q}} + \|\mathbf{y}\|_p \left[\sum (|(x_i + y_i)^{p-1}|)^q \right]^{\frac{1}{q}} \\
&= \|\mathbf{x}\|_p \left[\sum (|x_i + y_i|^{p-1})^q \right]^{\frac{1}{q}} + \|\mathbf{y}\|_p \left[\sum (|x_i + y_i|^{p-1})^q \right]^{\frac{1}{q}} \\
&= \|\mathbf{x}\|_p \left(\sum |x_i + y_i|^p \right)^{\frac{1}{q}} + \|\mathbf{y}\|_p \left(\sum |x_i + y_i|^p \right)^{\frac{1}{q}} \quad (\text{because } q(p-1) = p) \\
&= (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p) \|\mathbf{x} + \mathbf{y}\|_p^{\frac{p}{q}}
\end{aligned}$$

So

$$\|\mathbf{x} + \mathbf{y}\|_p = \|\mathbf{x} + \mathbf{y}\|_p^{p - \frac{p}{q}} \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

■

Definition 8.1.5. (Definition of p -norm) On \mathbb{C}^n , if and only if $p > 1$, we say the p -norm is

$$\|\mathbf{v}\|_p := \left(\sum |v_i|^p \right)^{\frac{1}{p}}$$

8.2 Definition of Differential-able on Euclidean Space

Definition 8.2.1. (Little-o Notation) If we write $o(g(x)) = f(x)$, we mean

$$\lim_{g(x) \rightarrow 0} \frac{f(x)}{g(x)} = 0$$

Definition 8.2.2. (Differentiable on \mathbb{R}) We say a real-valued function $f(x)$ on \mathbb{R} is differentiable on a if

$$\exists L \in \mathbb{R}, \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = L$$

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 8.2.3. (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

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - \mathbf{J}(\mathbf{h})\|_{\mathbb{R}^m}}{\|\mathbf{h}\|_{\mathbb{R}^n}} = 0$$

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

$$\forall \epsilon, \exists \delta, \forall \mathbf{h}, \|\mathbf{h}\|_{\mathbb{R}^n} < \delta \longrightarrow \frac{\|\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - \mathbf{J}(\mathbf{h})\|_{\mathbb{R}^m}}{\|\mathbf{h}\|_{\mathbb{R}^n}} < \epsilon$$

We also say \mathbf{J} is a differential or a derivative of \mathbf{f} at \mathbf{a} .

Definition 8.2.4. (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

$$\{(\mathbf{x}, \mathbf{J}(\mathbf{x})) \in \mathbb{R}^{n+m} : \mathbf{x} \in \mathbb{R}^n\} \text{ where } \mathbf{J}(\mathbf{a}) = \mathbf{f}(\mathbf{a})$$

is the tangent plane of \mathbf{f} at \mathbf{a} .

Definition 8.2.5. (Definition of Gradient) Let $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. The gradient of f is defined as

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

Notice that we can verify Gradient is Jacobian, and jacobian is always a derivative.

Definition 8.2.6. (Jacobian Matrix) For function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, The Jacobian matrix $\mathbf{J}_{\mathbf{f}}$ is

$$\mathbf{J}_{\mathbf{f}} := \begin{bmatrix} \nabla^T f_1 \\ \vdots \\ \nabla^T f_m \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

The notation is

$$D(\mathbf{f}) = \mathbf{J} = \mathbf{J}_{\mathbf{f}} = \nabla \mathbf{f} = \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)}$$

Theorem 8.2.7. (If Differentiable, Then Jacobian give a Differential) Let $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable at \mathbf{a} . Define

$$\mathbf{J} : \mathbb{R}^n \rightarrow \mathbb{R}_m, \mathbf{x} \mapsto \mathbf{J}_f(\mathbf{a}) \cdot \mathbf{x}$$

We have

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \mathbf{J}(\mathbf{x} - \mathbf{a})\|_{\mathbb{R}^m}}{\|\mathbf{x} - \mathbf{a}\|_{\mathbb{R}^n}} = 0$$

8.3 Example Exercises for Differentiable

Theorem 8.3.1. (Example of Differentiable Function)

$f(x, y) = x^2 + 2y$ is differentiable every where

Proof. Notice

$$\mathbf{J} = [2x \ 2]$$

For each point (a, b) , we let $\mathbf{H} := \mathbf{J}(a, b) = [2a \ 2]$, and we have

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \mathbf{H}(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|_{\mathbb{R}^n}} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{(x - a)^2}{\sqrt{(x - a)^2 + (y - b)^2}}$$

Notice that we have

$$\lim_{x \rightarrow a} \frac{(x - a)^2}{\sqrt{(x - a)^2}} = 0$$

In ϵ, δ language, we have

$$\forall \epsilon, \exists \delta, \forall x, |x - a| < \delta \implies \frac{(x - a)^2}{\sqrt{(x - a)^2}} < \epsilon$$

For all ϵ , we then can pick the δ above so we have

$$\sqrt{(x - a)^2 + (y - b)^2} < \delta \implies |x - a| < \delta \implies \frac{(x - a)^2}{\sqrt{(x - a)^2 + (y - b)^2}} \leq \frac{(x - a)^2}{\sqrt{(x - a)^2}} < \epsilon$$

This prove equation 4.40 approach to 0 ■

Theorem 8.3.2. (Example of Differentiable Function) Let $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$

If \mathbf{f} is linear, then \mathbf{f} is differentiable everywhere

Proof. Let $\mathbf{H} = \mathbf{f}$, then we have

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \mathbf{H}(\mathbf{x} - \mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x} - \mathbf{a} - (\mathbf{x} - \mathbf{a}))\|}{\|\mathbf{x} - \mathbf{a}\|} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{0})\|}{\|\mathbf{x} - \mathbf{a}\|} = 0$$
■

Theorem 8.3.3. (Example of Jacobian Exists, But Non-Differentiable)

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{if } (x, y) \neq \mathbf{0} \\ 0 & \text{if } (x, y) = \mathbf{0} \end{cases} \text{ exists jacobian but non-differentiable at } \mathbf{0}$$

Proof. Notice $\mathbf{J}(0,0) = \begin{bmatrix} 0 & 0 \end{bmatrix}$. Now we prove f is non-differentiable. Observe

$$\lim_{(x,y) \rightarrow (0,0)} \frac{f(x,y) - f(0,0) - \mathbf{H}(0,0)}{\sqrt{x^2 + y^2}} = \frac{f(x,y)}{\sqrt{x^2 + y^2}} = \frac{xy}{(x^2 + y^2)^{\frac{3}{2}}}$$

We know

$$\lim_{x \rightarrow 0} \frac{x^2}{(2x^2)^{\frac{3}{2}}} = \infty$$

In ϵ, δ language, we have

$$\forall \epsilon, \exists \delta, x = \delta \implies \frac{x^2}{(2x^2)^{\frac{3}{2}}} > \epsilon$$

Then we have

$$\forall \epsilon, \exists \delta, (x, y) = \left(\frac{\delta}{\sqrt{2}}, \frac{\delta}{\sqrt{2}}\right) \implies \frac{xy}{(x^2 + y^2)^{\frac{3}{2}}} = \frac{\delta^2}{2(\delta^2)^{\frac{3}{2}}} > \epsilon$$

which prove equation 4.47 does not approach to 0. ■

Theorem 8.3.4. (Chain Rule) Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and let $g : \mathbb{R}^k \rightarrow \mathbb{R}^n$. We have

$$D_{\mathbf{a}}(f \circ g) = D_{g(\mathbf{a})}(f) \circ D_{\mathbf{a}}(g)$$

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^1$, we can write

$$h(\mathbf{x}) := f \circ g(\mathbf{x}) = f(g_1(\mathbf{x}), \dots, g_k(\mathbf{x}))$$

and we have

$$\frac{\partial h}{\partial x_i} = \sum_{1 \leq j \leq n} \frac{\partial f}{\partial g_j} \frac{\partial g_j}{\partial x_i}$$

Theorem 8.3.5. (Example of Chain Rule) Let $\delta : (0, 1) \rightarrow \mathbb{R}^n$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable. Define

$$F(t) := f(\delta(t))$$

We have

$$L_{F'(t)} = D_{\delta(t)}(f) \circ D_t(\delta)$$

Proof. Notice that $D_t(F) = L_{F'(t)}$, and we are done. ■

Theorem 8.3.6. (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

$$\frac{\partial h}{\partial x} = 2xy^2 \sin x + (x^2y^2 + 2e^x \sin x) \cos x + e^x \sin^2 x$$

Proof. We have

$$\begin{aligned}
 \frac{\partial h}{\partial x} &= \frac{\partial h}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial h}{\partial v} \frac{\partial v}{\partial x} + \frac{\partial h}{\partial w} \frac{\partial w}{\partial x} \\
 &= (2uv)(y) + (u^2 + 2wv)(\cos x) + (v^2)(e^x) \\
 &= 2xy(\sin x)y + (x^2y^2 + 2e^x \sin x) \cos x + (\sin^2 x)e^x \\
 &= 2xy^2 \sin x + (x^2y^2 + 2e^x \sin x) \cos x + e^x \sin^2 x
 \end{aligned}$$

■

Theorem 8.3.7. (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

$$f_i(\mathbf{y}) - f_i(\mathbf{x}) = (Df_i)(\mathbf{x} + c_i(\mathbf{y} - \mathbf{x}))$$

8.4 Directional Derivative

Definition 8.4.1. (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

$$(D_{\mathbf{v}}f)(\mathbf{x}_0) := \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v})}{t}$$

Theorem 8.4.2. (Directional Derivative) Let U be open and $f : U \rightarrow \mathbb{R}^n$ be differentiable at x_0 . We have

$$(D_{\mathbf{v}}f)(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v} = \mathbf{J}_f(\mathbf{x}_0) \cdot \mathbf{v}$$

Proof. Outline: in the limit of differentiable equation, let $x = x_0 + tv$ and see what happen when $t \rightarrow 0$ ■

Theorem 8.4.3. (Example) Let

$$f(x, y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{if } (x, y) \neq 0 \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

We have $f(x, y)$ is non-differentiable at v

Proof. Observe

$$\begin{aligned} D_v f(0) &= \lim_{t \rightarrow 0} \frac{f(tv_1, tv_2) - f(0, 0)}{t} \\ &= \lim_{t \rightarrow 0} \frac{v_1^3}{v_1^2 + v_2^2} = v_1^3 \neq Df(0)(v) \end{aligned}$$
■

8.5 Gradient

Definition 8.5.1. (Level Set) Let f be a scalar function. A level set is a set of points

$$L(c) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = c\}$$

Theorem 8.5.2. (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 8.5.3. (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

$$f(x, y, z) := x^2 + y^2 + z^2$$

We have

$$S = L(f(1, 1, 1))$$

Compute

$$\nabla f(1, 1, 1) = (4, 4, 4)$$

■

Theorem 8.5.4. (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

$$f(x, y, z) = x^2 - y^2 + xyz$$

We have

$$S = L(f(1, 0, 1))$$

Compute

$$\nabla f(1, 0, 1) = (2, 1, 0)$$

The tangent plane is

$$2x + y = 2$$

■

8.6 Integral

Theorem 8.6.1. (Integral) Let

$$\forall x \in S, g_1(x) \leq g_2(x)$$

Let

$$A = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : \forall x \in S, g_1(x) \leq y \leq g_2(x)\}$$

Let $f : A \rightarrow \mathbb{R}$ be continuous, we have

$$\int_A f(x, y) d(x, y) = \int_S \left(\int_{g_1(y)}^{g_2(y)} f(x, y) dy \right) dx$$

This enable us to calculate integral by iteration.

Theorem 8.6.2. (Example) Let

$$A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, x \leq y \leq 1\}$$

Find

$$\int_A xy dA$$

Proof. Compute

$$\begin{aligned} \int_A xy dA &= \int_0^1 \int_0^y xy dx dy \\ &= \int_0^1 \frac{y^3}{2} dy \\ &= \frac{1}{8} \end{aligned}$$

■

Theorem 8.6.3. (Example) Let

$$A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, \sqrt{x} \leq y \leq 1\}$$

Find

$$\int_A e^{y^3} dA$$

Proof. Compute

$$\begin{aligned}\int_A e^{y^3} dA &= \int_0^1 \int_0^{y^2} e^{y^3} dx dy \\ &= \int_0^1 y^2 e^{y^3} dy \\ &= \frac{1}{3}(e - 1)\end{aligned}$$

■

Theorem 8.6.4. (Example) Let

$$A = \{(x_1, x_2, x_3) : x_1, x_2, x_3 \geq 0, x_1 + x_2 + x_3 \leq 1\}$$

Find

$$\int_A (x_1 + x_2 + x_3)^2 dx_1 dx_2 dx_3$$

Proof.

$$\begin{aligned}\int_S (x_1 + x_2 + x_3)^2 dx_1 dx_2 dx_3 &= \int_0^1 \int_0^{1-x_3} \int_0^{1-x_3-x_2} (x_1 + x_2 + x_3)^2 dx_1 dx_2 dx_3 \\ &= \int_0^1 \int_0^{1-x_3} \frac{1}{3} - \frac{(x_2 + x_3)^3}{3} dx_2 dx_3 \\ &= \int_0^1 \frac{x_3^4}{12} - \frac{x_3}{3} + \frac{1}{4} dx_3 \\ &= \frac{1}{10}\end{aligned}$$

The answer is $\frac{4}{5}$

■

Theorem 8.6.5. (Example) Let $V(n)$ be the volume of unit n -ball. Compute $V(n)$

Proof. We first write $V(n)$ as

$$V(n) = \int_{-1}^1 \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{n-1}^2}} dx_n dx_{n-1} \cdots dx_1$$

Then we can observe

$$\begin{aligned}V(n) &= V(n-1) \int_{-1}^1 (1-x_1^2)^{\frac{n-1}{2}} dx_1 \\ &= V(n-1) \int_0^1 (1-u)^{n-1} \frac{u^{-\frac{1}{2}}}{2} du (\because u = x_1^2)\end{aligned}$$

■

Chapter 9

Ve 2

9.1 Line Integral

Theorem 9.1.1. (Computation of Arc Length with Intuition) Let $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ be a function that represent the position of a particle. Let

$$C = \{\gamma(t) : a \leq t \leq b\}$$

The length of C is then

$$\int_C dS = \int_a^b \|\gamma'(t)\| dt$$

Proof. We have

$$dS = \sqrt{(dx)^2 + (dy)^2 + (dz)^2}$$

Then

$$dS = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt = \|\gamma'(t)\| dt$$

■

Corollary 9.1.2. (Path Integral) Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a scalar field, and

$$C = \{\gamma(t) : a \leq t \leq b\}$$

We have

$$\int_C f(\gamma) dS = \int_a^b f(\gamma(t)) \times |\gamma'(t)| dt$$

Corollary 9.1.3. (Line Integral) Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a vector field, and

$$C = \{\gamma(t) : a \leq t \leq b\}$$

We have

$$\int_C F(\gamma) \cdot dS = \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt$$

Above are specification. It is not proof or anything.

Below practice some path integral.

Theorem 9.1.4. (Example) Let

$$C = \{(x, y) : x = a \cos t, a \sin t, 0 \leq t \leq \pi\}$$

We have

$$\int_C x^2 y ds = \frac{2}{3} a^4$$

Proof. Observe

$$\gamma(t) = (a \cos t, a \sin t) \implies \gamma'(t) = (-a \sin t, a \cos t) \implies |\gamma'(t)| = a$$

We have

$$\begin{aligned} \int_C f(\gamma) dS &= \int_0^\pi f(\gamma(t)) \times a dt \\ &= \int_0^\pi a^4 \cos^2 t \sin t dt \\ &= a^4 \left. \frac{-\cos^3 t}{3} \right|_{t=0}^\pi \\ &= \frac{a^4}{3} (1 + 1) = \frac{2a^4}{3} \end{aligned}$$

■

Theorem 9.1.5. (Example) Let

$$C = \{(x, y) : x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}}\}$$

We have

$$\int_C dS = 6a$$

Proof. Let $\gamma(t) = (a \cos^3 t, a \sin^3 t)$. Then

$$C = \{\gamma(t) : 0 \leq t \leq 2\pi\}$$

We have

$$\gamma'(t) = (-3a \cos^2 t \sin t, 3a \sin^2 t \cos t) \implies |\gamma'(t)| = 3a |\cos t \sin t|$$

Then

$$\begin{aligned}
 \int_C dS &= \int_0^{2\pi} 3a |\cos t \sin t| dt \\
 &= 4 \int_0^{\frac{\pi}{2}} 3a \cos t \sin t dt \\
 &= 12a \int_0^{\frac{\pi}{2}} \frac{\sin 2t}{2} dt \\
 &= 12a \left. \frac{-\cos 2t}{4} \right|_0^{\frac{\pi}{2}} \\
 &= 6a
 \end{aligned}$$

■

Theorem 9.1.6. (Example) Let

$$A = (0, 0, 0) \text{ and } B = (1, 2, 3)$$

Let C be the segment that connect A and B . We have

$$\int_C (x^2 + y^3) dS = \frac{7\sqrt{14}}{3}$$

Proof. Let

$$\gamma(t) = (t, 2t, 3t)$$

We have

$$C = \{\gamma(t) : 0 \leq t \leq 1\} \text{ and } |\gamma'(t)| = \sqrt{14}$$

Then

$$\begin{aligned}
 \int_C (x^2 + y^3) dS &= \int_0^1 (t^2 + 8t^3) \sqrt{14} dt \\
 &= \sqrt{14} \left(\frac{t^3}{3} + 2t^4 \right) \Big|_0^1 \\
 &= \sqrt{14} \times \frac{7}{3} = \frac{7\sqrt{14}}{3}
 \end{aligned}$$

■

Theorem 9.1.7. (Example) Let

$$C = \{(a \cos^3 t, a \sin^3 t) : 0 \leq t \leq \frac{\pi}{2}\}$$

We have

$$\int_C \frac{x^2 dy - y^2 dx}{x^{\frac{5}{3}} + y^{\frac{5}{3}}}$$

Notice

$$\frac{dx}{dt} = -3a \cos^2 t \sin t \text{ and } \frac{dy}{dt} = 3a \sin^2 t \cos t$$

Then

$$\begin{aligned} \int_C \frac{x^2 dy - y^2 dx}{x^{\frac{5}{3}} + y^{\frac{5}{3}}} &= \int_0^{\frac{\pi}{2}} \frac{a^2 \cos^6 t 3a \sin^2 t \cos t - a^2 \sin^6 t (-3a) \cos^2 t \sin t}{a^{\frac{5}{3}} (\cos^5 t + \sin^5 t)} dt \\ &= 3a^{\frac{4}{3}} \int_0^{\frac{\pi}{2}} \frac{\cos^7 t \sin^2 t + \sin^7 t \cos^2 t}{\cos^5 t + \sin^5 t} dt \\ &= 3a^{\frac{4}{3}} \int_0^{\frac{\pi}{2}} \sin^2 t \cos^2 t dt \\ &= 3a^{\frac{4}{3}} \int_0^{\frac{\pi}{2}} \left(\frac{\sin 2t}{2}\right)^2 dt \\ &= \frac{3}{4} a^{\frac{4}{3}} \int_0^{\frac{\pi}{2}} \sin^2 2t dt \\ &= \frac{3a^{\frac{4}{3}}}{4} \int_0^{\frac{\pi}{2}} \frac{1 - \cos 4t}{2} dt \\ &= \frac{3a^{\frac{4}{3}}}{8} \int_0^{\frac{\pi}{2}} 1 - \cos 4t dt \\ &= \frac{3a^{\frac{4}{3}}}{8} \left(\frac{-\sin 4t}{4} + t \right) \Big|_0^{\frac{\pi}{2}} \\ &= \frac{3a^{\frac{4}{3}}}{8} \left(\frac{\pi}{2} \right) = \frac{3\pi a^{\frac{4}{3}}}{16} \end{aligned}$$

Below practice line integral.

Theorem 9.1.8. (Example) Let

$$F(x, y, z) = (xy, 2, 4z) \text{ and } \gamma(t) = (\cos t, \sin t, t)$$

Let

$$C = \{\gamma(t) : 0 \leq t \leq 2\pi\}$$

We have

$$\int_C F(\gamma) \cdot dS = 8\pi^2$$

Proof. We have

$$\gamma'(t) = (-\sin t, \cos t, 1)$$

Compute

$$\begin{aligned} \int_C F(\gamma) \cdot dS &= \int_0^{2\pi} F(\cos t, \sin t, t) \cdot (-\sin t, \cos t, 1) dt \\ &= \int_0^{2\pi} (\cos t \sin t, 2, 4t) \cdot (-\sin t, \cos t, 1) dt \\ &= \int_0^{2\pi} -\cos t \sin^2 t + 2 \cos t + 4t dt \\ &= \left. \frac{-\sin^3 t}{3} + 2 \sin t + 2t^2 \right|_0^{2\pi} = 8\pi^2 \end{aligned}$$

■

Theorem 9.1.9. (Example) Let

$$F(x, y) = (e^y, -\sin \pi x) \text{ and } C = \{(-1, 0) \rightarrow (1, 0)\} \cup \{(1, 0) \rightarrow (0, 1)\} \cup \{(0, 1) \rightarrow (-1, 0)\}$$

We have

$$\int_C F(\gamma) \cdot dS = 4 - 2e - \frac{4}{\pi}$$

Proof. Define

$$\gamma_1(t) = (-1 + 2t, 0) \text{ and } \gamma_2(t) = (1 - t, t) \text{ and } \gamma_3(t) = (-t, 1 - t)$$

We have

$$\begin{aligned} \int_{C_1} F(\gamma_1) \cdot dS &= \int_0^1 (1, -\sin(\pi(-1 + 2t))) \cdot (2, 0) dt \\ &= \int_0^1 2 dt = 2 \end{aligned}$$

We have

$$\begin{aligned} \int_{C_2} F(\gamma_2) \cdot dS &= \int_0^1 (e^t, -\sin \pi(1 - t)) \cdot (-1, 1) dt \\ &= \int_0^1 -e^t - \sin \pi(1 - t) dt \\ &= -e^t - \frac{1}{\pi} \cos \pi(1 - t) \Big|_0^1 \\ &= -e + 1 - \frac{2}{\pi} \end{aligned}$$

We have

$$\begin{aligned}
 \int_{C_3} F(\gamma_3) \cdot dS &= \int_0^1 (e^{1-t}, \sin \pi t) \cdot (-1, -1) dt \\
 &= - \int_0^1 e^{1-t} + \sin \pi t dt \\
 &= e^{1-t} + \frac{\cos \pi t}{\pi} \Big|_0^1 \\
 &= (1 - e) + \frac{-1 - 1}{\pi} \\
 &= -e + 1 - \frac{2}{\pi}
 \end{aligned}$$

This give us

$$\begin{aligned}
 \int_C F(\gamma) \cdot dS &= \int_{C_1} F(\gamma_1) \cdot dS + \int_{C_2} F(\gamma_2) \cdot dS + \int_{C_3} F(\gamma_3) \cdot dS \\
 &= 2 + 2(-e + 1 - \frac{2}{\pi}) = 4 - 2e - \frac{4}{\pi}
 \end{aligned}$$

■

Theorem 9.1.10. (Example) Let

$$F(x, y, z) = (xy, yz, xz) \text{ and } \gamma(t) = (t, t^2, t^3)$$

Let

$$C = \{\gamma(t) : 0 \leq t \leq 1\}$$

We have

$$\int_C F(\gamma) \cdot dS = \frac{27}{28}$$

Proof. Notice

$$\gamma'(t) = (1, 2t, 3t^2)$$

Then

$$\begin{aligned}\int_C F(\gamma) \cdot dS &= \int_0^1 F(t, t^2, t^3) \cdot (1, 2t, 3t^2) dt \\&= \int_0^1 (t^3, t^5, t^4) \cdot (1, 2t, 3t^2) dt \\&= \int_0^1 5t^6 + t^3 dt \\&= \frac{5}{7}t^7 + \frac{t^4}{4} \Big|_{t=0}^1 \\&= \frac{27}{28}\end{aligned}$$

■

Theorem 9.1.11. (Example) Let

$$F(x, y, z) = (y^2 + z^2, x^2 + z^2, x^2 + y^2) \text{ and } \gamma_1(t) = (t, t, t) \text{ and } \gamma_2(t) = (t, t^2, t^3)$$

Let

$$C_1 = \{\gamma_1(t) : 0 \leq t \leq 1\} \text{ and } C_2 = \{\gamma_2(t) : 0 \leq t \leq 1\}$$

We have

$$\int_{C_1} F(\gamma_1) \cdot dS = 2$$

We have

$$\int_{C_2} F(\gamma_2) \cdot dS = \frac{297}{140}$$

Proof. We have

$$\gamma_1'(t) = (1, 1, 1) \text{ and } \gamma_2'(t) = (1, 2t, 3t^2)$$

Then

$$\begin{aligned}\int_{C_1} F(\gamma_1) \cdot dS &= \int_0^1 F(\gamma_1(t)) \cdot \gamma_1'(t) dt \\&= \int_0^1 F(t, t, t) \cdot (1, 1, 1) dt \\&= \int_0^1 (2t^2, 2t^2, 2t^2) \cdot (1, 1, 1) dt \\&= \int_0^1 6t^2 dt \\&= 2t^3 \Big|_{t=0}^1 = 2\end{aligned}$$

Then

$$\begin{aligned}\int_{C_2} F(\gamma_2) \cdot dS &= \int_0^1 F(\gamma_2(t)) \cdot \gamma_2'(t) dt \\&= \int_0^1 F(t, t^2, t^3) \cdot (1, 2t, 3t^2) dt \\&= \int_0^1 (t^4 + t^6, t^2 + t^6, t^2 + t^4) \cdot (1, 2t, 3t^2) dt \\&= \int_0^1 t^4 + t^6 + 2t^3 + 2t^7 + 3t^4 + 3t^6 dt \\&= \int_0^1 2t^7 + 4t^6 + 4t^4 + 2t^3 dt \\&= \left. \frac{t^8}{4} + \frac{4t^7}{7} + \frac{4t^5}{5} + \frac{t^4}{2} \right|_{t=0}^1 \\&= \frac{35 + 80 + 112 + 70}{140} = \frac{297}{140}\end{aligned}$$

■

9.2 Fundamental Theorem of Line Integral

Theorem 9.2.1. (Fundamental Theorem of Calculus) Let $C = \{\gamma(t) : t \in [a, b]\}$ be a piece-wise smooth curve. If f is differential, then

$$\int_C \nabla f(\gamma) \cdot d\gamma = f(\gamma(b)) - f(\gamma(a))$$

Proof. Notice the multivariate chain rule

$$\frac{d}{dt}f \circ \gamma = \nabla f(\gamma(t)) \cdot \gamma'(t)$$

Then

$$\begin{aligned} \int_C \nabla f(\gamma) \cdot d\gamma &= \int_a^b \nabla f(\gamma(t)) \cdot \gamma'(t) dt \\ &= \int_a^b \frac{d}{dt}f(\gamma(t)) dt \\ &= f(\gamma(b)) - f(\gamma(a)) \end{aligned}$$

■

Definition 9.2.2. (Definition of Conservative Vector Field) We say a vector field F is conservative if there exists a scalar field f such that

$$F = \nabla f$$

Definition 9.2.3. (Definition of Vector Field Being Independent of Path) We say a vector field F is independent of path if for each pair of paths C_1, C_2 of same starting and ending points a, b , we have

$$\int_{C_1} F(\gamma) \cdot d\gamma = \int_{C_2} F(\gamma) \cdot d\gamma$$

Theorem 9.2.4. (Independent of Path Theorem)

$$F \text{ is conservative} \iff F \text{ is independent of path}$$

Proof. (\longrightarrow)

This follows from Theorem 9.2.1 (Fundamental Theorem of Line Integral)

(\longleftarrow)

Define

$$f(0) = 0 \text{ and } f(s) = \int_0^s F(\gamma) \cdot d\gamma$$

We now see

$$\frac{\partial f}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x + h(1, 0, 0)) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{\int_0^{x+h(1,0,0)} F(\gamma) \cdot d\gamma - \int_0^x F(\gamma) \cdot d\gamma}{h}$$

Because F is independent of path, we have

$$\int_0^{x+h(1,0,0)} F(\gamma) \cdot d\gamma - \int_0^x F(\gamma) \cdot d\gamma = \int_x^{x+h(1,0,0)} F(\gamma) \cdot d\gamma$$

By changing the path of $0 \rightarrow (x + h(1, 0, 0))$ to $0 \rightarrow x \cup x \rightarrow (x + h(1, 0, 0))$. Then we have

$$\begin{aligned} \frac{\partial f}{\partial x} &= \lim_{h \rightarrow 0} \frac{\int_x^{x+h(1,0,0)} F(\gamma) \cdot d\gamma}{h} \\ &= \lim_{h \rightarrow 0} \frac{\int_0^1 F(x + t(1, 0, 0)) \cdot (h, 0, 0) dt}{h} \\ &= \lim_{h \rightarrow 0} \frac{\int_0^1 h F_1(x + (th, 0, 0)) dt}{h} \\ &= \lim_{h \rightarrow 0} \int_0^1 F_1(x + (th, 0, 0)) dt = F_1(x) \end{aligned}$$

■

Theorem 9.2.5. (Formula for Line Integral) Let

$$F = (P, Q, R) \text{ and } \gamma(t) = (x(t), y(t), z(t))$$

We have

$$\int_C F(\gamma) \cdot d\gamma = \int_a^b P(\gamma(t))x'(t) + Q(\gamma(t))y'(t) + R(\gamma(t))z'(t)dt$$

Below practice formula for line integral.

Theorem 9.2.6. (Example) Let

$$\gamma(t) = (1 - t, t) \text{ and } C = \{\gamma(t) : t \in [0, 1]\}$$

We have

$$\int_C x^2 y dx + x y dy = \frac{1}{12}$$

Proof. We have

$$\frac{dx}{dt} = -1 \text{ and } \frac{dy}{dt} = 1$$

Then

$$\begin{aligned} \int_C x^2 y dx + xy dy &= \int_0^1 (1-t)^2 t (-1) + (1-t)t dt \\ &= \int_0^1 -t^3 + t^2 dt \\ &= \left. \frac{-t^4}{4} + \frac{t^3}{3} \right|_0^1 \\ &= \frac{1}{12} \end{aligned}$$

■

Theorem 9.2.7. (Example) Let

$$\gamma_1(t) = (3t, 2) \text{ and } \gamma_2(t) = (3, 2 + 3t) \text{ and } \gamma_3(t) = (3 - 3t, 5 - 3t)$$

Let

$$C_1 = \{\gamma_1(t) : t \in [0, 1]\} \text{ and } C_2 = \{\gamma_2(t) : t \in [0, 1]\} \text{ and } C_3 = \{\gamma_3(t) : t \in [0, 1]\}$$

We have

$$\int_C xy^2 dx + xy^2 dy$$

Proof. Notice

$$\frac{dx}{dt} = 3 \text{ or } 0 \text{ or } -3 \text{ and } \frac{dy}{dt} = 0 \text{ or } 3 \text{ or } -3$$

We have

$$\begin{aligned} \int_{C_1} xy^2 dx + xy^2 dy &= \int_0^1 3t \cdot 2^2 \cdot 3 + 3t \cdot 2^2 \cdot 0 dt \\ &= \int_0^1 36t dt = 18t^2 \Big|_0^1 = 18 \end{aligned}$$

We have

$$\begin{aligned} \int_{C_2} xy^2 dx + xy^2 dy &= \int_0^1 3(2 + 3t)^2 \cdot 3 dt \\ &= (2 + 3t)^3 \Big|_0^1 = 117 \end{aligned}$$

We have

$$\begin{aligned}
 \int_{C_3} xy^2 dx + xy^2 dy &= \int_0^1 -6(3-3t)(5-3t)^2 dt \\
 &= \int_0^1 -18(-9t^3 + 39t^2 - 55t + 25) dt \\
 &= -18 \left(\frac{-9}{4} t^4 + 13t^3 - \frac{55}{2} t^2 + 25t \right) \Big|_0^1 = \frac{-297}{2}
 \end{aligned}$$

■

Lastly

Theorem 9.2.8. (Curl = 0 If and Only If Conservative)

$$\forall s, \operatorname{curl} F(s) = 0 \iff F \text{ is conservative}$$

Proof. (\leftarrow)

Let f satisfy

$$\nabla f = F$$

We have

$$\begin{aligned}
 \operatorname{curl} F &= \nabla \times F \\
 &= \nabla \times (\nabla f) \\
 &= \left| \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_x & f_y & f_z \end{bmatrix} \right| \\
 &= (f_{yz} - f_{zy}, f_{xz} - f_{zx}, f_{yx} - f_{xy}) = (0, 0, 0)
 \end{aligned}$$

(\rightarrow)

Let Σ be a smooth oriented surface in \mathbb{R}^3 . Let Γ be the boundary of Σ . Stokes' Theorem give us

$$\int \int_{\Sigma} \nabla \times F d\Sigma = \int_{\Gamma} F \cdot dS$$

Because curl equals 0, we see F is independent of the path by observing seeing

$$\int_{\Gamma} F \cdot dS = 0$$

By viewing Γ to be the union of two different path with one of them reversed.

■

Theorem 9.2.9. (Test Example) Let

$$F = (4x^3 + 9x^2y^2, 6x^3y + 6y^5)$$

We have

$$f = x^4 + 3x^3y^2 + y^5 + C \implies F = \nabla f$$

Proof. We see

$$F_X = 4x^3 + 9x^2y^2 \implies f = x^4 + 3x^3y^2 + C$$

and

$$F_Y = 6x^3y + 6y^5 \implies f = y^5 + 3x^3y^2 + C$$

Then

$$f = x^4 + 3x^3y^2 + y^5 + C$$

■

9.3 Green's Theorem

Definition 9.3.1. (Definition of Curve) By a curve C , we mean a continuous function p from $I \subseteq \mathbb{R}$ to a topological space.

Definition 9.3.2. (Definition of Simple Curve) We say a curve C is simple if p is one-to-one.

Definition 9.3.3. (Definition of Closed Curve) We say a curve C is closed if

$$C = p[I] \text{ and } p(a) = p(b) \text{ where } I = [a, b]$$

Definition 9.3.4. (Definition of Simple Closed Curve) We say a curve C is simple closed if

$$C = p[I] \text{ and } p(a) = p(b) \text{ and } I = [a, b] \text{ and } p$$

Definition 9.3.5. (Definition of Simple Region) We say a connected region $\Omega \subseteq \mathbb{R}^2$ is simple if Ω is both x -simple and y -simple. By Ω is y -simple, we mean

$$\Omega = \{(x, y) : x \in [a, b] \text{ and } y \in [f_1(x), f_2(x)]\}$$

Similarly, by Ω is x -simple, we mean

$$\Omega = \{(x, y) : y \in [a, b] \text{ and } x \in [f_1(y), f_2(y)]\}$$

Theorem 9.3.6. (Green's Theorem) Let Ω be a simple closed region with smooth boundary. Let P and $Q : \mathbb{R}^2 \rightarrow \mathbb{R}$ be differentiable. Let C be counterclockwise boundary of Ω . We have

$$\iint_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_C P dx + Q dy$$

In other words, let $F = (P, Q)$, we have

$$\iint_{\Omega} \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ P & Q \end{vmatrix} dA = \int_C F(\gamma) \cdot d\gamma$$

Proof. We wish to prove

$$\iint_{\Omega} \frac{\partial Q}{\partial x} dx dy = \int_C Q dy \text{ and } \iint_{\Omega} \frac{-\partial P}{\partial y} dx dy = \int_C P dx$$

Because Ω is y -simple. We can characterize Ω as

$$\Omega = \{(x, y) : x \in [a, b] \text{ and } y \in [g(x), f(x)]\}$$

Then

$$\begin{aligned}\iint_{\Omega} \frac{\partial P}{\partial y} dA &= \int_a^b \int_{g(x)}^{f(x)} \frac{\partial P}{\partial y} dy dx \\ &= \int_a^b P(x, f(x)) - P(x, g(x)) dx\end{aligned}$$

Partition C into $C_1 \cup C_2 \cup C_3 \cup C_4$. Clearly

$$\int_{C_2} P dx = 0 = \int_{C_4} P dx$$

We have

$$\int_{C_1} P dx = \int_a^b P(x, g(x)) dx \text{ and } \int_{C_3} P dx = \int_b^a P(x, f(x)) dx$$

Then

$$\int_C P dx = \int_a^b P(x, g(x)) - P(x, f(x)) dx = - \iint_{\Omega} \frac{\partial P}{\partial y} dA$$

It remains to prove

$$\iint_{\Omega} \frac{\partial Q}{\partial x} dx dy = \int_C Q dy$$

Because Ω is x -simple. We can characterize Ω as

$$\Omega = \{(x, y) : y \in [a, b] \text{ and } x \in [g(y), f(y)]\}$$

Then

$$\begin{aligned}\iint_{\Omega} \frac{\partial Q}{\partial x} dA &= \int_a^b \int_{g(y)}^{f(y)} \frac{\partial Q}{\partial x} dx dy \\ &= \int_a^b Q(f(y), y) - Q(g(y), y) dy\end{aligned}$$

Clearly

$$\int_{C_1} Q dy = 0 = \int_{C_3} Q dy$$

We have

$$\int_{C_2} Q dy = \int_a^b Q(f(y), y) dy \text{ and } \int_{C_4} Q dy = \int_b^a Q(g(y), y) dy$$

Then

$$\int_C Q dy = \int_a^b Q(f(y), y) - Q(g(y), y) dy = \iint_{\Omega} \frac{\partial Q}{\partial x} dA$$



Corollary 9.3.7. Let $F = (P, Q)$ be a \mathbb{R}^2 vector field. We have

$$F \text{ is conservative} \iff \frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$$

Proof. (\leftarrow)

Let C_1 and C_2 be two curve from a to b . We wish to show

$$\int_{C_1} Pdx + Qdy = \int_{C_2} Pdx + Qdy$$

We see

$$\int_{C_1 \cup -C_2} Pdx + Qdy = \iint_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = 0$$

(\rightarrow)

Let

$$F = \nabla f$$

We have

$$P = f_x \text{ and } Q = f_y$$

Then we see

$$\frac{\partial P}{\partial y} = f_{xy} = f_{yx} = \frac{\partial Q}{\partial x}$$

■

We now practice the usage of Green's Theorem.

Theorem 9.3.8. (Example) Let

$$a > 0 \text{ and } C = \{(x, y) : x^2 + y^2 = a^2\}$$

Compute

$$\int_C (3x^2 + y)dx + (2x + y^2)dy$$

Proof. We have

$$\begin{aligned} \int_C (3x^2 + y)dx + (2x + y^2)dy &= \iint_{\Omega} \frac{\partial 2x + y^2}{\partial x} - \frac{\partial 3x^2 + y}{\partial y} dA \\ &= \iint_{\Omega} dA = a^2 \pi \end{aligned}$$

■

Below practice how to compute area of simple region using Green's Theorem.

Theorem 9.3.9. (Area Calculation) Let Ω be a simple closed region with piece-wise smooth curve C . Then

$$A(\Omega) = \iint_{\Omega} dA = \frac{1}{2} \int_C xdy - ydx$$

Theorem 9.3.10. (Example) Let

$$\Omega = \{(x, y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\}$$

We have

$$A(\Omega) = ab\pi$$

Proof.

$$C = \{(x, y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$$

Let

$$\gamma(t) = (a \cos t, b \sin t)$$

We have

$$\begin{aligned} A(\Omega) &= \iint_{\Omega} dA \\ &= \frac{1}{2} \int_C xdy - ydx \\ &= \frac{1}{2} \int_0^{2\pi} ab \cos^2 t + ab \sin^2 t dt \\ &= \frac{1}{2} \int_0^{2\pi} ab dt \\ &= ab\pi \end{aligned}$$

■

Theorem 9.3.11. (Example) Let C be a simple closed curve.

$$\int_C \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = \begin{cases} 0 & \text{if } (0, 0) \notin \Omega \\ 2\pi & \text{if } (0, 0) \in \Omega \end{cases}$$

Proof. Let

$$P = \frac{-y}{x^2 + y^2} \text{ and } Q = \frac{x}{x^2 + y^2}$$

Notice that P_y, Q_x exists if and only if $(x, y) \neq (0, 0)$.

Then if $(0, 0) \notin \Omega$, by Green's Theorem, we have

$$\begin{aligned} \int_C \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy &= \int_C P dx + Q dy \\ &= \iint_{\Omega} \frac{y^2 - x^2}{(x^2 + y^2)^2} - \frac{y^2 - x^2}{(x^2 + y^2)^2} dA = 0 \end{aligned}$$

If $(0, 0) \in \Omega$, we draw a circle of radius r centering $(0, 0)$ contained by Ω . Draw two segments, one starting from $(r, 0)$ and head straight right, while the other start from $(-r, 0)$ and head straight left. Let O_r represent the clockwise boundary of the circle. We have

$$C = C_1 + C_2 - O_r$$

Because C_1 and C_2 doesn't enclose $(0, 0)$, we have

$$\int_{C_1} P dx + Q dy = 0 = \int_{C_2} P dx + Q dy$$

Then

$$\begin{aligned} \int_C \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy &= \int_{O_r} F(\gamma) \cdot d\gamma \\ &= \int_0^{2\pi} (-\sin t, \cos t) \cdot (-\sin t, \cos t) dt \\ &= \int_0^{2\pi} 1 dt = 2\pi \end{aligned}$$

■

Theorem 9.3.12. (Example) Let

$$C = (0, 0) \rightarrow (1, 0) + (1, 0) \rightarrow (0, 1) + (0, 1) \rightarrow (0, 0)$$

We have

$$\int_C xy^2 dx + (x^2 + y) dy = \frac{1}{4}$$

Proof. Let $P = xy^2$ and $Q = x^2 + y$. Notice that $Q_x = 2x$ and $P_y = 2xy$. We have

$$\begin{aligned}
 \int_C xy^2 dx + (x^2 + y) dy &= \iint_{\Omega} Q_x - P_y dA \\
 &= \iint_{\Omega} 2x(1 - y) dA \\
 &= \int_0^1 \int_0^{1-x} 2x(1 - y) dy dx \\
 &= \int_0^1 2x \left. \frac{-y^2 + 2y}{2} \right|_0^{1-x} dx \\
 &= \int_0^1 2x \frac{-(1-x)^2 + 2(1-x)}{2} dx \\
 &= \int_0^1 x(-x^2 + 1) dx \\
 &= \int_0^1 -x^3 + x dx \\
 &= \left. \frac{-x^4}{4} + \frac{x^2}{2} \right|_{x=0}^1 \\
 &= \frac{1}{4}
 \end{aligned}$$

■

Theorem 9.3.13. (Example) Let

$$C = \{(x, y) : x^2 + y^2 = a^2\}$$

We have

$$\int_C (2y + \sqrt{9 + x^2}) dx + (5x + e^{\tan^{-1} y}) dy$$

Proof. Let

$$P = 2y + \sqrt{9 + x^2} \text{ and } Q = 5x + e^{\tan^{-1} y}$$

We have

$$Q_x = 5 \text{ and } P_y = 2$$

Then

$$\begin{aligned}\int_C (2y + \sqrt{9 + x^2})dx + (5x + e^{\tan^{-1}y})dy &= \iint_{\Omega} Pdx + Qdy \\ &= \iint_{\Omega} Q_x - P_y dA \\ &= \iint_{\Omega} 3dA \\ &= 3a^2\pi\end{aligned}$$

■

Theorem 9.3.14. (Example) Let C be the boundary of $\{(x, y) : x < y < x^2\}$. We have

$$\int_C (2xy - x^2)dx + (x + y^2)dy$$

Proof. Let

$$P = 2xy - x^2 \text{ and } Q = x + y^2$$

We have

$$Q_x = 1 \text{ and } P_y = 2x$$

Then

$$\begin{aligned}\int_C (2xy - x^2)dx + (x + y^2)dy &= \int_C Pdx + Qdy \\ &= \iint_{\Omega} Q_x - P_y dA \\ &= \iint_{\Omega} 1 - 2x dA \\ &= \int_0^1 \int_{x^2}^x 1 - 2x dy dx \\ &= \int_0^1 (1 - 2x)(x^2 - x) dx \\ &= \int_0^1 -2x^3 + 3x^2 - x dx \\ &= \left. \frac{-x^4}{2} + x^3 - \frac{x^2}{2} \right|_{x=0}^1 = 0\end{aligned}$$

■

Theorem 9.3.15. (Example) Let

$$\gamma(t) = \left(\frac{3at}{1+t^3}, \frac{3at^2}{1+t^3} \right)$$

This is the parametrization of

$$x^3 + y^3 = 3axy$$

where

$$t = \frac{y}{x}$$

Find area of Ω .

Proof. We have

$$dx = \frac{3a(1+t^3) - 9at^3}{(1+t^3)^2} dt \text{ and } dy = \frac{6at(1+t^3) - 9at^4}{(1+t^3)^2} dt$$

$$\begin{aligned} \iint_{\Omega} dA &= \frac{1}{2} \int_C -ydx + xdy \\ &= \frac{1}{2} \int_0^{\infty} \frac{-3at^2}{1+t^3} \times \frac{3a(1+t^3) - 9at^3}{(1+t^3)^2} + \frac{3at}{1+t^3} \times \frac{6at(1+t^3) - 9at^4}{(1+t^3)^2} dt \\ &= \frac{1}{2} \int_0^{\infty} \frac{9a^2(t^2+t^5)}{(1+t^3)^3} dt \\ &= \frac{9a^2}{2} \int_0^{\infty} \frac{t^2}{(1+t^3)^2} dt \\ &= \frac{9a^2}{2} \int_1^{\infty} \frac{1}{3u^2} du \text{ use } u = 1+t^3 \\ &= \frac{9a^2}{2} \times \frac{1}{-3u} \Big|_{u=1}^{\infty} \\ &= \frac{3a^2}{2} \end{aligned}$$

■

Chapter 10

Ve 3

10.1 Flux and Surface Integral

Theorem 10.1.1. (Normal Vector of a Surface) Consider the surface

$$S := \{\mathbf{r}(u, v) : (u, v) \in \Omega\}$$

The unit normal vector $\mathbf{n}(\mathbf{r}(u, v))$ of S at $\mathbf{r}(u, v)$ is

$$\mathbf{n}(u, v) = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}$$

Proof. See Theorem 8.5.2 ■

Definition 10.1.2. (Definition of Surface Integral) Given

- (a) $S = \{\mathbf{r}(u, v) : (u, v) \in \Omega\}$
- (b) $\mathbf{F} = (P, Q, R)$ is a vector field

We define

$$\iint_S f(x, y, z) dS := \iint_{\Omega} f(\mathbf{r}(u, v)) |\mathbf{r}_u \times \mathbf{r}_v| du dv$$

We define

$$\begin{aligned} \iint_S \mathbf{F} \cdot d\mathbf{S} &:= \iint_S \mathbf{F} \cdot \mathbf{n} dS \\ &= \iint_{\Omega} \mathbf{F} \cdot \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|} |\mathbf{r}_u \times \mathbf{r}_v| du dv \\ &= \iint_{\Omega} \mathbf{F} \cdot \mathbf{r}_u \times \mathbf{r}_v du dv \\ &= \iint_{\Omega} \begin{vmatrix} P & Q & R \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} du dv \end{aligned}$$

In short

$$\frac{dS}{dudv} = |\mathbf{r}_u \times \mathbf{r}_v|$$

Above is important, please remind.

Theorem 10.1.3. (Parametrization Formula of $z(x, y)$) Define

$$\mathbf{r}(x, y) = (x, y, z(x, y))$$

We have

$$|\mathbf{r}_x \times \mathbf{r}_y| = \sqrt{1 + z_x^2 + z_y^2}$$

Proof. Compute

$$\begin{aligned} |\mathbf{r}_x \times \mathbf{r}_y| &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & z_x \\ 0 & 1 & z_y \end{vmatrix} \\ &= |(-z_x, -z_y, 1)| = \sqrt{1 + z_x^2 + z_y^2} \end{aligned}$$

■

Theorem 10.1.4. (Example of Computing Surface Area with $z(x, y)$) Find the surface area of sphere of radius a .

The surface area of sphere of radius a is $4a^2\pi$

Proof. We have

$$z = \sqrt{a^2 - x^2 - y^2}$$

Then

$$f_x = \frac{-x}{\sqrt{a^2 - x^2 - y^2}} \text{ and } f_y = \frac{-y}{\sqrt{a^2 - x^2 - y^2}}$$

Then

$$\sqrt{1 + f_x^2 + f_y^2} = \sqrt{\frac{a^2}{a^2 - x^2 - y^2}} = \frac{a}{\sqrt{a^2 - (x^2 + y^2)}}$$

This give us

$$\begin{aligned}
 \iint_{\Omega} \sqrt{1 + f_x^2 + f_y^2} dA &= \iint_{\Omega} \frac{a}{\sqrt{a^2 - (x^2 + y^2)}} dA \\
 &= \int_0^a \int_0^{2\pi} \frac{ar}{\sqrt{a^2 - r^2}} d\theta dr \\
 &= 2a\pi \int_0^a \frac{r}{\sqrt{a^2 - r^2}} dr \\
 &= -2a\pi \sqrt{a^2 - r^2} \Big|_{r=0}^a \\
 &= 2a^2\pi
 \end{aligned}$$

■

Below are 6 example, with one left unsolved. Notice the first example use another way to find the unit normal vector \mathbf{n}

Lemma 10.1.5. Given

(a) $S = \{(x, y, z) : f(x, y, z) = \alpha\}$

The unit normal vector $\mathbf{n}(x, y, z)$ of S is

$$\frac{\nabla f(x, y, z)}{|\nabla f(x, y, z)|}$$

Theorem 10.1.6. (Example) Let

(a) $\mathbf{A} = (xy, -x^2, x + z)$

(b) $S = \{(x, y, z) : 2x + 2y + z = 6 \text{ and } x, y, z \geq 0\}$

We have

$$\iint_S \mathbf{A} \cdot \mathbf{n} dS = \frac{27}{4}$$

Proof. Let $f(x, y, z) = 2x + 2y + z$. We have

$$\nabla f = (2, 2, 1)$$

Then

$$\mathbf{n} = \frac{(2, 2, 1)}{3}$$

Let A be on x, y plane. Let $z = 6 - 2x - 2y$. We have

$$\frac{dS}{dA} = \sqrt{1 + z_x^2 + z_y^2} = \sqrt{1 + 2(-2)^2} = 3$$

Then

$$\begin{aligned}
\iint_S \mathbf{A} \cdot \mathbf{n} dS &= \frac{1}{3} \iint_{\Omega} 2xy - 2x^2 + (x + z) dS \\
&= \iint_{\Omega} 2xy - 2x^2 + (6 - x - 2y) dA \quad (\because 2x + 2y + z = 6) \\
&= \int_0^3 \int_0^{3-x} 2xy - 2x^2 - x - 2y + 6 dy dx \\
&= \int_0^3 xy^2 - 2x^2y - xy - y^2 + 6y \Big|_{y=0}^{3-x} dx \\
&= \int_0^3 x(3-x)^2 - 2x^2(3-x) - x(3-x) - (3-x)^2 + 6(3-x) dx \\
&= \int_0^3 3x^3 - 12x^2 + 6x + 9 dx \\
&= \frac{3x^4}{4} - 4x^3 + 3x^2 + 9x \Big|_{x=0}^3 \\
&= \frac{27}{4}
\end{aligned}$$

■

Theorem 10.1.7. (Example) Given

(a) $S = \{\mathbf{r}(r, \theta) : r \in [0, 1], \theta \in [0, 2\pi]\}$

(b) $\mathbf{r}(r, \theta) = (r \cos \theta, r \sin \theta, \theta)$

We have

$$\iint_S \sqrt{x^2 + y^2 + 1} dS = \frac{8}{3}\pi$$

Proof. We have

$$\mathbf{r}_r \times \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos \theta & \sin \theta & 0 \\ -r \sin \theta & r \cos \theta & 1 \end{vmatrix} = (\sin \theta, -\cos \theta, r)$$

Then

$$|\mathbf{r}_r \times \mathbf{r}_\theta| = \sqrt{r^2 + 1}$$

We now have

$$\begin{aligned}
 \iint_S \sqrt{x^2 + y^2 + 1} dS &= \iint_{\Omega} \sqrt{r^2 + 1} \sqrt{r^2 + 1} dr d\theta \\
 &= \int_0^{2\pi} \int_0^1 r^2 + 1 dr d\theta \\
 &= 2\pi \int_0^1 r^2 + 1 dr \\
 &= 2\pi \left(\frac{r^3}{3} + r \right) \Big|_{r=0}^1 = \frac{8}{3}\pi
 \end{aligned}$$

■

Theorem 10.1.8. (Example) Given surface S in 3-Euclidean Space

$$S = \{(x, y, z) : (x^2 + y^2 + z^2)^2 = x^2 - y^2\}$$

We have

$$A(S) = \iint_S dS =$$

Proof. Consider Spherical Coordinate

(a) $x = r \sin \phi \cos \theta$

(b) $y = r \sin \phi \sin \theta$

(c) $z = r \cos \phi$

We have

$$\begin{aligned}
 S &= \{(x, y, z) : r^4 = r^2 \sin^2 \phi (\cos^2 \theta - \sin^2 \theta)\} \\
 &= \{(x, y, z) : r^2 = \sin^2 \phi \cos 2\theta\}
 \end{aligned}$$

■

Theorem 10.1.9. (Example) Given

(a) $S = \{(x, y, z) : x^2 + y^2 + z^2 = t^2\}$

(b) $f(x, y, z) = \begin{cases} x^2 + y^2 & \text{if } z \geq \sqrt{x^2 + y^2} \\ 0 & \text{if } z < \sqrt{x^2 + y^2} \end{cases}$

We have

$$\iint_S f(x, y, z) dS = \frac{8 - 5\sqrt{2}}{6} t^4$$

Proof. Notice that

$$(x, y, z) \in S \text{ and } f(x, y, z) > 0 \implies z = \sqrt{t^2 - x^2 - y^2}$$

We can just parametrize S with r, θ . Suppose

$$\mathbf{r}(r, \theta) = (r \cos \theta, r \sin \theta, \sqrt{t^2 - r^2})$$

We have

$$\begin{aligned} |\mathbf{r}_r \times \mathbf{r}_\theta| &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos \theta & \sin \theta & \frac{-r}{\sqrt{t^2 - r^2}} \\ -r \sin \theta & r \cos \theta & 0 \end{vmatrix} \\ &= \left| \left(\frac{r^2 \cos \theta}{\sqrt{t^2 - r^2}}, \frac{-r^2 \sin \theta}{\sqrt{t^2 - r^2}}, r \right) \right| \\ &= \sqrt{\frac{r^4}{t^2 - r^2} + r^2} \\ &= \frac{tr}{\sqrt{t^2 - r^2}} \end{aligned}$$

Compute the boundary

$$\begin{aligned} z &= \sqrt{t^2 - x^2 - y^2} \geq \sqrt{x^2 + y^2} \\ \iff t^2 &\geq 2x^2 + 2y^2 \\ \iff \sqrt{x^2 + y^2} &\leq \frac{t}{\sqrt{2}} \\ \iff r &\leq \frac{r}{\sqrt{2}} \end{aligned}$$

Then

$$\begin{aligned}
\iint_S f(x, y, z) dS &= \iint_{\Omega} r^2 \times \frac{tr}{\sqrt{t^2 - r^2}} dr d\theta \\
&= \int_0^{2\pi} \int_0^{\frac{t}{\sqrt{2}}} \frac{tr^3}{\sqrt{t^2 - r^2}} dr d\theta \\
&= 2\pi t \int_0^{\frac{t}{\sqrt{2}}} \frac{r^3}{\sqrt{t^2 - r^2}} dr \\
&= 2\pi t \int_{t^2}^{\frac{t^2}{2}} \frac{-u + t^2}{-2\sqrt{u}} du (\because u = t^2 - r^2) \\
&= 2\pi t \int_{t^2}^{\frac{t^2}{2}} \frac{\sqrt{u}}{2} + \frac{-t^2 u^{-\frac{1}{2}}}{2} du \\
&= 2\pi t \left(\frac{u^{\frac{3}{2}}}{3} - t^2 u^{\frac{1}{2}} \right) \Big|_{u=t^2}^{\frac{t^2}{2}} \\
&= 2\pi t \left(\frac{8 - 5\sqrt{2}}{12} \right) t^3 = \frac{8 - 5\sqrt{2}}{6} t^4
\end{aligned}$$

■

Theorem 10.1.10. (Example) Given

- (a) $S = \{(x, y, z) : x^2 + y^2 = 1 \text{ and } 0 \leq z \leq 1\}$
- (b) $\mathbf{F} = (xy, x, xyz)$

We have

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = 0$$

Proof. Let

$$\mathbf{r}(\theta, z) = (\cos \theta, \sin \theta, z)$$

We have

$$S = \{\mathbf{r}(\theta, z) : \theta \in [0, 2\pi] \text{ and } z \in [0, 1]\}$$

Recall

$$\begin{aligned}
d\mathbf{S} &= \mathbf{n} dS \\
&= \mathbf{r}_\theta \times \mathbf{r}_z d\theta dz \\
&= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{vmatrix} d\theta dz \\
&= (\cos \theta, \sin \theta, 0) d\theta dz
\end{aligned}$$

Then

$$\begin{aligned}
\iint_S \mathbf{F} \cdot d\mathbf{S} &= \iint_{\Omega} \mathbf{F} \cdot (\cos \theta, \sin \theta, 0) d\theta dz \\
&= \iint_{\Omega} \cos^2 \theta \sin \theta + \cos \theta \sin \theta d\theta dz \\
&= \int_0^{2\pi} \cos^2 \sin \theta + \cos \theta \sin \theta d\theta \\
&= \frac{-\cos^3 \theta}{3} + \frac{-\cos 2\theta}{4} \Big|_{\theta=0}^{2\pi} = 0
\end{aligned}$$

■

Theorem 10.1.11. (Example) Given

(a) $S = \{(x, y, z) : x^2 + y^2 + z^2 = 4\}$

(b) $\mathbf{F} = (x, y, z)$

We have

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \frac{64}{3}\pi$$

Proof. Let

$$\mathbf{r}(\phi, \theta) = (2 \sin \phi \cos \theta, 2 \sin \phi \sin \theta, 2 \cos \phi)$$

Then

$$\begin{aligned}
\iint_S \mathbf{F} \cdot d\mathbf{S} &= \iint_{\Omega} \begin{vmatrix} 2 \sin \phi \cos \theta & 2 \sin \phi \sin \theta & 2 \cos \phi \\ 2 \cos \phi \cos \theta & 2 \cos \phi \sin \theta & -2 \sin \phi \\ -2 \sin \phi \sin \theta & 2 \sin \phi \cos \theta & 0 \end{vmatrix} d\phi d\theta \\
&= 8 \int_0^{2\pi} \int_0^{\pi} \sin^3 \phi + \sin \phi \cos \phi \cos \theta d\phi d\theta \\
&= 8 \int_0^{2\pi} \left(-\cos \phi + \frac{\cos^3 \phi}{3} - \frac{\cos 2\phi \cos \theta}{4} \right) \Big|_{\phi=0}^{\pi} d\theta \\
&= 8 \int_0^{2\pi} 2 - \frac{2}{3} d\theta \\
&= \frac{64}{3}\pi
\end{aligned}$$

Also, one can compute the following to get exactly the same value

$$\begin{aligned}
 \mathbf{F} \cdot d\mathbf{S} &= \mathbf{F} \cdot \mathbf{n} dS \\
 &= \mathbf{F} \cdot \frac{\mathbf{r}_\phi \times \mathbf{r}_\theta}{|\mathbf{r}_\phi \times \mathbf{r}_\theta|} dS \\
 &= \mathbf{F} \cdot \frac{\mathbf{r}_\phi \times \mathbf{r}_\theta}{|\mathbf{r}_\phi \times \mathbf{r}_\theta|} |\mathbf{r}_\phi \times \mathbf{r}_\theta| d\phi d\theta \\
 &= \begin{vmatrix} 2 \sin \phi \cos \theta & 2 \sin \phi \sin \theta & 2 \cos \theta \\ 2 \cos \phi \cos \theta & 2 \cos \phi \sin \theta & -2 \sin \phi \\ -2 \sin \phi \sin \theta & 2 \sin \phi \cos \theta & 0 \end{vmatrix} d\phi d\theta
 \end{aligned}$$

■

Theorem 10.1.12. (Example) Given

- (a) $\mathbf{r}(u, v) = (u + v, u - v, u^2)$
- (b) $S = \{\mathbf{r}(u, v) : (u, v) \in [0, 1] \times [0, 1]\}$
- (c) $\mathbf{F} = (1, xy, 0)$

Compute

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \frac{7}{6}$$

Proof.

$$\begin{aligned}
 \iint_S \mathbf{F} \cdot d\mathbf{S} &= \int_0^1 \int_0^1 \begin{vmatrix} 1 & u^2 - v^2 & 0 \\ 1 & 1 & 2u \\ 1 & -1 & 0 \end{vmatrix} du dv \\
 &= \int_0^1 \int_0^1 -2uv^2 + 2u^3 + 2u du dv \\
 &= \int_0^1 -v^2 + \frac{3}{2} dv = \frac{7}{6}
 \end{aligned}$$

■

10.2 Gauss Divergence Theorem

Definition 10.2.1. (Definition of Divergence) We define the divergence of a vector field \mathbf{F} by

$$\operatorname{div} \mathbf{F} = \mathbf{F} \cdot \nabla = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

Theorem 10.2.2. (Gauss Divergence Theorem) If V is a smooth area in \mathbb{R}^3 and $S = \partial V$ is always positive. Given that P, Q, R are both differentiable, we have

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \iiint_V \operatorname{div} \mathbf{F} dV$$

We now verify Gauss Divergence Theorem with following cylindrical example in which the flux can be directly computed with method in Definition 10.1.2 and Theorem 10.1.3.

Theorem 10.2.3. (Verification of Gauss Divergence Theorem with Cylindrical Example) Given

(a) $\mathbf{F} = (4x, -2y^2, z^2)$

(b) $S = \{(x, y, z) : x^2 + y^2 \leq 4 \text{ and } z \in [0, 3]\}$

We have

$$\iint_S \mathbf{F} d\mathbf{S} =$$

Proof. (Direct Method) Divide S into

$$S_1 = \{(r \cos \theta, r \sin \theta, 3) \in \mathbb{R}^3 : r \in [0, 2], \theta \in [0, 2\pi]\}$$

$$S_2 = \{(r \cos \theta, r \sin \theta, 0) \in \mathbb{R}^3 : r \in [0, 2], \theta \in [0, 2\pi]\}$$

$$S_3 = \{\mathbf{r}(\theta, z) \in \mathbb{R}^3 : [0, 2\pi], z \in [0, 3]\}$$

where

$$\mathbf{r}(\theta, z) = (2 \cos \theta, 2 \sin \theta, z)$$

For S_3 , we have

$$\mathbf{n} = \frac{\mathbf{r}_\theta \times \mathbf{r}_z}{|\mathbf{r}_\theta \times \mathbf{r}_z|}$$

where

$$\begin{aligned}\mathbf{r}_\theta \times \mathbf{r}_z &= \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -2 \sin \theta & 2 \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= (2 \cos \theta, 2 \sin \theta, 0)\end{aligned}$$

This give us

$$\begin{aligned}\iint_{S_3} \mathbf{F} d\mathbf{S} &= \iint_{S_3} \mathbf{F} \cdot \mathbf{n} dS \\ &= \iint_{\Omega} \mathbf{F} \cdot \frac{\mathbf{r}_\theta \times \mathbf{r}_z}{|\mathbf{r}_\theta \times \mathbf{r}_z|} |\mathbf{r}_\theta \times \mathbf{r}_z| d\theta dz \\ &= \int_0^{2\pi} \int_0^3 (8 \cos \theta, -8 \sin^2 \theta, z^2) \cdot (2 \cos \theta, 2 \sin \theta, 0) dz d\theta \\ &= 3 \int_0^{2\pi} 8r \cos^2\end{aligned}$$

■

Chapter 11

Advanced Calculus HW

11.1 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 $\frac{|Ax|}{|x|} = \left| \frac{Ax}{|x|} \right| = |A\hat{x}| \leq \sup\{|Ay| : |y| = 1\}$, since $|\hat{x}| = \left| \frac{x}{|x|} \right| = \frac{|x|}{|x|} = 1 \implies |A\hat{x}| \in \{|Ay| : |y| = 1\}$. 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

$$\forall M, |A\mathbf{0}| \leq M |\mathbf{0}|.$$

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

$$\frac{|(A+B)x'|}{|x'|} > \frac{|Ax'|}{|x'|} + \frac{|Bx'|}{|x'|} \quad \text{CaC}$$



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 $\sup A \cup B \geq \sup A$ and $\sup A \cup B \geq \sup B$ ■

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

$$f(x') > \sup\{f(x) : x \in D\}$$

We first construct a sequence $\{x_i\}_{i=1}^{\infty}$ in D such that

$$\lim_{i \rightarrow \infty} x_i = x' \text{ which reads } \forall \epsilon, \exists N, n > N \longrightarrow |x_n - x'| < \epsilon$$

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

$$\forall i, |x_i - x'| < \frac{1}{i}$$

For all ϵ , we can pick a natural $N > \frac{1}{\epsilon}$, so we have

$$n > N \implies |x_n - x'| < \frac{1}{n} < \frac{1}{N} < \frac{1}{\epsilon} \text{ (done)}$$

We now prove

$$\lim_{i \rightarrow \infty} f(x_i) = f(x') \text{ which reads } \forall \epsilon, \exists N, n > N \longrightarrow |f(x_n) - f(x')| < \epsilon$$

Because f is continuous, we have

$$\forall \epsilon, \exists \delta, \forall u \in \mathbb{R}, |u - x'| < \delta \implies |f(u) - f(x')| < \epsilon$$

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

$$n > N \implies |x_n - x'| < \delta \implies |f(x_n) - f(x')| < \epsilon \text{ (done)}$$

Let $H = \sup\{f(x) : x \in D\}$. Now we prove

$$\lim_{i \rightarrow \infty} f(x_i) \leq H$$

Assume $f(x') = \lim_{i \rightarrow \infty} f(x_i) > H$. We know there exists N such that

$$n > N \longrightarrow |f(x_n) - f(x')| < |H - f(x')| = f(x') - H$$

The last equality hold true due to the premise equation (5.2). Notice

$$|f(x_n) - f(x')| < f(x') - H \implies H - f(x') < f(x_n) - f(x') \implies f(x_n) > H$$

so in fact we know there exists N such that

$$n > N \longrightarrow f(x_n) > H = \sup\{f(x) : x \in D\} \text{ CaC (done)}$$

Now, using all our proven facts, we have

$$\sup\{f(x) : x \in D\} < f(x') = \lim_{i \rightarrow \infty} f(x_i) \leq H = \sup\{f(x) : x \in D\} \text{ CaC}$$

where the first inequality follows from premise equation (5.2) ■

Proof. Let $H = \sup\{f(x) : x \in D\}$. We first prove

$$\forall i \in \mathbb{N}, \{f(x) : x \in D \text{ and } H - f(x) < \frac{1}{i}\} \neq \emptyset$$

Assume there exists some $n \in \mathbb{N}$ such that the set is empty. We then have

$$\forall x \in D, H \geq f(x) + \frac{1}{n}$$

So we have

$$\forall x \in D, H - \frac{1}{2n} \geq f(x) + \frac{1}{2n} > f(x)$$

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

$$\forall \epsilon, n > \left[\frac{1}{\epsilon}\right] + 1 \implies n > \frac{1}{\epsilon} \implies |f(x_n) - H| < \frac{1}{n} < \epsilon$$

This written in limit sign is

$$\lim_{i \rightarrow \infty} f(x_i) = H = \sup\{f(x) : x \in D\} = \sup\{f(x) : x \in \mathbb{R}\}$$

■

11.2 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

$$A + B = \bigcup \{\{a + b : a \in A\} : b \in B\}$$

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

$$B_r(a) \subseteq A$$

Let

$$B_r(a) + b := \{x + b : x \in B_r(a)\}$$

We now prove

$$B_r(a) + b = B_r(a + b)$$

Arbitrarily pick $x + b \in B_r(a) + b$. We have

$$|(x + b) - (a + b)| = |x - a| < r$$

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

$$y = z + b \text{ and } |z - a| = |y - (a + b)| < r$$

The latter tell us $z \in B_r(a)$, so we have

$$y = z + b \in B_r(a) + b$$

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

$$B_r(a) \subseteq A$$

and it is clear that

$$B_r(a) + b \subseteq A + b$$

So we have

$$B_r(a + b) = B_r(a) + b \subseteq A + b$$

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 3.9.1, we have proved the set

$$\mathcal{O} := \{B_r(p) : p \in \mathbb{Q} \text{ and } r \in \mathbb{Q}^+\}$$

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

$$a \in B_{r_a}(p_a) \subseteq A$$

Collect all such open ball for each points in A and call this collection \mathcal{O}' .

Define a relation \sim on \mathcal{O}' by

$$(a, b) \sim (c, d) \text{ if } \exists S \in \mathcal{P}(\mathcal{O}') \text{ such that} \\ \bigcup S \text{ is an open interval and } (a, b) \subseteq \bigcup S \text{ and } (c, d) \subseteq \bigcup S$$

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

$$(a_1, b_1) \sim (a_2, b_2) \text{ and } (a_2, b_2) \sim (a_3, b_3) \implies (a_1, b_1) \sim (a_3, b_3)$$

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

$$S_3 := S_1 \cup S_2$$

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

$$(a_1, b_1) \subseteq \bigcup_{284} S_1 \subseteq \bigcup S_3$$

and deduce

$$(a_3, b_3) \subseteq \bigcup S_2 \subseteq \bigcup S_3 \text{ (done)}$$

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

$$B = \bigcup B' \text{ and } C = \bigcup C'$$

Assume B, C is not disjoint, say, $x \in B \cap C$, we have

$$\exists (a, b) \in B', x \in (a, b) \subseteq B$$

and have

$$\exists (c, d) \in C', x \in (c, d) \subseteq C$$

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

$$\exists (y_1, z_1) \in B', (y_1, z_1) \subseteq (\inf B, x) \text{ and } \exists (y_2, z_2) \in B', (y_2, z_2) \subseteq (x, \sup B)$$

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

$$B = A \cup (B \setminus A) \subseteq A \cup A' = \overline{A}$$

If B is closed, we have

$$\overline{A} \subseteq \overline{B} = B$$

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

$$A \subseteq A \cup B \implies \overline{A} \subseteq \overline{A \cup B}$$

and deduce

$$B \subseteq A \cup B \implies \overline{B} \subseteq \overline{A \cup B}$$

so deduce

$$\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$$

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

$$\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$$

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

$$\overline{A \cap B} \subseteq \overline{A} \text{ and } \overline{A \cap B} \subseteq \overline{B}$$

Then have

$$\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$$

Let $A = (0, 1)$ and $B = (1, 2)$. We have

$$\overline{A \cap B} = \overline{\emptyset} = \emptyset \text{ and } \overline{A} \cap \overline{B} = [0, 1] \cap [1, 2] = \{1\}$$

■

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

$$\{x_i\}_{i=1}^{\infty} \subseteq F \subseteq K$$

Because K is sequentially compact, we know there exists a sub-sequence $\{x_{n_i}\}_{i=1}^{\infty}$ converge to some point a

$$\lim_{i \rightarrow \infty} \{x_{n_i}\} = a$$

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

$$\forall p \in M, B_1(p) = \{p\}$$

So we know no set in M has a limit point. Let $K \subseteq M$ be compact. By Theorem 3.7.7, we know K is limit point compact. Assume K is infinite. Then K has a limit point $x \in K$. ■

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

$$E \subseteq \{x_n : n \in \mathbb{N}\} \cup \{x\}$$

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 3.7.7, $\{x_n : n \in \mathbb{N}\} \cup \{x\}$ is compact. ■

Question 31

- Let A and B be two subsets of a metric space (M, d) . The distance between A and B is defined by

$$d(A, B) \equiv \inf\{d(x, y) : x \in A, y \in B\}.$$

- Give an example of two disjoint, nonempty, closed sets A and B in \mathbb{R}^n for which $d(A, B) = 0$.
- 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

$$A = \{a_k = k - 1 + \sum_{i=1}^{k-1} 2^{-i-1} : k \in \mathbb{N}\}$$

Let

$$B = \{b_k = k - \sum_{i=1}^{k-1} 2^{-i-1} : u \in \mathbb{N}\}$$

Notice the pattern

$$\begin{aligned} a_1 &= 0 < \frac{1}{2} < 1 = b_1 \\ a_2 &= 1 + \frac{1}{4} < 1 + \frac{1}{2} < 2 - \frac{1}{4} = b_2 \\ a_3 &= 2 + \frac{1}{4} + \frac{1}{8} < 2 + \frac{1}{2} < 3 - \frac{1}{4} - \frac{1}{8} = b_3 \\ &\vdots \end{aligned}$$

and compute

$$b_k - a_k = 2^{1-k}$$

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

$$b_k - a_k = 2^{1-k} < r \iff 1 - k < \log_2 r \iff k > 1 - \log_2 r$$

And we are done, by picking great enough k .

Now we do (b).

Because

$$\{d(a, b) : a \in A, b \in B\} = \bigcup \{\{d(a, b) : a \in A\} : b \in B\}$$

so we have

$$\begin{aligned} d(A, B) &= \inf\{d(a, b) : a \in A, b \in B\} \\ &= \inf \bigcup \{\{d(a, b) : a \in A\} : b \in B\} \\ &= \inf\{\inf\{d(a, b) : a \in A\} : b \in B\} \\ &= \inf\{d(A, \{b\}) : b \in B\} \end{aligned}$$

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 3.7.7, 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

$$d(A, \{p\}) \leq d(A, \{b'\}) + d(b', p)$$

Because

$$d(A, \{b'\}) + d(b', p) = \inf\{d(a, b') : a \in A\} + d(b', p) = \inf\{d(a, b') + d(b', p) : a \in A\}$$

where for each a , we have

$$d(A, \{p\}) \leq d(a, p) \leq d(a, b') + d(b', p)$$

Then from

$$\forall b \in B, r \leq d(A, \{p\}) \leq d(A, \{b\}) + d(\{b\}, \{p\})$$

we have

$$\forall b \in B [d(A, \{b\}) < \frac{r}{2} \implies d(\{b\}, \{p\}) > \frac{r}{2}]$$

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

$$\inf\{d(A, \{b\}) : b \in B\} = d(A, B) = u$$

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

$$d(A, \{b_n\}) < u + \frac{1}{n}$$

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 3.7.7, 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

$$\forall b \in B, m = d(A, \{p\}) < d(A, \{b\}) + d(b, p) \text{ This was proved at } *$$

we have

$$\forall b \in B, d(A, \{b\}) < u + \frac{1}{n} \implies d(b, p) > m - u - \frac{1}{n}$$

Let k be a natural such that $u + \frac{1}{k} < m$. For each natural t greater than k , we have

$$d(b_t, p) > m - u - \frac{1}{t} > m - u - \frac{1}{k}$$

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

$$d(a_n, b) < u + \frac{1}{n}$$

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 3.7.7, 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

$$\forall a_n, m = d(a_k, b) < d(b, a_n) + d(a_n, a_k)$$

give us

$$\forall a_n, d(a_n, b) < u + \frac{1}{n} \implies d(a_n, a_k) > m - u - \frac{1}{n}$$

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

$$d(a_w, b) < u + \frac{1}{w} < u + \frac{1}{s}$$

so we have

$$d(a_w, a_k) > m - u - \frac{1}{s}$$

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 3.7.7, 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. ■

11.3 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

$$A = (A \cap S_1) \cup (A \cap S_2)$$

We wish to show

$$A \cap S_1 \text{ and } A \cap S_2 \text{ are separated.}$$

Notice

$$\overline{A \cap S_1} \subseteq \overline{S_1}$$

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

$$\overline{A \cap S_2} \subseteq \overline{S_2}$$

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. ■

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 $A_2 = \{(0, y) \in \mathbb{R}^2 : y \in \mathbb{R}\}$.

To show A_1, A_2 are both connected, we can show they are both path connected.

Let $(\cos \alpha, \sin \alpha) \in A_1$ and $(\cos \beta, \sin \beta) \in A_1$. Define $f : [0, 1] \rightarrow A_1$ by

$$f(x) = (\cos(\alpha + x(\beta - \alpha)), \sin(\alpha + x(\beta - \alpha)))$$

Clearly f is continuous, and $f(0) = (\cos \alpha, \sin \alpha)$, $f(1) = (\cos \beta, \sin \beta)$.

Let $(0, z) \in A_2$ and $(0, y) \in A_2$. Define $g : [0, 1] \rightarrow A_2$ by

$$g(x) = (0, z + x(y - z))$$

Clearly g is continuous and $g(0) = (0, z)$, $g(1) = (0, y)$.

We see $A_1 \cap A_2 = \{(0, 1), (0, -1)\}$. We see $\{(0, 1)\}$ and $\{(0, -1)\}$ are separated, because they are closed.

For (b), see next three Theorems. ■

Theorem 11.3.1. (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 . ■

Theorem 11.3.2. (Subspace Topology) Let Y be a subspace of (X, d) , let $E \subseteq X$, and let $p \in Y$. We have

$$\begin{aligned} p \text{ is an interior point of } E \text{ in } X &\implies p \text{ is an interior point of } E \cap Y \text{ in } Y \\ Y \cap E^\circ \text{ in } X &\text{ is a subset of the interior of } E \cap Y \text{ in } Y \\ E \text{ is open in } X &\implies E \cap Y \text{ is open in } Y \end{aligned}$$

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

$$n > N \implies x_n \in E$$

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. ■

Theorem 11.3.3. (Union of Connected Sets that have Nonempty Intersection is Connected) Let \mathcal{F} be a class of connected sets. We have

$$\bigcap \mathcal{F} \neq \emptyset \implies \bigcup \mathcal{F} \text{ is connected}$$

Proof. Assume $\bigcup \mathcal{F}$ is not connected. Let

$$\bigcup \mathcal{F} = A \dot{\cup} B \text{ and } A \neq \emptyset \neq B$$

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

$$A \cap \bigcap \mathcal{F} \neq \emptyset$$

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 11.3.2, 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 ■

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

$$\begin{aligned} |s_n - L| &= \left| \frac{\sum_{k=1}^n a_k}{n} - L \right| = \left| \frac{(\sum_{k=1}^n a_k) - nL}{n} \right| \\ &= \left| \frac{\sum_{k=1}^n (a_k - L)}{n} \right| \\ &= \frac{1}{n} \left| \sum_{k=1}^n (a_k - L) \right| \\ &\leq \frac{1}{n} \sum_{k=1}^n |a_k - L| \end{aligned}$$

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

$$k > N \implies |s_k - L| < R$$

Because $\lim_{n \rightarrow \infty} a_n = L$, we know there exists N_0 such that

$$n > N_0 \iff |a_n - L| < \frac{R}{2}$$

Let

$$H = \sum_{k=1}^{N_0} |a_k - L| \text{ and } m = \frac{2}{R}(H - N_0 R)$$

We wish to prove

$$n > N_0 + m \implies |s_n - L| < R$$

Let $n = N_0 + u$ where $u > m$. Observe

$$\begin{aligned} |s_n - L| &\leq \frac{1}{n} \sum_{k=1}^n |a_k - L| = \frac{\sum_{k=1}^n |a_k - L|}{N_0 + u} \\ &= \frac{\sum_{k=1}^{N_0} |a_k - L| + \sum_{k=N_0+1}^{N_0+u} |a_k - L|}{N_0 + u} \\ &\leq \frac{H + u \frac{R}{2}}{N_0 + u} \end{aligned}$$

and Observe

$$\begin{aligned} u > m = \frac{2}{R}(H - N_0 R) &\implies \frac{Ru}{2} > H - N_0 R \\ &\implies N_0 R + Ru > H + \frac{Ru}{2} \\ &\implies R > \frac{H + \frac{Ru}{2}}{N_0 + u} \end{aligned}$$

In other words, we have

$$n > N_0 + m \implies u > m \implies |s_n - L| \leq \frac{H + \frac{Ru}{2}}{N_0 + u} < R \text{ (done)}$$

For (b), we raise an counter-example. Let

$$a_n = \begin{cases} 1 & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases}$$

To see $\{a_n\}$ does not converge to any number, consider the sub-sequences

$$\{a_{n_k}\} \text{ where } n_k = 2k - 1 \text{ and } \{a_{n_u}\} \text{ where } n_u = 2u$$

We have

$$\forall k, a_{n_k} = 1 \text{ and } \forall u, a_{n_u} = 0$$

Then we see

$$\lim_{k \rightarrow \infty} a_{n_k} = 1 \text{ and } \lim_{u \rightarrow \infty} a_{n_u} = 0$$

If $\{a_n\}$ converge, then these two sub-sequence should converge to the same number.

We wish to show

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{2}$$

With simple logical computation, we have

$$s_n = \begin{cases} \frac{n-1}{2n} & \text{if } n \text{ is odd} \\ \frac{1}{2} & \text{if } n \text{ is even} \end{cases}$$

Notice

$$\frac{n-1}{2n} = \frac{1}{2} - \frac{1}{2n}$$

Fix $\epsilon \in \mathbb{R}^+$. We see that

$$n > \frac{1}{2\epsilon} \implies \left| s_n - \frac{1}{2} \right| = \begin{cases} \frac{1}{2n} & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases} < \epsilon$$

Lastly, we prove

$$\lim_{n \rightarrow \infty} a_n = a \implies \lim_{n \rightarrow \infty} t_n = a$$

First notice

$$\begin{aligned} |t_n - a| &= \left| \frac{(2n-1)a_1 + (2n-3)a_2 + \cdots + a_n}{n^2} - a \right| \\ &= \left| \frac{1}{n^2} [(2n-1)(a_1 - a) + (2n-3)(a_2 - a) + \cdots + (a_n - a)] \right| \\ &= \frac{1}{n^2} \left| \sum_{k=1}^n (2n-2k+1)(a_k - a) \right| \\ &\leq \frac{1}{n^2} \sum_{k=1}^n |(2n-2k+1)(a_k - a)| \\ &= \frac{1}{n^2} \sum_{k=1}^n (2n-2k+1) |a_k - a| \\ &\leq \frac{1}{n^2} \sum_{k=1}^n (2n-1) |a_k - a| = \frac{2n-1}{n^2} \sum_{k=1}^n |a_k - a| \end{aligned}$$

Arbitrarily pick $\epsilon \in \mathbb{R}^+$. Because $\lim_{n \rightarrow \infty} a_n = a$. We know there exists N_0 such that

$$n > N_0 \implies |a_n - a| < \frac{\epsilon}{4} \implies \left(\frac{2n-1}{n} \right) |a_n - a| = \left(2 - \frac{1}{n} \right) |a_n - a| < 2 |a_n - a| < \frac{\epsilon}{2}$$

Let

$$H = \sum_{k=1}^{N_0} |a_k - a|$$

We have

$$\begin{aligned} |t_n - a| &\leq \frac{2n-1}{n^2} \sum_{k=1}^n |a_k - a| = \frac{2n-1}{n^2} \sum_{k=1}^{N_0} |a_k - a| + \frac{2n-1}{n^2} \sum_{k=N_0+1}^n |a_k - a| \\ &= \frac{2n-1}{n^2} H + \frac{2n-1}{n^2} \sum_{k=N_0+1}^n |a_k - a| \\ &= \frac{2n-1}{n^2} H + \frac{1}{n} \sum_{k=N_0+1}^n \frac{2n-1}{n} |a_k - a| \\ &\leq \frac{2n-1}{n^2} H + \frac{1}{n} (n - N_0) \frac{\epsilon}{2} \end{aligned}$$

Then if we let

$$X_n := \frac{2n-1}{n^2} H + \frac{(n - N_0)\epsilon}{2n}$$

For all $n > N_0$, we have

$$|t_n - a| < X_n$$

Notice

$$X_n = \frac{2H}{n} - \frac{H}{n^2} + \frac{\epsilon}{2} - \frac{N_0\epsilon}{2n}$$

So we have

$$\lim_{n \rightarrow \infty} X_n = \frac{\epsilon}{2}$$

Then we know there exists some N_1 such that

$$n > N_1 \implies \left| X_n - \frac{\epsilon}{2} \right| < \frac{\epsilon}{2} \implies |t_n - a| \leq X_n < \epsilon \text{ (done)}$$

■

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. The fact

$$\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \limsup_{n \rightarrow \infty} \sqrt[n]{a_n}$$

follows from definition.

We first prove

$$\liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{a_n}$$

Let

$$\alpha = \liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

Notice that $\{a_k\}$ being positive give us $\alpha \geq 0$ and $0 \leq \liminf_{n \rightarrow \infty} \sqrt[n]{a_n}$. If $\alpha = 0$, the proof is done trivially. We only have to consider when α is positive.

Arbitrarily pick positive β smaller than α :

$$\beta < \alpha = \liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

Then we know there exists N such that

$$\forall n \geq N, \frac{a_{n+1}}{a_n} > \beta$$

This implies

$$\forall k, a_{N+k} > \beta^k a_N$$

Then for all $n > N$, we have

$$\sqrt[n]{a_n} > \sqrt[n]{\beta^{n-N} a_N} = \beta \sqrt[n]{\beta^{-N} a_N}$$

Because

$$\lim_{n \rightarrow \infty} \beta \sqrt[n]{\beta^{-N} a_N} = \beta$$

We see

$$\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \geq \beta$$

Notice that β is arbitrarily pick from $\{x \in \mathbb{R} : 0 \leq x < \alpha\}$, so we have in fact proved

$$0 \leq x < \alpha \implies x \leq \liminf_{n \rightarrow \infty} \sqrt[n]{a_n}$$

If $\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} < \alpha$, there should exists $x < \alpha$ such that $\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} < x$, which we have prove is impossible. (done)

We now prove

$$\limsup_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

Let $\gamma = \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$. If $\gamma = \infty$, the proof is done trivially. We only have to consider when $\gamma < \infty$.

Notice that $\{a_k\}$ being positive give us $\gamma \geq 0$. Arbitrarily pick positive δ greater than γ :

$$\delta > \gamma = \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

Then we know there exists N such that

$$\forall n \geq N, \frac{a_{n+1}}{a_n} < \delta$$

This implies

$$\forall k, a_{N+k} < \delta^k a_N$$

Then for all $n > N$, we have

$$\sqrt[n]{a_n} < \sqrt[n]{\delta^{n-N} a_N} = \delta \sqrt[n]{\delta^{-N} a_N}$$

Because

$$\lim_{n \rightarrow \infty} \delta \sqrt[n]{\delta^{-N} a_N} = \delta$$

We see

$$\limsup_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \delta$$

Notice that δ is arbitrarily picked from $\{x \in \mathbb{R} : x > \gamma\}$, so we have in fact proved

$$x > \gamma \implies x \geq \limsup_{n \rightarrow \infty} \sqrt[n]{a_n}$$

If $\limsup_{n \rightarrow \infty} \sqrt[n]{a_n} > \gamma$, there should exists some $x > \gamma$ such that $\limsup_{n \rightarrow \infty} \sqrt[n]{a_n} > x$, which we have proved is impossible. (done)

Let

$$a_k = \begin{cases} 2 & \text{if } k \text{ is odd} \\ 1 & \text{if } k \text{ is even} \end{cases}$$

We see

$$\limsup_{n \rightarrow \infty} \sqrt[n]{a_n} = 1 < 2 \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

■

Question 39

6. If $s_1 = \sqrt{2}$, and

$$s_{n+1} = \sqrt{2 + \sqrt{s_n}}, \quad n = 1, 2, 3,$$

prove that s_n converge and bounded above by 2

Proof. We first prove $\{s_n\}$ increase monotonically by induction.

Base case:

$$s_1^2 = 2 < 2 + \sqrt{2} = s_2^2 \implies s_1 < s_2$$

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

$$s_k \leq s_{k+1} = \sqrt{2 + \sqrt{s_k}}$$

Then we can deduce

$$\begin{aligned} s_k &\leq \sqrt{2 + \sqrt{s_k}} \\ \implies \sqrt{s_k} &\leq \sqrt{\sqrt{2 + \sqrt{s_k}}} \\ \implies 2 + \sqrt{s_k} &\leq 2 + \sqrt{\sqrt{2 + \sqrt{s_k}}} \\ \implies \sqrt{2 + \sqrt{s_k}} &\leq \sqrt{2 + \sqrt{\sqrt{2 + \sqrt{s_k}}}} \\ \implies s_{k+1} = \sqrt{2 + \sqrt{s_k}} &\leq \sqrt{2 + \sqrt{s_{k+1}}} = s_{k+2} \text{ (done)} \end{aligned}$$

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

$$\begin{aligned}
 s_k \leq 2 &\implies s_k \leq 4 \\
 &\implies \sqrt{s_k} \leq \sqrt{4} = 2 \\
 &\implies 2 + \sqrt{s_k} \leq 4 \\
 &\implies s_{k+1} = \sqrt{2 + \sqrt{s_k}} \leq \sqrt{4} = 2 \text{ (done)}
 \end{aligned}$$

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

$$S_n = \sum_{k=1}^n \frac{a_k}{1 + ka_k}$$

$$T_n = \sum_{k=1}^n \frac{a_k}{1 + k^2 a_k}$$

If $s_n = \sum_{k=1}^n a_k$ converge, does $J_n = \sum_{k=1}^n ka_k$ converge

Proof. We prove

$$t_n \text{ converge} \implies s_n \text{ converge}$$

Notice that

$$a_n = \frac{a_n + a_n^2}{1 + a_n} = \frac{a_n}{1 + a_n} + \frac{a_n^2}{1 + a_n}$$

So we have

$$s_n = t_n + \sum_{k=1}^n \frac{a_k^2}{1 + a_k}$$

Because t_n converge, above tell us we only have to prove $\sum_{k=1}^n \frac{a_k^2}{1+a_k}$ converge.

Because t_n converge, we know

$$\lim_{n \rightarrow \infty} \frac{a_n}{1 + a_n} = 0$$

Assume $\lim_{n \rightarrow \infty} a_n \neq 0$. Then there exists ϵ such that

$$\forall N \in \mathbb{N}, \exists n > N, a_n > \epsilon$$

Notice

$$a_n > \epsilon \implies \frac{a_n}{1+a_n} = 1 - \frac{1}{1+a_n} > 1 - \frac{1}{1+\epsilon} \implies \left| \frac{a_n}{1+a_n} - 0 \right| > 1 - \frac{1}{1+\epsilon}$$

In other words, there exists a sub-sequence $\{a_{f(n)}\}$ such that

$$\lim_{n \rightarrow \infty} \frac{a_{f(n)}}{1+a_{f(n)}} \neq 0 \text{ CaC}$$

We have proved $\lim_{n \rightarrow \infty} a_n = 0$. Then we know there exists N_1 such that

$$\forall n > N_1, a_n < 1$$

In other words,

$$\forall n > N_1, \frac{a_n^2}{1+a_n} < \frac{a_n}{1+a_n}$$

By comparison test, our proof is done (done) .

We show

It is possible s_n diverge and S_n converge.

Let

$$a_k = \begin{cases} 1 & \text{if } \exists u \in \mathbb{N}, k = u^2 \\ \frac{1}{k^2} & \text{otherwise} \end{cases}$$

Clearly, $\lim_{n \rightarrow \infty} s_n = \infty$. Yet, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} S_n &= \sum_{k=1}^{\infty} \frac{a_k}{1+ka_k} = \sum_{u=1}^{\infty} \frac{a_{u^2}}{1+u^2a_{u^2}} + \sum_{k \in \mathbb{N} \setminus \{r^2: r \in \mathbb{N}\}} \frac{a_k}{1+ka_k} \\ &= \sum_{u=1}^{\infty} \frac{1}{1+u^2} + \sum_{k \in \mathbb{N} \setminus \{r^2: r \in \mathbb{N}\}} \frac{\frac{1}{k^2}}{1+\frac{1}{k}} \\ &= \sum_{u=1}^{\infty} \frac{1}{1+u^2} + \sum_{k \in \mathbb{N} \setminus \{r^2: r \in \mathbb{N}\}} \frac{1}{k(k+1)} \end{aligned}$$

Notice that

$$\frac{1}{1+u^2} < \frac{1}{u^2} \text{ and } \frac{1}{k(k+1)} < \frac{1}{k^2}$$

Then by comparison test, we know

$$\text{both } \sum_{u=1}^{\infty} \frac{1}{1+u^2} \text{ and } \sum_{k \in \mathbb{N} \setminus \{r^2: r \in \mathbb{N}\}} \frac{1}{k(k+1)} \text{ converge}$$

Then we know

S_n also converge (done)

Notice that for all n

$$k \geq 1 \implies 1 + k^2 a_k > 1 + k a_k \implies \frac{a_k}{1 + k^2 a_k} < \frac{a_k}{1 + k a_k} \implies T_n < S_n$$

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 $\{a_{n_k}\}_{k \in \mathbb{N}}$

$$\forall k \in \mathbb{N}, a_{n_k} \notin B_\epsilon(a)$$

Because A is (sequentially) compact, we know there must exist a sub-sequence $\{a_{n_{k_u}}\}_{u \in \mathbb{N}}$ such that

$$\{a_{n_{k_u}}\} \text{ converge}$$

Notice that

$$\{a_{n_{k_u}}\} \text{ is a sub-sequence of } \{a_n\}$$

So we know

$$\lim_{u \rightarrow \infty} a_{n_{k_u}} = a \text{ CaC to } \forall n \in \mathbb{N}, a_{n_k} \neq B_\epsilon(a)$$

Let $A = \mathbb{Q}$, and let

$$a_n = \begin{cases} 0 & \text{if } n \text{ is odd} \\ n & \text{if } n \text{ is even} \end{cases}$$

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

$$\forall n > N, \frac{\ln\left(\frac{1}{|a_n|}\right)}{\ln n} > \alpha > 1$$

This give us

$$\forall n > N, \ln\left(\frac{1}{|a_n|}\right) > \alpha \ln n$$

This give us

$$\forall n > N, \frac{\ln\left(\frac{1}{|a_n|}\right)}{\alpha} > \ln n$$

This give us

$$\forall n > N, \ln(|a_n|^{-\frac{1}{\alpha}}) > \ln n$$

This give us

$$\forall n > N, |a_n|^{\frac{-1}{\alpha}} > n$$

This give us

$$\forall n > N, |a_n| < n^{-\alpha} < n^{-1}$$

By comparison test, we are done.

Let $p < \beta < 1$. Then we know there exists N such that

$$\forall n > N, \frac{\ln\left(\frac{1}{|a_n|}\right)}{\ln n} < \beta < 1$$

This give us

$$\forall n > N, \frac{\ln(|a_n|^{-1})}{\beta} < \ln n$$

This give us

$$\forall n > N, \ln(|a_n|^{\frac{-1}{\beta}}) < \ln n$$

This give us

$$\forall n > N, |a_n|^{\frac{-1}{\beta}} < n$$

This give us

$$\forall n > N, |a_n| > n^{-\beta} > n^{-1}$$

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

$$\alpha < \beta < -1$$

Because

$$\lim_{x \rightarrow \infty} \frac{xf'(x)}{f(x)} = \alpha$$

We know there exists R such that

$$\forall x > R, \frac{xf'(x)}{f(x)} \leq \beta$$

Then we have

$$\forall x > R, \frac{d}{dx} \ln f(x) = \frac{f'(x)}{f(x)} \leq \frac{\beta}{x}$$

This tell us

$$\forall y > R, \int_R^y \frac{d}{dx} \ln f(x) dx \leq \int_R^y \frac{\beta}{x} dx$$

Which means

$$\forall y > R, \ln\left(\frac{f(y)}{f(R)}\right) = \ln f(y) - \ln f(R) \leq \beta(\ln(y) - \ln(R)) = \ln\left(\frac{y}{R}\right)^\beta$$

This shows

$$\forall y > R, \frac{f(y)}{f(R)} \leq \left(\frac{y}{R}\right)^\beta$$

It means

$$\forall y > R, f(y) \leq \frac{f(R)}{R^\beta} y^\beta$$

Notice $\beta < -1$, we know the series

$$\sum_{n>R} \frac{f(R)}{R^\beta} n^\beta = \frac{f(R)}{R^\beta} \sum_{n>R} n^\beta \text{ converge}$$

Then because $f(y)$ is always positive (given by the question), we know $\sum_{n=1}^{\infty} f(n)$ converge by comparison test. ■

Lemma 11.3.4. (Bernoulli Inequality) Let $r \geq 1$ and $x \geq -1$. We have

$$(1+x)^r \geq 1+rx$$

Proof. Let $r \geq 1$, and let

$$f(x) = (1+x)^r - 1 - rx$$

We wish to prove

$$\forall x \geq -1, f(x) \geq 0$$

We now split the proof into two parts

$$\begin{aligned} \forall x \geq 0, f(x) &\geq 0 \\ \forall x \in [-1, 0), f(x) &\geq 0 \end{aligned}$$

First notice by computation

$$f'(x) = r(1+x)^{r-1} - r = r((1+x)^{r-1} - 1)$$

And notice by some algebra

$$f'(x) \geq 0 \iff (1+x)^{r-1} \geq 1 \iff x \geq 0$$

We first prove

$$\forall x \geq 0, f(x) \geq 0$$

By computation,

$$f(0) = 0$$

This give us

$$\forall x \geq 0, f(x) = f(x) - 0 = f(x) - f(0) = \int_0^x f'(t)dt \geq 0 \text{ (done)}$$

We now prove

$$\forall x \in [-1, 0), f(x) \geq 0$$

Notice that above have shown

$$f'(x) \geq 0 \iff x \geq 0$$

So we know

$$\forall x \in [-1, 0), f'(x) < 0$$

Observe that $f(x)$ is a polynomial, so we know $f(x)$ is continuous. Then if for some $y \in [-1, 0)$, we have $f(y) < 0 = f(0)$, there must exist $u \in (y, 0) \subseteq [-1, 0)$ such that $f'(u) > 0$, which is impossible. ■

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. Pick α that satisfy

$$\lim_{k \rightarrow \infty} k \left(1 - \left| \frac{a_{k+1}}{a_k} \right| \right) > \alpha > 1$$

We know there exist N such that

$$\forall n > N, n \left(1 - \left| \frac{a_{n+1}}{a_n} \right| \right) > \alpha$$

With a little algebra,

$$\forall n > N, 1 - \frac{\alpha}{n} > \left| \frac{a_{n+1}}{a_n} \right|$$

Plugin Lemma 11.3.4 (Bernoulli Inequality) with $r = \alpha > 1$ and $x = \frac{-1}{n} \geq -1$. We have

$$\forall n > N, \left| \frac{a_{n+1}}{a_n} \right| < 1 + \frac{-\alpha}{n} \leq \left(1 - \frac{1}{n} \right)^\alpha = \left(\frac{n-1}{n} \right)^\alpha$$

Then we have

$$(*) \forall n > N, |a_{n+1}| n^\alpha < |a_n| (n-1)^\alpha$$

For each $n > N$, define

$$b_n = |a_{n+1}| n^\alpha$$

Then we have

$$|a_n| (n-1)^\alpha = b_{n-1}$$

So by $(*)$ we have

$$\forall n > N+1, b_n < b_{n-1}$$

This tell us $\{b_n\}_{n>N+1}$ is a decreasing sequence. Then we know $\{b_n : n > N+1\}$ is bounded above. More precisely, let

$$M = b_{N+1}$$

We have

$$\forall n > N+1, b_n < M$$

Recall the definition of b_n , we have

$$\forall n > N+1, |a_{n+1}| n^\alpha < M$$

In other words

$$\forall n > N+2, |a_n| < M n^{-\alpha}$$

Notice that because $\alpha > 1$. We know

$$\sum_{n \in \mathbb{N}} M n^{-\alpha} \text{ converge}$$

Then by comparison test, we know

$$\sum_{n \in \mathbb{N}} |a_n| \text{ converge}$$

■

11.4 HW4

Theorem 11.4.1. Let $B_n = \sum_{k=1}^n b_k$. Given

$$a_n \searrow 0 \text{ or } a_n \nearrow 0 \text{ as } n \rightarrow \infty$$

and

$$\exists M, \forall n, |B_n| < M$$

We have

$$\sum_{n=1}^{\infty} a_n b_n \text{ converge}$$

Proof. Let N satisfy

$$a_{n+N} \nearrow 0 \text{ or } a_{n+N} \searrow 0$$

We see

$$\sum_{k=N+1}^n b_k = B_n - B_N \text{ is also bounded}$$

If $a_{n+N} \searrow 0$, then by Dirichlet's test we see

$$\sum_{k=1}^{\infty} a_k b_k = \sum_{k=1}^N a_k b_k + \sum_{n=N+1}^{\infty} a_k b_k \text{ converge}$$

If $a_{n+N} \nearrow 0$, we see $-a_{n+N} \searrow 0$. Then by Dirichlet's test we see

$$\sum_{k=1}^{\infty} a_k b_k = \sum_{k=1}^N a_k b_k - \sum_{k=N+1}^{\infty} (-a_k) b_k \text{ converges}$$

■

Question 45

Suppose that $\sum_{k=1}^{\infty} a_k$ converges and that $b_k \searrow b$ as $k \rightarrow \infty$. Prove that

$$\sum_{k=1}^{\infty} a_k b_k \text{ converges.}$$

Proof. Define

$$b'_n := b_n - b$$

Deduce

$$\begin{aligned}\sum_{k=1}^{\infty} a_k b_k &= \sum_{k=1}^{\infty} a_k (b + b'_k) \\ &= b \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} a_k b'_k\end{aligned}$$

Notice

$$b_k \searrow b \text{ as } k \rightarrow \infty \implies b'_k \searrow 0 \text{ as } k \rightarrow \infty$$

Then by Dirichlet's test, we know

$$\sum_{k=1}^{\infty} a_k b'_k \text{ converge}$$

Because by premise $\sum_{k=1}^{\infty} a_k$ converge, we know $\sum_{k=1}^{\infty} a_k b_k = b \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} a_k b'_k$ converge. ■

Question 46

Show that under the hypotheses of Dirichlet's test,

$$\sum_{k=1}^{\infty} a_k b_k = \sum_{k=1}^{\infty} s_k (b_k - b_{k+1}),$$

where

$$s_k = \sum_{j=1}^k a_j.$$

Proof. Abel's formula give us

$$\begin{aligned}\sum_{k=1}^n a_k b_k &= \sum_{k=1}^n (s_k - s_{k-1}) b_k \\ &= \sum_{k=1}^n s_k b_k - \sum_{k=1}^{n-1} s_k b_{k+1} \\ &= \sum_{k=1}^{n-1} s_k (b_k - b_{k+1}) + s_n b_n\end{aligned}$$

Define

$$X_n := \sum_{k=1}^n a_k b_k \text{ and } Y_n := \sum_{k=1}^n s_k (b_k - b_{k+1})$$

Abel's formula then can be rewritten into

$$\forall n, Y_n = X_{n+1} - s_n b_n$$

Define

$$\alpha := \lim_{n \rightarrow \infty} X_n$$

The question ask us to prove

$$\lim_{n \rightarrow \infty} Y_n = \alpha$$

The hypothesis of Dirichlet's test are

$$s_n = \sum_{k=1}^n a_k \text{ are bounded and } b_n \searrow 0$$

Let $M \in \mathbb{R}^+$ satisfy

$$\forall n, |s_n| < M$$

Arbitrarily pick $\epsilon \in \mathbb{R}^+$. We know there exists N_1 such that

$$\forall n > N_1, 0 \leq b_n < \frac{\epsilon}{2M}$$

Because $\alpha = \lim_{n \rightarrow \infty} X_n$, we also know there exists N_2 such that

$$\forall n > N_2, |X_n - \alpha| < \frac{\epsilon}{2}$$

We then see

$$\begin{aligned} \forall n > \max\{N_1, N_2\}, |Y_n - \alpha| &= |X_{n+1} - \alpha - s_n b_n| \\ &\leq |X_{n+1} - \alpha| + |s_n b_n| \\ &< \frac{\epsilon}{2} + M \times \frac{\epsilon}{2M} = \epsilon \end{aligned}$$

Because ϵ is arbitrarily picked from \mathbb{R}^+ . We have proved $\lim_{n \rightarrow \infty} Y_n = \alpha$ ■

Lemma 11.4.2. Given n

$$\sum_{k=1}^{\infty} x_k \text{ converge} \implies \sum_{k=n}^{\infty} x_k = \sum_{k=1}^{\infty} x_k - \sum_{k=1}^{n-1} x_k$$

Proof. For all $u \in \mathbb{N}$, define

$$s_u := \sum_{k=n}^{n+u} x_k \text{ and } v_u := \sum_{k=1}^u x_k$$

From definition, we have

$$\forall u \in \mathbb{N}, s_u = v_{n+u} - \sum_{k=1}^{n-1} x_k \text{ and } \lim_{u \rightarrow \infty} v_u \text{ converge}$$

Notice

$$\lim_{u \rightarrow \infty} v_{n+u} = \lim_{u \rightarrow \infty} v_u$$

Then we have

$$\lim_{u \rightarrow \infty} s_u = \lim_{u \rightarrow \infty} (v_{n+u} - \sum_{k=1}^{n-1} x_k) = (\lim_{u \rightarrow \infty} v_{n+u}) - \sum_{k=1}^{n-1} x_k = \lim_{u \rightarrow \infty} v_u - \sum_{k=1}^{n-1} x_k$$

Which written in series is

$$\sum_{k=n}^{\infty} x_k = \sum_{k=1}^{\infty} x_k - \sum_{k=1}^{n-1} x_k$$

■

Lemma 11.4.3.

$$\sum_{k=1}^{\infty} x_k \text{ converge} \implies \lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} x_k = 0$$

Proof. By Lemma 11.4.2, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} x_k &= \lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{n-1} a_k \\ &= \sum_{k=1}^{\infty} a_k - \lim_{n \rightarrow \infty} \sum_{k=1}^{n-1} a_k \\ &= \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} a_k = 0 \end{aligned}$$

■

Question 47

Suppose that $\sum_{k=1}^{\infty} a_k$ converges, that $b_k \nearrow \infty$ and that $\sum_{k=1}^{\infty} a_k b_k$ converge. Prove

$$\lim_{m \rightarrow \infty} b_m \sum_{k=m}^{\infty} a_k = 0$$

Proof. Define

$$c_k = a_k b_k \text{ and } C_k = \sum_{j=k}^{\infty} c_j$$

By Lemma 11.4.3, we have

$$\lim_{k \rightarrow \infty} C_k = 0$$

Observe

$$\begin{aligned} \sum_{k=n}^{\infty} a_k &= \sum_{k=n}^{\infty} \frac{c_k}{b_k} \\ &= \sum_{k=n}^{\infty} \frac{C_k - C_{k+1}}{b_k} \\ &= \lim_{N \rightarrow \infty} \sum_{k=n}^N \frac{C_k - C_{k+1}}{b_k} \\ &= \lim_{N \rightarrow \infty} \sum_{k=n}^N \frac{C_k}{b_k} - \sum_{k=n+1}^{N+1} \frac{C_k}{b_{k-1}} \\ &= \lim_{N \rightarrow \infty} \sum_{k=n+1}^N C_k \left(\frac{1}{b_k} - \frac{1}{b_{k-1}} \right) - \frac{C_{N+1}}{b_N} + \frac{C_n}{b_n} \end{aligned}$$

Because $C_{N+1} \rightarrow 0$ and $b_N \rightarrow \infty$ as $N \rightarrow \infty$, we have

$$\begin{aligned} \sum_{k=n}^{\infty} a_k &= \lim_{N \rightarrow \infty} \sum_{k=n+1}^N C_k \left(\frac{1}{b_k} - \frac{1}{b_{k-1}} \right) + \frac{C_n}{b_n} \\ &= \sum_{k=n+1}^{\infty} C_k \left(\frac{1}{b_k} - \frac{1}{b_{k-1}} \right) + \frac{C_n}{b_n} \end{aligned}$$

This give us

$$\begin{aligned} \left| b_n \sum_{k=n}^{\infty} a_n \right| &= \left| b_n \sum_{k=n+1}^{\infty} C_k \left(\frac{1}{b_k} - \frac{1}{b_{k-1}} \right) + b_n \times \frac{C_n}{b_n} \right| \\ &\leq \left| b_n \sum_{k=n+1}^{\infty} C_k \left(\frac{1}{b_k} - \frac{1}{b_{k-1}} \right) \right| + |C_n| \end{aligned}$$

Arbitrarily pick $\epsilon \in \mathbb{R}^+$. Because $C_n \rightarrow 0$ and $b_n \nearrow \infty$ as $n \rightarrow \infty$, we know there exists N such that $|C_n| < \frac{\epsilon}{2}$ and $b_n > 0$ when $n > N$.

Let $n > N$. We now have

$$\begin{aligned}
\left| b_n \sum_{k=n}^{\infty} a_n \right| &\leq b_n \left| \sum_{k=n+1}^{\infty} C_k \left(\frac{1}{b_k} - \frac{1}{b_{k-1}} \right) \right| + \frac{\epsilon}{2} \\
&\leq b_n \sum_{k=n+1}^{\infty} \left| C_k \left(\frac{1}{b_k} - \frac{1}{b_{k-1}} \right) \right| + \frac{\epsilon}{2} \\
&\leq b_n \sum_{k=n+1}^{\infty} |C_k| \times \left| \frac{1}{b_k} - \frac{1}{b_{k-1}} \right| + \frac{\epsilon}{2} \\
&= b_n \sum_{k=n+1}^{\infty} |C_k| \times \left(\frac{1}{b_{k-1}} - \frac{1}{b_k} \right) + \frac{\epsilon}{2}
\end{aligned}$$

Notice that because $C_n \rightarrow 0$, we know $\{C_n : n \in \mathbb{N}\}$ is bounded. Then we know for each n , the supremum $\sup_{k \geq n} |C_k|$ exists. Define

$$X_n = \sup_{k \geq n+1} |C_k|$$

For $n > N$, we now have

$$\begin{aligned}
\left| b_n \sum_{k=n}^{\infty} a_n \right| &\leq b_n \sum_{k=n+1}^{\infty} |C_k| \times \left(\frac{1}{b_{k-1}} - \frac{1}{b_k} \right) + \frac{\epsilon}{2} \\
&\leq b_n \sum_{k=n+1}^{\infty} X_n \times \left(\frac{1}{b_{k-1}} - \frac{1}{b_k} \right) + \frac{\epsilon}{2} \\
&= (X_n b_n \sum_{k=n+1}^{\infty} \frac{1}{b_{k-1}} - \frac{1}{b_k}) + \frac{\epsilon}{2} \\
&= \lim_{u \rightarrow \infty} X_n b_n \left(\frac{1}{b_n} - \frac{1}{b_u} \right) + \frac{\epsilon}{2} \\
&= X_n \frac{b_n}{b_n} + \frac{\epsilon}{2} \\
&= X_n + \frac{\epsilon}{2}
\end{aligned}$$

Notice that

$$\lim_{n \rightarrow \infty} X_n = \lim_{n \rightarrow \infty} X_{n-1} = \lim_{n \rightarrow \infty} \sup_{k \geq n} |C_k| = \lim_{n \rightarrow \infty} \sup |C_n|$$

Then because $C_n \rightarrow 0$, we know

$$\lim_{n \rightarrow \infty} X_n = \lim_{n \rightarrow \infty} \sup |C_n| = \lim_{n \rightarrow \infty} |C_n| = 0$$

This tell us there exists N_1 such that $X_n < \frac{\epsilon}{2}$. Then for $n > \max\{N, N_1\}$, we have

$$\left| b_n \sum_{k=n}^{\infty} a_k \right| \leq X_n + \frac{\epsilon}{2} = \epsilon$$

This finish the proof. ■

Question 48

Suppose that $a_k > 0$ and

$$\sum_{k=1}^{\infty} a_k$$

converges. Prove that there exist b_k such that

$$\lim_{k \rightarrow \infty} \frac{b_k}{a_k} = \infty$$

and

$$\sum_{k=1}^{\infty} b_k$$

converges.

Proof. Define

$$r_n := \sum_{k=n}^{\infty} a_k$$

Because $\{a_k\}$ are positive, by Lemma 11.4.2, we know r_n monotonically decrease.

Moreover, by Lemma 11.4.3, we know $r_n \searrow 0$.

Observe

$$\sqrt{r_n} - \sqrt{r_{n+1}} = \frac{r_n - r_{n+1}}{\sqrt{r_n} + \sqrt{r_{n+1}}} = \frac{a_n}{\sqrt{r_n} + \sqrt{r_{n+1}}}$$

Define $\{b_n\}$ by

$$b_n := \frac{a_n}{\sqrt{r_n} + \sqrt{r_{n+1}}} = \sqrt{r_n} - \sqrt{r_{n+1}}$$

Observe

$$\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} (\sqrt{r_k} - \sqrt{r_{k+1}}) = \sqrt{r_1}$$

Notice

$$\frac{b_n}{a_n} = \frac{1}{\sqrt{r_n} + \sqrt{r_{n+1}}}$$

And deduce

$$r_n \searrow 0 \implies \sqrt{r_n} \searrow 0 \implies \frac{b_n}{a_n} = \frac{1}{\sqrt{r_n} + \sqrt{r_{n+1}}} \nearrow \infty$$

■

Question 49

Suppose that $a_k > 0$ and

$$\sum_{k=1}^{\infty} a_k$$

diverges. Prove that there exist b_k such that

$$\lim_{k \rightarrow \infty} \frac{b_k}{a_k} = 0$$

and

$$\sum_{k=1}^{\infty} b_k$$

diverges.

Proof. Define

$$A_n := \sum_{k=1}^n a_k$$

And define

$$b_n := \frac{a_n}{A_n}$$

Notice that $\{a_n\}$ are positive and $\sum_{k=1}^{\infty} a_k$ diverges implies

$$\lim_{n \rightarrow \infty} A_n = \sum_{k=1}^{\infty} a_k = \infty$$

So we have

$$\lim_{n \rightarrow \infty} \frac{b_n}{a_n} = \lim_{n \rightarrow \infty} \frac{1}{A_n} = 0$$

We now show the following, which will be essential for proving $\sum_{k=1}^{\infty} b_k$ diverges.

Define

$$X_p^q := \sum_{n=p}^q b_n$$

Let $q > p$. Because $\{a_n\}$ are positive, we have

$$A_p < A_q$$

This give us

$$X_p^q = \sum_{n=p}^q \frac{a_n}{A_n} \geq \sum_{n=p}^q \frac{a_n}{A_q} = \frac{A_q - A_{p-1}}{A_q} = 1 - \frac{A_{p-1}}{A_q}$$

Then because $\lim_{n \rightarrow \infty} A_n = \infty$, we have

$$\forall p, \liminf_{q \rightarrow \infty} X_p^q \geq \liminf_{q \rightarrow \infty} 1 - \frac{A_{p-1}}{A_q} = \lim_{q \rightarrow \infty} 1 - \frac{A_{p-1}}{A_q} = 1$$

Define $p_1 = 1$. We now know there exists $p_2 > p_1$ such that

$$\sum_{n=p_1}^{p_2-1} b_n = X_{p_1}^{p_2-1} > \frac{1}{2}$$

Again, because $\forall p, \liminf_{q \rightarrow \infty} X_p^q = 1$, we know there exists $p_3 > p_2$ such that

$$\sum_{n=p_2}^{p_3-1} b_n = X_{p_2}^{p_3-1} > \frac{1}{2}$$

Proceeding the argument above, we know there exists an increasing index sequence $\{p_k\}_{k \in \mathbb{N}}$ such that

$$\forall k, \sum_{n=p_k}^{p_{k+1}-1} b_k > \frac{1}{2}$$

Then we have

$$\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} \sum_{n=p_k}^{p_{k+1}-1} b_k \geq \sum_{k=1}^{\infty} \frac{1}{2} = \infty$$

■

Question 50

Prove that

$$\sum_{k=1}^{\infty} a_k \cos(kx)$$

converges for every $x \in (0, 2\pi)$ and every $a_k \searrow 0$. What happens when $x = 0$?

Proof. Let $x \in (0, 2\pi)$, and let

$$B_n = \sum_{k=1}^n \cos(kx)$$

We wish to prove

$\{B_n\}$ is bounded

Deduce

$$\begin{aligned} \forall n, |B_n| &= \left| \sum_{k=1}^n \cos(kx) \right| = \left| \sum_{k=1}^n \operatorname{Re} e^{i(kx)} \right| \\ &= \left| \sum_{k=1}^n \operatorname{Re} (e^{ix})^k \right| \\ &= \left| \operatorname{Re} \sum_{k=1}^n (e^{ix})^k \right| \\ &= \left| \operatorname{Re} \frac{e^{ix}(1 - e^{nix})}{1 - e^{ix}} \right| \\ &\leq \left| \frac{e^{ix}(1 - e^{nix})}{1 - e^{ix}} \right| \\ &= \frac{|e^{ix} - e^{(n+1)ix}|}{|1 - e^{ix}|} \\ &\leq \frac{|e^{ix}| + |e^{(n+1)ix}|}{|1 - e^{ix}|} \\ &\leq \frac{2}{|1 - e^{ix}|} \end{aligned}$$

Notice that because $x \in (0, 2\pi)$, we have

$$e^{ix} \neq 1$$

So we know $\frac{2}{|1-e^{ix}|} \in \mathbb{R}^+$. Let $M = \frac{2}{|1-e^{ix}|}$. By the long inequality above, we have

$$\forall n, |B_n| \leq M \text{ (done)}$$

Because $\{B_n\}$ is bounded, by Dirichlet's Test, we are done.

If $x = 0$, it is possible that $a_k = \frac{1}{k}$, and we will have

$$\sum_{k=1}^{\infty} a_k \cos(kx) = \sum_{k=1}^{\infty} a_k \text{ diverge when } a_k \searrow 0$$

■

Question 51

Prove that

$$\sum_{k=1}^{\infty} a_k \sin((2k+1)x)$$

converges for every $x \in \mathbb{R}$ and every $a_k \searrow 0$.

Proof. Notice that

$$x \equiv 0 \pmod{\pi} \implies \forall k, \sin((2k+1)x) = 0$$

This make our proof trivial. We only have to consider when

$$x \not\equiv 0 \pmod{\pi}$$

Observe

$$\begin{aligned} \forall n, \left| \sum_{k=1}^n \sin((2k+1)x) \right| &= \left| \sum_{k=1}^n \operatorname{Im} e^{i(2k+1)x} \right| \\ &= \left| \operatorname{Im} \sum_{k=1}^n e^{i(2k+1)x} \right| \\ &\leq \left| \sum_{k=1}^n e^{i(2k+1)x} \right| \\ &= \left| \frac{e^{i3x}(1 - e^{i2nx})}{1 - e^{i2x}} \right| \\ &\leq \frac{|e^{i3x}| + |e^{i(3x+2nx)}|}{|1 - e^{i2x}|} \\ &\leq \frac{2}{|1 - e^{i2x}|} \end{aligned}$$

Because $x \not\equiv 0 \pmod{\pi}$, we know $\frac{2}{|1-e^{i2x}|} \in \mathbb{R}^+$. We have proved the partial sums $\sum_{k=1}^n \sin((2k+1)x)$ is bounded. The proof for this problems follows from Dirichlet's Test. ■

Question 52

Suppose that

$$\sum_{k=1}^{\infty} a_k^2$$

and

$$\sum_{k=1}^{\infty} b_k^2$$

converges. Prove that the following series

1.

$$\sum_{k=1}^{\infty} |a_k b_k|$$

2.

$$\sum_{k=1}^{\infty} (a_k + b_k)^2$$

3.

$$\sum_{k=1}^{\infty} \frac{|a_k|}{k}$$

converge.

Proof. Deduce

$$\begin{aligned} 0 \leq (a_k - b_k)^2 &\implies 0 \leq a_k^2 - 2a_k b_k + b_k^2 \\ &\implies a_k b_k \leq \frac{a_k^2 + b_k^2}{2} \end{aligned}$$

Deduce

$$\begin{aligned} 0 \leq (a_k + b_k)^2 &\implies 0 \leq a_k^2 + 2a_k b_k + b_k^2 \\ &\implies a_k b_k \geq -\frac{a_k^2 + b_k^2}{2} \end{aligned}$$

Above give us

$$|a_k b_k| \leq \frac{a_k^2 + b_k^2}{2}$$

Clearly

$$\sum_{k=1}^{\infty} \frac{a_k^2 + b_k^2}{2} \text{ converges}$$

Then by comparison test, we know $\sum_{k=1}^{\infty} |a_k b_k|$ converges.

Notice that

$$0 \leq (a_k + b_k)^2 = a_k^2 + 2a_k b_k + b_k^2 \leq a_k^2 + 2|a_k b_k| + b_k^2$$

Then, because

$$\sum_{k=1}^{\infty} a_k^2 + 2|a_k b_k| + b_k^2 \text{ converges}$$

By comparison test, we now know

$$\sum_{k=1}^{\infty} (a_k + b_k)^2 \text{ converges}$$

Notice

$$\sum_{k=1}^{\infty} a_k^2 \text{ converge} \implies \lim_{k \rightarrow \infty} a_k^2 = 0$$

Then we know there exists N such that

$$\forall n > N, a_n^2 < 1$$

Because $a_n^2 < 1 \implies |a_n| < 1$. This give us

$$\forall n, |a_n| \leq \max\{|a_1|, |a_2|, \dots, |a_N|, 1\}$$

We have proved $|a_n|$ is bounded. Notice that

$$\frac{1}{k} \searrow 0$$

Then by Dirichlet's Test, we are done. ■

Question 53

Does the series

$$\sum_{k=1}^{\infty} \frac{(-1)^k \ln(k+1)}{k}$$

converge? Does it converge absolutely? Justify your answer.

Proof. First notice

$$\frac{d}{dt} \frac{\ln(t+1)}{t} = \frac{\frac{t}{t+1} - \ln(t+1)}{t^2}$$

Because

$$\frac{t}{t+1} \nearrow 1 \text{ and } \ln(t+1) \nearrow \infty$$

We know for large k ,

$$\left. \frac{d}{dt} \frac{\ln(t+1)}{t} \right|_{t=k} = \frac{1}{k^2} \left(\frac{k}{k+1} - \ln(k+1) \right) < 0$$

By fundamental theorem of calculus,

$$\frac{\ln(k+2)}{k+1} - \frac{\ln(k+1)}{k} = \int_k^{k+1} \frac{d}{dt} \frac{\ln(t+1)}{t} dt < 0 \text{ for large } k$$

We now know

$$\frac{\ln(k+1)}{k} \text{ monotonically decrease for large } k$$

Notice that

$$\lim_{k \rightarrow \infty} \frac{\ln(k+1)}{k} \stackrel{H}{=} \lim_{k \rightarrow \infty} \frac{\frac{1}{k+1}}{1} = 0$$

We have proved $\frac{\ln(k+1)}{k} \searrow 0$ as $k \rightarrow \infty$. Then by Alternating Series Test, we know

$$\sum_{k=1}^{\infty} (-1)^k \frac{\ln(k+1)}{k} \text{ converges}$$

Notice

$$\int \frac{\ln k}{k} dk = \frac{(\ln k)^2}{2}$$

Then we have

$$\int_1^{\infty} \frac{\ln k}{k} dk = \frac{(\ln k)^2}{2} \Big|_{k=1}^{\infty} = \infty$$

Notice that

$$\frac{\ln k}{k} \leq \frac{\ln(k+1)}{k} \text{ for large } k$$

Then we know

$$\int_1^{\infty} \frac{\ln(k+1)}{k} dk = \infty$$

By integral test, we now know

$$\sum_{k=1}^{\infty} \left| (-1)^k \frac{\ln(k+1)}{k} \right| = \ln 2 + \sum_{k=2}^{\infty} \frac{\ln(k+1)}{k} \text{ diverges}$$

■

Question 54

Find all values of $p \in \mathbb{R}$ that make the following series converge absolutely:

$$\sum_{k=1}^{\infty} (-1)^k \frac{1}{\ln^p(k)}$$

Proof. Suppose $p \leq 0$. Let $q = -p$. We have

$$\left| (-1)^k \frac{1}{\ln^p(k)} \right| = \ln^q(k) \text{ where } q \geq 0$$

Then we have

$$\lim_{k \rightarrow \infty} \left| (-1)^k \frac{1}{\ln^p(k)} \right| = \lim_{k \rightarrow \infty} \ln^q(k) \neq 0$$

This tell us if $p \leq 0$, the series do not converge.

Suppose $p > 0$. We see

$$\ln(x^{\frac{1}{p}}) \leq x^{\frac{1}{p}} \text{ for large } x$$

Then for large x we have

$$\frac{1}{p} \ln x \leq x^{\frac{1}{p}}$$

This give us

$$\ln x \leq p x^{\frac{1}{p}}$$

This give us

$$(\ln x)^p \leq p^p x$$

This give us

$$\frac{1}{(\ln x)^p} \geq \frac{1}{p^p x}$$

Then because

$$\frac{1}{p^p} \sum_{x=1}^{\infty} \frac{1}{x} \text{ diverge}$$

By Comparison Test, we see

$$\sum_{x=1}^{\infty} \left| (-1)^x \frac{1}{\ln^p(x)} \right| = \sum_{x=1}^{\infty} \frac{1}{(\ln x)^p} \text{ diverge}$$

In conclusion, for all p , the series diverge. ■

Question 55

Let a_k and b_k be real sequences. Decide which of the following statements are true and which are false. Prove the true ones and give counterexamples to the false ones.

1. If $a_k \searrow 0$ as $k \rightarrow \infty$, and $\sum_{k=1}^{\infty} b_k$ converges conditionally, then

$$\sum_{k=1}^{\infty} a_k b_k$$

converges.

2. If $a_k \rightarrow 0$ as $k \rightarrow \infty$, then

$$\sum_{k=1}^{\infty} (-1)^k a_k$$

converges.

3. If $a_k \rightarrow 0$ as $k \rightarrow \infty$, and $a_k \geq 0$ for all $k \in \mathbb{N}$, then

$$\sum_{k=1}^{\infty} (-1)^k a_k$$

converges.

4. If $a_k \rightarrow 0$ as $k \rightarrow \infty$, and

$$\sum_{k=1}^{\infty} (-1)^k a_k$$

converges, then $a_k \searrow 0$ as $k \rightarrow \infty$.

Proof. From now, we let

$$A_n = \sum_{k=1}^n a_k \text{ and } B_n = \sum_{k=1}^n b_k$$

We claim

$$a_k \searrow 0 \text{ as } k \rightarrow \infty \text{ and } \sum_{k=1}^{\infty} b_k \text{ converges} \implies \sum_{k=1}^{\infty} a_k b_k \text{ converges}$$

Notice

$$\sum_{k=1}^{\infty} b_k \text{ converges} \implies \exists L \in \mathbb{R}, \lim_{n \rightarrow \infty} B_n = L$$

Let $\epsilon = 1$, we know there exists N such that

$$\forall n > N, B_n < L + 1$$

Then we see

$$\forall k, B_k < \max\{B_1 + 1, B_2 + 1, \dots, B_N + 1, L + 1\}$$

Then by Dirichlet's Test, we are done. (done)

Define

$$a_k = (-1)^k \frac{1}{k}$$

We have

$$\lim_{k \rightarrow \infty} a_k = 0 \text{ and } \sum_{k=1}^{\infty} (-1)^k a_k = \sum_{k=1}^{\infty} (-1)^{2k} \frac{1}{k} = \sum_{k=1}^{\infty} \frac{1}{k} \text{ diverges}$$

Define

$$a_k = \begin{cases} \frac{1}{\sqrt{\lceil \frac{k}{2} \rceil + 1} - 1} & \text{if } k \text{ is odd} \\ \frac{1}{\sqrt{\lceil \frac{k}{2} \rceil + 1} + 1} & \text{if } k \text{ is even} \end{cases}$$

The sequence we have is

$$a_1 = \frac{1}{\sqrt{2} - 1}, a_2 = \frac{1}{\sqrt{2} + 1}, a_3 = \frac{1}{\sqrt{3} - 1}, a_4 = \frac{1}{\sqrt{3} + 1}, a_5 = \frac{1}{\sqrt{4} - 1}, \dots$$

Because

$$\lim_{k \rightarrow \infty} \lceil \frac{k}{2} \rceil = \infty$$

We have

$$\lim_{k \rightarrow \infty} a_k = 0$$

Clearly, $\forall k, a_k \geq 0$. Define

$$B_n := \sum_{k=1}^n (-1)^k a_k$$

Notice

$$B_{2n} = \sum_{k=1}^n \frac{-1}{\sqrt{k+1}-1} + \frac{1}{\sqrt{k+1}+1} = \sum_{k=1}^n \frac{-2}{k}$$

Then we have

$$\lim_{n \rightarrow \infty} B_{2n} = \sum_{k=1}^n \frac{-2}{k} \text{ diverge}$$

This tell us

$$\sum_{n=1}^{\infty} (-1)^n a_n = \lim_{n \rightarrow \infty} B_n \text{ diverge}$$

Otherwise the sub-sequence $\lim_{n \rightarrow \infty} B_{2n}$ would have converge.

Define

$$a_n = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \frac{1}{(\frac{n}{2})^2} & \text{if } n \text{ is even} \end{cases}$$

The form of this sequence is

$$a_1 = 0, a_2 = \frac{1}{1^2}, a_3 = 0, a_4 = \frac{1}{2^2}, a_5 = 0, a_6 = \frac{1}{3^2}, a_7 = 0, \dots$$

Clearly, $a_n \rightarrow 0$. We see

$$\sum_{n=1}^{\infty} a_n = \sum_{n=0}^{\infty} a_{2n+1} + \sum_{n=1}^{\infty} a_{2n} = \sum_{n=0}^{\infty} 0 + \sum_{n=1}^{\infty} \frac{1}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} \text{ converge}$$

Yet, for all even n , we have

$$a_n = \frac{1}{(\frac{n}{2})^2} > 0 = a_{n+1}$$

So we don't have $a_n \searrow 0$ as $n \rightarrow \infty$. This serve as a counter example for the last question. ■

11.5 HW5

Question 56

1. Let (X, d) , (Y, d) be metric spaces; let $f : X \rightarrow Y$. Show that f is continuous on X if and only if for every open set $V \subseteq Y$, $f^{-1}(V)$ is open.

Proof. See Theorem 2.4.4 ■

Question 57

2. Let X and Y be metric spaces; let $f : X \rightarrow Y$. Show that the following statements are equivalent:
 1. f is continuous.
 2. For every subset A of X , one has $f(A) \subseteq \overline{f(A)}$.
 3. For every closed set B of Y , the set $f^{-1}(B)$ is closed in X .

Proof. See Theorem 2.4.4 ■

Question 58

3. Suppose that E is a nonempty subset of \mathbb{R} , let $a \in E$, and that $f : E \rightarrow \mathbb{R}$. If f has the property that if $\{x_n\} \subseteq E$ and x_n converges to a , then $f(x_n) \rightarrow f(a)$ as $n \rightarrow \infty$. Is f continuous at a ? Justify your result.

Proof. We prove that f is continuous at a . Assume f is not continuous at a . Then we know a is a limit point of E and

$$\lim_{x \rightarrow a} f(x) \neq f(a)$$

Then there exists a sequence $\{x_n\} \subseteq E$ that converge to a and satisfy

$$\forall n, a \neq x_n \text{ and } \lim_{n \rightarrow \infty} f(x_n) \neq f(a) \text{ CaC to premise}$$

Question 59

4. Let $f : M \rightarrow N$ be a map from metric space M to another metric space N with the property that if a sequence $\{p_n\}$ in M converges, then the sequence $\{f(p_n)\}$ in N also converges. Is f continuous? Justify your result.

Proof. We prove f is continuous. Assume f is not continuous at p . Then we know there exists a sequence $\{p_n\} \subseteq M$ such that

$$\lim_{n \rightarrow \infty} p_n = p \text{ and } \forall n \in \mathbb{N}, p_n \neq p$$

From premise, we know

$$\lim_{n \rightarrow \infty} f(p_n) \text{ exists}$$

Because f is not continuous at p , we know

$$\lim_{n \rightarrow \infty} f(p_n) \neq f(p)$$

Consider the sequence

$$q_n = \begin{cases} p_n & \text{if } n \text{ is odd} \\ p & \text{if } n \text{ is even} \end{cases}$$

$\{q_n\}$ clearly is a sequence that converge to p . By premise, we should have

$$\lim_{n \rightarrow \infty} f(q_n) \text{ exists}$$

Yet, we see

$$\lim_{n \rightarrow \infty} f(q_{2n}) = f(p) \neq \lim_{n \rightarrow \infty} f(p_n) = \lim_{n \rightarrow \infty} f(q_{2n+1}) \text{ CaC}$$

■

Question 60

5. Suppose f is a real function defined on \mathbb{R} which satisfies

$$\lim_{h \rightarrow 0} [f(x+h) - f(x-h)] = 0$$

for every $x \in \mathbb{R}$. Does this imply that f is continuous?

Proof. No. Define

$$x \mapsto \begin{cases} 0 & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

Clearly f is not continuous at 0. We now prove f satisfy the premise.

Let $y \neq 0$. We see

$$|h| \leq |y| \implies f(y+h) - f(y-h) = 0 - 0 = 0 \in B_\epsilon(0)$$

This shows

$$\lim_{h \rightarrow 0} [f(y+h) - f(y-h)] = 0$$

Let $y = 0$. We see

$$0 < h \implies f(y+h) - f(y-h) = 1 - 1 = 0 \in B_\epsilon(0)$$

■

Question 61

6. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : I(\subseteq \mathbb{R}) \rightarrow \mathbb{R}$ be a map. Please state the definition of following nouns. (There is no standard answer, you can write down all equivalent statements as possible as you can)

1. f is continuous on \mathbb{R} .
2. f is uniformly continuous on \mathbb{R} .
3. f is Lipschitz continuous on \mathbb{R} .
4. g is absolutely continuous on I .

Proof. For continuous, see Definition 3.11.7, Definition 3.11.6 and Definition 3.11.1.

For uniform continuous, see Definition 5.5.1.

For Absolute continuous, see Definition 5.5.1.

For Lipschitz continuous, see Definition 5.5.4

■

Question 62

7. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : I(\subseteq \mathbb{R}) \rightarrow \mathbb{R}$ be a map. Prove or disprove following statement.

- (i) if f is continuous on \mathbb{R} , then f is uniformly continuous on \mathbb{R} .
- (ii) if f is uniformly continuous on \mathbb{R} , then f is continuous on \mathbb{R} .
- (iii) if g is continuous on I and I is compact, then g is uniformly continuous on I .
- (iv) if f is continuous on \mathbb{R} , then f is Lipschitz continuous on \mathbb{R} .
- (v) if f is Lipschitz continuous on \mathbb{R} , then f is continuous on \mathbb{R} .
- (vi) if g is Lipschitz continuous on I , then g is uniformly continuous on I .
- (vii) if g is uniformly continuous on I , then g is Lipschitz continuous on I .
- (viii) if g is absolutely continuous on I , then g is uniformly continuous on I .

- (ix) if g is uniformly continuous on I , then g is absolutely continuous on I .
- (x) if g is absolutely continuous on I , then g is Lipschitz continuous on I .
- (xi) if g is Lipschitz continuous on I , then g is absolutely continuous on I .

Proof. (i) No, use x^2 on non-compact set. See Theorem 5.5.4.

(ii) Yes, see Theorem 5.5.3.

(iii) Yes. This is Heine-Cantor Theorem. See Theorem 5.5.1.

(iv) No, use \sqrt{x} . See Theorem 5.5.8, and notice that absolutely continuous is stronger than continuous. This fact is proved by Theorem 5.5.2 and Theorem 5.5.3.

(v) Yes. See Theorem 5.5.5, and Theorem 5.5.3

(vi) Yes. See Theorem 5.5.5.

(vii) No. Use \sqrt{x} . See Theorem 5.5.8 and notice that absolutely continuous is stronger than uniformly continuous. See Theorem 5.5.2.

(viii) Yes. See Theorem 5.5.2.

(ix) No. See Theorem 5.5.3.

(x) No. Use \sqrt{x} . See Theorem 5.5.8

(xi) Yes. See Theorem 5.5.6

■

Question 63

8. Suppose that $f : [0, \infty) \rightarrow \mathbb{R}$ is continuous and there is an $L \in \mathbb{R}$ such that $f(x) \rightarrow L$ as $x \rightarrow \infty$. Prove that f is uniformly continuous on $[0, \infty)$.

Proof. Fix ϵ . Because $f(x) \rightarrow L$ as $x \rightarrow \infty$, we know there exists $R \in \mathbb{R}^+$ such that

$$\forall x > R, |f(x) - L| < \frac{\epsilon}{2}$$

Also, because $[0, R+1]$ is compact and f is continuous, we know f is uniformly continuous on $[0, R+1]$, then we know there exists δ' such that

$$\forall y', z' \in [0, R+1], |y' - z'| < \delta' \implies |f(y') - f(z')| < \epsilon$$

We now prove

$$\delta := \min\{\delta', 1\} \text{ works}$$

WOLG, let $z + \delta > y > z \geq 0$. If $z < R$, then $y < z + \delta < R + \delta \leq R + 1$. This tells us $z, y \in [0, R + 1]$. Then because $y - z < \delta < \delta'$, we have $|f(y) - f(z)| < \epsilon$. If $z \geq R$. We have $y > z > R$. Then we have

$$|f(y) - f(z)| \leq |f(y) - L| + |f(z) - L| < \epsilon$$

■

Question 64

9. Show that if f is monotone on an interval I , then f has at most countably many points of discontinuity on I .

Proof. See Theorem 3.11.17.

■

11.6 HW6

Question 65

1. Let $f(x, y) = \frac{x^2y}{x^4+y^2}$ for $(x, y) \neq (0, 0)$, prove that $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$ does NOT exist. (Hint: Choose $a_n = (0, \frac{1}{n})$ and $b_n = (\frac{1}{n}, \frac{1}{n^2})$, evaluate $\lim_{n \rightarrow \infty} f(a_n)$ and $\lim_{n \rightarrow \infty} f(b_n)$, and conclude by sequential characterization of limit.)

Proof. Let

(a) $a_n = (0, \frac{1}{n})$

(b) $b_n = (\frac{1}{n}, \frac{1}{n^2})$

It is clear that

(a) $a_n \rightarrow 0$

(b) $b_n \rightarrow 0$

(c) $f(a_n) = 0$

(d) $f(b_n) = \frac{1}{2}$

This conclude $\lim_{(x,t) \rightarrow (0,0)} f(x, y)$ D.N.E. ■

Question 66

2. Let $f(x, y) = \frac{x^3 - x^2y + y^3}{x^2 + y^2}$ for $(x, y) \neq (0, 0)$. Use $\varepsilon - \delta$ definition of continuity to prove that $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$.

Proof.

$$\begin{aligned} |f(x, y)| &\leq \left| \frac{x^3}{x^2 + y^2} \right| + \left| \frac{x^2y}{x^2 + y^2} \right| + \left| \frac{y^3}{x^2 + y^2} \right| \\ &\leq \left| \frac{x^3}{x^2} \right| + \left| \frac{x^2y}{x^2} \right| + \left| \frac{y^3}{y^2} \right| \quad (\because x^2 + y^2 \geq x^2 \text{ and } x^2 + y^2 \geq y^2) \\ &= |x| + 2|y| \end{aligned}$$

Let $\delta = \frac{\epsilon^2}{9}$. We now see

$$x^2 + y^2 < \delta \implies x^2 < \delta \implies |x| < \sqrt{\delta} = \frac{\epsilon}{3}$$

And also, $|y| < \frac{\epsilon}{3}$. Then we have

$$|f(x, y)| \leq |x| + 2|y| < \epsilon$$

■

Question 67

3. (a) (Converse statement of extreme value theorem) If $f : [a, b] \rightarrow \mathbb{R}$ admits its maximum and minimum, does this imply f continuous? Justify your result.
 (b) (Converse statement of intermediate value theorem) Let $f : [a, b] \rightarrow \mathbb{R}$ with $f(a) < f(b)$, if for all L in $(f(a), f(b))$, there always exists some c in (a, b) such that $f(c) = L$, does this imply f continuous? Justify your result.

Proof. For (a), No. We see

$$f(x) = \begin{cases} 2 & \text{if } x = b \\ 1 & \text{if } x \in (a, b) \\ 0 & \text{if } x = a \end{cases}$$

has maximum $f(b) = 2$ and minimum $f(a) = 0$, yet it is clearly not continuous at a and b .

For (b), No. We see

$$f(x) = \begin{cases} 2 & \text{if } x = 3 \\ 1 & \text{if } x \in (2, 3) \\ 2(x-1) & \text{if } x \in [1, 2] \\ 1 & \text{if } x \in (0, 1) \\ 0 & \text{if } x = 0 \end{cases}$$

We have

$$f(a) = 0 < f(b) = 2$$

and for all $L \in (0, 2) = (f(a), f(b))$

$$x = \frac{L}{2} + 1 \implies x \in (1, 2) \implies f(x) = L$$

That is $f(\frac{L}{2} + 1) = L$. Notice that f has jump discontinuous at $\{0, 1, 2, 3\}$ and we are done. ■

Question 68

4. Let $f : [a, b] \rightarrow \mathbb{R}$ satisfy

(a) $\forall (c_1, c_2) \subseteq [a, b], \forall L \in [f(c_1), f(c_2)], \exists c \in (c_1, c_2), f(c) = L$

(b) $\forall r \in \mathbb{Q}, f^{-1}[r]$ is closed

Prove f is continuous.

(Hint: If not, there exists a sequence $\{x_n\}_{n \in \mathbb{N}} \rightarrow x_0$ but $f(x_n)$ doesn't converge to $f(x_0)$, which means for large n , there's still a gap between $f(x_n)$ and $f(x_0)$. Can you relate this with \mathbb{Q} ?)

Proof. Assume f has a discontinuity at x . Then there exists a sequence $x_n \rightarrow x$ such that $f(x_n) \not\rightarrow f(x)$. Because $f(x_n) \not\rightarrow f(x)$, we know there exists a sub-sequence $x_{n_k} \rightarrow x$ such that $f(x_{n_k}) \rightarrow y \neq f(x)$. Notice that it is possible $y = \pm\infty$, which won't damage the validity of the proof. WOLG, suppose $y > f(x)$. Let N satisfy

$$\forall k > N, f(x_{n_k}) > \frac{y + f(x)}{2}$$

Because \mathbb{Q} is dense in \mathbb{R} , we can let L satisfy

$$L \in (f(x), \frac{f(x) + y}{2}) \cap \mathbb{Q}$$

For each $k > N$, because $f(x_{n_k}) > \frac{y + f(x)}{2}$, we see

$$L \in (f(x), f(x_{n_k}))$$

By premise (a), we know there exists y_{n_k} between x and x_{n_k} such that $f(y_{n_k}) = L$. Notice that because y_{n_k} is bounded by x_{n_k} and x , where $x_{n_k} \rightarrow x$, we know $y_{n_k} \rightarrow x$. Notice that $y_{n_k} \in f^{-1}[L]$ and that $x \notin f^{-1}[L]$ **CaC** to that $f^{-1}[L]$ is closed. ■

Question 69

5. Let (X, d) be a metric space and $f : X \rightarrow \mathbb{R}$ be a continuous function. Show that $Z(f) = \{x \in X \mid f(x) = 0\}$ is closed. Please state at least two different methods.

Proof. (Method 1) Assume $Z(f)$ is not closed. Then there exists a convergent sequence $x_n \rightarrow x$ in $Z(f)$ such that $x \notin Z(f)$. Because $x \notin Z(f)$, we see $f(x) \neq 0$. Notice that x_n is a sequence in $Z(f)$, so we have

$$\forall n, f(x_n) = 0$$

Then

$x_n \rightarrow x$ and $f(x_n) \rightarrow 0 \neq f(x)$ **CaC** to sequential definition of continuous

(Method 2) Because $\{0\}$ is closed, we see the inverse image $Z(f) = \{x \in X : f(x) = 0\} = f^{-1}[\{0\}]$ must be closed. ■

Question 70

6. Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \rightarrow Y$. Show that if f is continuous and $f = 0$ on some dense subset A of X , then $f \equiv 0$ on X .

Proof. Let $x \in X \setminus A$. We wish to prove

$$f(x) = 0$$

Because A is dense in X , we know there exists a sequence $a_n \rightarrow x$ in A that converge to x . Because f is continuous, we now have

$$f(x) = \lim_{n \rightarrow \infty} f(a_n) = 0 (\because \forall n, f(a_n) = 0) \text{ (done)}$$

■

Question 71

7. (a) Define $f(x) = \begin{cases} \frac{1}{n} & \text{if } x = \frac{m}{n} \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$ Determine whether f is continuous at each irrational point. Justify your result.

(b) Define $f(x) = \begin{cases} \frac{1}{n} & \text{if } x = \frac{m}{2^n} \in \mathbb{Q} \cap [0, 1] \text{ for some } m, n \in \mathbb{N} \\ 0 & \text{if } x \in [0, 1] \setminus \mathbb{Q} \end{cases}$ Determine whether f is continuous at each irrational point. Justify your result.

Proof. We prove **f in (a) is continuous at all irrational point**. Let $x_n \rightarrow x \in \mathbb{R} \setminus \mathbb{Q}$. We wish to prove $f(x_n) \rightarrow f(x)$. Because f is 0 at all irrational, WOLG, we can suppose x_n are all dyadic rationals. Fix ϵ , we wish to find N such that

$$\forall n > N, f(x_n) < \epsilon$$

Let m be the smallest natural such that $\epsilon > \frac{1}{m}$. For each $k \leq m$, let $\frac{q_k}{k}$ satisfy

$$x \in \left(\frac{q_k}{k}, \frac{q_k + 1}{k} \right)$$

Define

$$\alpha = \min \left(\left\{ \left| x - \frac{q_k}{k} \right| \in \mathbb{R} : \forall k \leq m \right\} \cup \left\{ \left| x - \frac{q_k + 1}{k} \right| \in \mathbb{R} : \forall k \leq m \right\} \right)$$

Let N satisfy

$$\forall n > N, x_n \in (x - \alpha, x + \alpha)$$

We prove such N works. Let $n > N$. We wish to prove

$$f(x_n) < \epsilon$$

Because x_n is rational, we can express x_n in the form

$$x_n = \frac{q}{p} \text{ where } \gcd(q, p) = 1 \text{ and } p \in \mathbb{N}$$

Assume $p \leq m$. Notice that $x \in (\frac{q_p}{p}, \frac{q_p+1}{p})$ and $\alpha \leq \min\{d(x, \frac{q_p}{p}), d(x, \frac{q_p+1}{p})\}$, this tell us

$$x + \alpha \leq \frac{q_p + 1}{p} \text{ and } x - \alpha \geq \frac{q_p}{p}$$

Then we see

$$x_n \in (x - \alpha, x + \alpha) \subseteq (\frac{q_p}{p}, \frac{q_p + 1}{p})$$

Which is impossible. **CaC**

We now have $p > m$. Then we see

$$f(x_n) = \frac{1}{p} < \frac{1}{m} < \epsilon \text{ (done)}$$

We now prove f in (b) is continuous at all irrationals. Let $x_n \rightarrow x \in \mathbb{R} \setminus \mathbb{Q}$. We wish to prove $f(x_n) \rightarrow 0$. Because f is 0 at all points that isn't dyadic irrationals, WOLG, we can suppose x_n are all dyadic rationals. Fix ϵ , we wish to find N such that

$$\forall n > N, f(x_n) < \epsilon$$

Let m be the smallest naturals such that

$$\frac{1}{m} < \epsilon$$

For each $k \leq m$, let $\frac{q_k}{2^k}$ satisfy

$$x \in (\frac{q_k}{2^k}, \frac{q_k + 1}{2^k})$$

Define

$$\alpha = \min \left(\left\{ d\left(x, \frac{q_k}{2^k}\right) \in \mathbb{R} : k \leq m \right\} \cup \left\{ d\left(x, \frac{q_k+1}{2^k}\right) \in \mathbb{R} : k \leq m \right\} \right)$$

Because $x_n \rightarrow x$, we know there exists N such that

$$n > N \implies d(x_n, x) < \alpha$$

We prove such N works. Let $n > N$. We wish to prove

$$f(x_n) < \epsilon$$

Because x_n is dyadic rational, we can express x_n in the form of

$$x_n = \frac{q}{2^p} \text{ where } q \text{ is odd}$$

Assume $p \leq m$. Notice that $x \in (\frac{q_p}{2^p}, \frac{q_p+1}{2^p})$ and $\alpha \leq \min\{d(x, \frac{q_p}{2^p}), d(x, \frac{q_p+1}{2^p})\}$. This tell us

$$x_n \in (x - \alpha, x + \alpha) \subseteq (\frac{q_p}{2^p}, \frac{q_p+1}{2^p})$$

which is impossible. **CaC**

We now have $p > m$. Then we see

$$f(x_n) = \frac{1}{p} < \frac{1}{m} < \epsilon \text{ (done)}$$

■

Question 72

8. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. If for all open subset U in \mathbb{R} , $f(U)$ is open in \mathbb{R} , prove that f is monotone. (Hint: Observe $f(x) = x^2$, which is not monotone on \mathbb{R} , does $f((-1, 1))$ open in \mathbb{R} ? You may use extreme value theorem to help you complete the whole proof.)

Proof. Assume f is not monotone. WOLG, suppose f has a local maximum at x

$$\forall y \in (x - \epsilon, x + \epsilon), f(y) \leq f(x)$$

Because f is continuous, we know

$$f[(a, b)] \text{ is an interval}$$

By the premise that f is an open mapping, we then can express

$$f[(a, b)] = (c, d) \text{ CaC } x \text{ is a local maximum}$$

since (c, d) has no maximum.

■

Question 73

9. We say $f : (a, b) \rightarrow \mathbb{R}$ is "convex" if $\forall x, y \in (a, b), \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$

(a) For $a < s < t < u < b$, show that

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t}$$

(b) Prove that every convex function is continuous. (Hint: Use the result in (a): fix u , as $t \rightarrow s^+$, what happens on $f(t) - f(s)$?)

Proof. See the proof of Lemma 5.2.8 and Theorem 5.2.9 ■

Question 74

10. If $f : (a, b) \rightarrow \mathbb{R}$ is continuous and $\forall x, y \in (a, b)$,

$$f\left(\frac{x + y}{2}\right) \leq \frac{f(x) + f(y)}{2}$$

Prove that f is convex. (Hint: for $\lambda \in [0, 1]$, write $\lambda = \sum_{n=0}^{\infty} \frac{a_n}{2^n}$, where $a_n = 0$ or 1 for all n .)

Proof. See Theorem 5.2.9 and Theorem 5.2.7. ■

11.7 HW7

Question 75

Let f be a real-valued function defined on an open interval E such that

$$|f(x) - f(y)| \leq |x - y|^\alpha, \forall x, y \in E$$

Then show that (a) f is continuous on E , if $\alpha > 0$ (b) f is constant on E , if $\alpha > 1$.

Proof. Suppose $\alpha > 0$. Let $\delta = \epsilon^{\frac{1}{\alpha}}$. We see

$$|x - y| \leq \delta = \epsilon^{\frac{1}{\alpha}} \implies |f(x) - f(y)| \leq |x - y|^\alpha \leq (\epsilon^{\frac{1}{\alpha}})^\alpha = \epsilon$$

We have proved f is uniformly continuous on E , which implies f is continuous on E .

Suppose $\alpha > 1$. Arbitrarily pick $x \in E$. We only wish to prove

$$f'(x) = 0$$

Observe

$$\forall y \in E, \left| \frac{f(x) - f(y)}{x - y} \right| \leq |x - y|^{\alpha-1}$$

Then because $\alpha - 1 > 0$, we have

$$|x - y|^{\alpha-1} \rightarrow 0 \text{ as } y \rightarrow x$$

This give us

$$f'(x) = \lim_{y \rightarrow x} \frac{f(x) - f(y)}{x - y} = 0 \text{ (done)}$$

Now, by Mean Value Theorem, we see f must stay constant on E . ■

Question 76

Let f be a differentiable real function defined in (a, b) . Prove that f is convex if and only if f' is monotonically increasing. Assume next that $f''(x)$ exists for every $x \in (a, b)$, and prove that f is convex if and only if $f''(x) \geq 0$ for $x \in (a, b)$.

Proof. We first prove

$$f \text{ is convex} \implies f' \text{ increase}$$

Let $(c, d) \subseteq (a, b)$. We wish to prove

$$f'(c) \leq f'(d)$$

Let $x \in (c, d)$, by Lemma 5.2.8, we see

$$\frac{f(x) - f(c)}{x - c} \leq \frac{f(d) - f(c)}{d - c}$$

Taking limit $x \rightarrow c$, we see

$$f'(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c} \leq \frac{f(d) - f(c)}{d - c}$$

Again by Lemma 5.2.8, we see

$$\frac{f(d) - f(c)}{d - c} \leq \frac{f(d) - f(x)}{d - x}$$

Again taking limit $x \rightarrow d$, we see

$$f'(d) = \lim_{x \rightarrow d} \frac{f(d) - f(x)}{d - x} \geq \frac{f(d) - f(c)}{d - c}$$

We now have

$$f'(c) \leq \frac{f(d) - f(c)}{d - c} \leq f'(d) \text{ (done)}$$

We now prove

$$f' \text{ increase} \implies f \text{ is convex}$$

Fix $(x_1, x_2) \subseteq (a, b)$, $\lambda \in [0, 1]$, $x_\lambda = \lambda x_1 + (1 - \lambda)x_2$. We only wish to prove

$$f(x_\lambda) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

By Mean Value Theorem, we know there exists $c_1 \in (x_1, x_\lambda)$, $c_2 \in (x_\lambda, x_2)$ such that

$$\begin{aligned} f'(c_1) &= \frac{f(x_\lambda) - f(x_1)}{x_\lambda - x_1} \\ f'(c_2) &= \frac{f(x_2) - f(x_\lambda)}{x_2 - x_\lambda} \end{aligned}$$

Because $c_1 < x_\lambda < c_2$ and f' is increasing, we have

$$\frac{f(x_\lambda) - f(x_1)}{x_\lambda - x_1} = f'(c_1) \leq f'(c_2) = \frac{f(x_2) - f(x_\lambda)}{x_2 - x_\lambda}$$

Acknowledging $x_1 < x_\lambda < x_2$, we now have

$$(x_2 - x_\lambda)[f(x_\lambda) - f(x_1)] \leq (x_\lambda - x_1)[f(x_2) - f(x_\lambda)]$$

Substituting x_λ with $\lambda x_1 + (1 - \lambda)x_2$, we now can deduce

$$\begin{aligned} \lambda(x_2 - x_1)[f(x_\lambda) - f(x_1)] &\leq (1 - \lambda)(x_2 - x_1)[f(x_2) - f(x_\lambda)] \\ \implies \lambda[f(x_\lambda) - f(x_1)] &\leq (1 - \lambda)[f(x_2) - f(x_\lambda)] \\ \implies f(x_\lambda) &\leq \lambda f(x_1) + (1 - \lambda)f(x_2) \text{ (done)} \end{aligned}$$

Following the result above and given the hypothesis, we now prove

$$f' \text{ increase} \implies \forall x \in (a, b), f''(x) \geq 0$$

Because f' increase, we can deduce

$$f''(x) = \lim_{t \rightarrow x^-} \frac{f(x) - f(t)}{x - t} \geq 0 \text{ (done)}$$

We now prove

$$\forall x \in (a, b), f''(x) \geq 0 \implies f' \text{ increase}$$

Assume $\exists(x_1, x_2) \subseteq (a, b), f'(x_1) > f'(x_2)$. By Mean Value Theorem, we know there exists $c \in (x_1, x_2)$ such that

$$f''(c) = \frac{f'(x_2) - f'(x_1)}{x_2 - x_1} < 0 \text{ CaC}$$

■

Question 77

Let f be a bounded real-valued function defined on (a, ∞) for some $a \in \mathbb{R}$ such that f' and f'' exist on (a, ∞) . Further, if M_0, M_1, M_2 are the least upper bounds of $|f|, |f'|, |f''|$, on (a, ∞) , respectively. Prove that $M_1^2 \leq 4M_0M_2$.

Proof. Arbitrarily pick $x \in (a, \infty)$ and let $h \in \mathbb{R}^+$. By Theorem 6.5.1 (Taylors' Theorem) and Corollary 6.5.3 (Lagarange Form of Remainder), expanding f at x and approximate $f(x + 2h)$, we have

$$(a) f(x + 2h) = P_1(x + 2h) + R_1(x + 2h)$$

$$(b) P_1(x + 2h) = f(x) + 2hf'(x)$$

$$(c) \exists \xi \in (x, x + 2h), R_1(x + 2h) = 2h^2 f''(\xi)$$

Express (a) in the form

$$f(x + 2h) = f(x) + 2hf'(x) + 2h^2 f''(\xi)$$

This give us

$$f'(x) = \frac{f(x + 2h) - f(x)}{2h} + hf''(\xi)$$

Then we can deduce

$$\begin{aligned} |f'(x)| &\leq \frac{|f(x + 2h)| + |f(x)|}{2h} + h|f''(\xi)| \\ &= \frac{M_0}{h} + hM_2 \end{aligned}$$

Then we have

$$|f'(x)|^2 \leq \frac{M_0^2}{h^2} + h^2 M_2^2 + 2M_0 M_2$$

We can deduce

$$|f'(x)|^2 - 4M_0 M_2 \leq \left(\frac{M_0}{h} - hM_2\right)^2$$

Notice that if $M_2 = 0$, because f is bounded, we must have $M_1 = 0$ and f is constant, then the proof become trivial. We only have to consider when $M_2 > 0$. Let $h = \sqrt{\frac{M_0}{M_2}}$. We see

$$|f'(x)|^2 - 4M_0 M_2 \leq \left(\frac{M_0}{h} - hM_2\right)^2 = 0$$

This give us

$$|f'(x)|^2 \leq 4M_0 M_2$$

Because x is arbitrarily picked from (a, ∞) , we now have

$$\forall x \in (a, \infty), |f'(x)| \leq 2\sqrt{M_0 M_2}$$

Then by the definition of least upper bound M_1

$$M_1 \leq 2\sqrt{M_0 M_2}$$

The result follows from squaring both side.

■

Question 78

Let f be a real-valued function defined on $[-1, 1]$ such that $f^{(n)}$ exists on $[-1, 1]$ for $n = 1, 2, 3$, and

$$\begin{aligned} f(-1) &= 0, \\ f(0) &= 0, \\ f(1) &= 1, \\ f'(0) &= 0. \end{aligned}$$

Prove that $f^{(3)}(x) \geq 3$ for some $x \in (-1, 1)$.

Proof. Taylor expanding f at 0 to polynomial of degree 2 and express the remainder in Lagrange form, we have

$$(a) \exists b \in (0, 1), f(1) = f(0) + f'(0) + \frac{f''(0)}{2} + \frac{f'''(b)}{6}$$

$$(b) \exists a \in (-1, 0), f(-1) = f(0) - f'(0) + \frac{f''(0)}{2} - \frac{f'''(a)}{6}$$

Plugging the value given and minus the two equations, we have

$$1 = \frac{f'''(b) + f'''(a)}{6}$$

This give us

$$f'''(a) + f'''(b) = 6$$

Then if $f'''(a) \geq 3$, we get the solution a . If $f'''(a) < 3$, we have $f'''(b) > 3$, which implies b is a solution. ■

Question 79

1. Let $f(x) = \exp\left(-\frac{1}{|x|}\right)$, $x \neq 0$ and $f(0) = 0$. Then calculate $f^{(n)}(0)$ by definition of derivative $\forall n \in \mathbb{N}$.
2. Show that $f(x)$ can NOT have its Taylor expansion at $x = 0$ (or, say the expand radius of $f(x)$ at $x = 0$ is zero).

Proof. We show that

$$\forall n \in \mathbb{N} \cup \{0\}, f^{(n)}(x) = \begin{cases} g_n(x)e^{\frac{-1}{x}} & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ p_n(x)e^{\frac{1}{x}} & \text{if } x < 0 \end{cases} \quad \text{where } \forall n \in \mathbb{N} \cup \{0\}, g_n, p_n \text{ are polynomials}$$

by induction.

Base Case:

$$g_0(x) = 1 = p_0(x)$$

Induction Case: Suppose

$$f^{(k)}(x) = \begin{cases} g_k(x)e^{-\frac{1}{x}} & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ p_k(x)e^{\frac{1}{x}} & \text{if } x < 0 \end{cases} \quad \text{where } g_k, p_k \text{ are polynomial}$$

Compute

$$f^{(k+1)}(x) = \begin{cases} g'_k(x)e^{-\frac{1}{x}} + (x^{-2})g_k(x)e^{-\frac{1}{x}} = [g'_k(x) + (x^{-2})g_k(x)]e^{-\frac{1}{x}} & \text{if } x > 0 \\ p'_k(x)e^{\frac{1}{x}} + (-x^{-2})p_k(x)e^{\frac{1}{x}} = [p'_k(x) + (-x^{-2})p_k(x)]e^{\frac{1}{x}} & \text{if } x < 0 \end{cases}$$

Notice that polynomials forms a ring and is closed under differentiation. We have proved g_{k+1}, p_{k+1} are polynomials. Because g_k, p_k are polynomial, we can compute

$$\lim_{h \rightarrow 0^+} \frac{g_k(h)e^{-\frac{1}{h}}}{h} = 0 = \lim_{h \rightarrow 0^-} \frac{p_k(h)e^{\frac{1}{h}}}{h}$$

Then we see

$$f^{(k+1)}(0) = \lim_{h \rightarrow 0} \frac{f^{(k)}(h) - f^{(k)}(0)}{h} = \lim_{h \rightarrow 0} \frac{f^{(k)}(h)}{h} = 0 \text{ (done)}$$

Following the result above, we see the Maclaurin Series of $f(x)$ is

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)x^k}{k!} = 0 \neq f(y) \text{ if } y \neq x$$

■

Question 80

Let f have a continuous derivative in the interval $[a, b]$, and let $f''(x) \geq 0$ for every value of x . Then if ξ is any point in the interval, the curve nowhere falls below its tangent at the point $x = \xi, y = f(\xi)$. (Hint: It suffices to show $f(x) \geq f(\xi) + f'(\xi)(x - \xi)$ on $[a, b]$, i.e. consider the Taylor expansion of $f(x)$ at $x = \xi$).

Proof. Taylor expand f at ξ and express the remainder in Lagrange Mean Value Form (Corollary 6.5.3). We have

$$\exists \xi_0 \in (x, \xi) \text{ or } (\xi, x), f(x) = f(\xi) + f'(\xi)(x - \xi) + f''(\xi_0)(x - \xi)^2$$

Then because $f''(\xi_0) \geq 0$ and $(x - \xi)^2 \geq 0$, we now see

$$f(x) \geq f(\xi) + f'(\xi)(x - \xi)$$

■

Lemma 11.7.1. Given

- (a) f is continuous on $[a, b]$
- (b) $\exists c \in [a, b], f(c) > 0$

Then there exists δ such that

$$\forall y \in [c - \delta, c + \delta], f(y) > \frac{f(c)}{2}$$

Proof. Because f is continuous at c , we know

$$\lim_{x \rightarrow c} f(x) = f(c) > \frac{f(c)}{2}$$

The Lemma now follows from definition of limit and taking $\delta = \frac{\delta(\frac{f(c)}{2})}{2}$

■

Question 81

Let f be a continuous function on $[a, b]$, then prove that $f(x) = 0$ for all $x \in [a, b]$ if and only if

- (a) $\int_a^b |f(x)| dx = 0$,
- (b) $\int_a^c f(x) dx = 0, \quad \forall c \in [a, b]$.

Proof. We first prove

$$\forall x \in [a, b], f(x) = 0 \implies \int_a^b |f(x)| dx = 0$$

Let P be an arbitrarily partition of $[a, b]$. Express

$$P = \{x_0, x_1, \dots, x_n\}$$

Because $\forall i, \{f(x) : x \in [x_{i-1}, x_i]\} = \{0\}$, we know

$$\forall i, m_i = 0 = M_i$$

Then we have

$$L(P, f) = \sum_{i=1}^n m_i \Delta x_i = 0$$

$$U(P, f) = \sum_{i=1}^n M_i \Delta x_i = 0 \text{ (done)}$$

We now prove

$$\int_a^b |f(x)| dx = 0 \implies \forall x \in [a, b], f(x) = 0$$

Assume $\exists c \in [a, b], f(c) = p \neq 0$. WOLOG, suppose $p > 0$. Let P be an arbitrary partition of $[a, b]$. Express

$$P = \{x_0, x_1, \dots, x_n\}$$

Let δ be from Lemma 11.7.1. Let j be the greatest and k be the smallest such that

$$x_j \leq c - \delta \text{ and } x_k \geq c + \delta$$

We then see

$$\begin{aligned} U(P, f) &= \sum_{i \leq j \text{ or } k < i \leq n} M_i \Delta x_i + \sum_{i=j+1}^k M_i \Delta x_i \\ &\geq \sum_{i=j+1}^k M_i \Delta x_i (\because |f| \geq 0) \\ &\geq \sum_{i=j+1}^k \frac{f(c)}{2} \Delta x_i \\ &= \frac{f(c)}{2} (x_k - x_j) \geq f(c) \delta \end{aligned}$$

Because P is an arbitrary partition, we now see

$$0 < f(c) \delta \leq \overline{\int_a^b |f| dx} \text{ CaC } \int_a^b |f| dx = 0 \text{ (done)}$$

The proof for

$$\forall x \in [a, b], f(x) = 0 \implies \forall c \in [a, b], \int_a^c f dx = 0$$

is trivial. We now prove

$$\forall c \in [a, b], \int_a^c f(x) dx = 0 \implies \forall x \in [a, b], f(x) = 0$$

Assume $\exists c \in [a, b], f(c) = p \neq 0$. WOLG, suppose $p > 0$. Let δ be from Lemma 11.7.1. We see $\frac{f(c)}{2}$ is a lower bound of f on $[c - \delta, c + \delta]$. Then because f is continuous, we can write

$$\int_{c-\delta}^{c+\delta} f dx \geq f(c)\delta > 0$$

Then because $\int_0^{c+\delta} f dx = 0$, we have

$$\int_0^{c-\delta} f dx = \int_0^{c+\delta} f dx - \int_{c-\delta}^{c+\delta} f dx < 0 \text{ CaC (done)}$$

■

Question 82

Prove that if $\lim_{b \rightarrow \infty} \int_1^b |f(x)| dx$ exists and finite, then $\lim_{n \rightarrow \infty} \int_1^\infty f(x^n) dx = 0$.

Proof. Let $u = x^n, du = nx^{n-1} dx$, we have

$$\begin{aligned} \int_1^\infty f(x^n) dx &= \int_1^\infty f(u) \frac{1}{nu^{\frac{n-1}{n}}} du \\ &\leq \int_1^\infty f(u) \frac{1}{nu} du \end{aligned}$$

■

Question 83

Prove that if f is continuous on $[a, b]$, and ϕ is monotone increasing on $[a, b]$, then there exists $\xi \in [a, b]$ such that

$$\int_a^b f d\phi = f(\xi)(\phi(b) - \phi(a))$$

Proof. WOLG, suppose ϕ increase on $[a, b]$. If $\phi(a) = \phi(b)$, then we see ϕ stay constant, so for every partition P , we have

$$\forall i, \Delta\phi(x_i) = 0$$

Then we have $\int_a^b f d\phi = 0$. It is clear that $\phi(a) = \phi(b) \implies f(\xi)(\phi(b) - \phi(a)) = 0$. We have proved the $\phi(a) = \phi(b)$ case. We now can suppose $\phi(a) < \phi(b)$. Notice that we wish to find $\xi \in [a, b]$ such that

$$f(\xi) = \frac{\int_a^b f d\phi}{\phi(b) - \phi(a)}$$

Because f is continuous on $[a, b]$ thus satisfying IVT, we can reduced the problem into proving

$$\inf_{x \in [a, b]} f \leq \frac{\int_a^b f d\phi}{\phi(b) - \phi(a)} \leq \sup_{x \in [a, b]} f$$

WOLG, we only prove

$$\frac{\int_a^b f d\phi}{\phi(b) - \phi(a)} \geq \inf_{x \in [a, b]} f$$

This is equivalent to

$$\int_a^b f d\phi \geq (\phi(b) - \phi(a)) \inf_{x \in [a, b]} f$$

Let $P = \{x_0, \dots, x_n\}$ be an arbitrary partition of $[a, b]$. We see

$$\begin{aligned} U(P, f) &= \sum_{i=1}^n m_i \Delta\phi(x_i) \\ &\geq \sum_{i=1}^n \inf_{x \in [a, b]} f \Delta\phi(x_i) \\ &= (\phi(b) - \phi(a)) \inf_{x \in [a, b]} f \end{aligned}$$

Because P is arbitrary, we now have

$$\overline{\int_a^b f d\phi} \geq (\phi(b) - \phi(a)) \inf_{x \in [a, b]} f \text{ (done)}$$

■

Lemma 11.7.2. Suppose f, g are both bounded on $[x_{i-1}, x_i]$. Let M_f, M_g, M_{fg} be the supremum of the values of there corresponding function on $[x_{i-1}, x_i]$. Let m_f, m_g, m_{fg} similarly be the same vice versa. We have

$$M_{fg} - m_{fg} \leq (M_f - m_f)(M_g - m_g)$$

Proof. Let $[x, y] \subseteq [x_{i-1}, x_i]$. Observe

$$|fg(x) - fg(y)| = |f(x)(g(x) - g(y))| + |(f(x) - f(y))g(y)|$$

Then the result follows from Theorem 7.1.7. ■

Question 84

Let $\{g_n\}$ be a sequence of non-negative and Stieltjes integrable with respect to an increasing function ϕ on $[a, b]$ such that

$$\lim_{n \rightarrow \infty} \int_a^b g_n d\phi = 0$$

Prove that if f is also integrable, then

$$\lim_{n \rightarrow \infty} \int_a^b fg_n d\phi = 0$$

Proof. Because g_n are non-negative, we know

$$\int_a^b -|f|g_n d\phi \leq \int_a^b fg_n d\phi \leq \int_a^b |f|g_n d\phi$$

Because of Squeeze Theorem, we can now suppose f is non-negative.

For each interval $[x, y] \subseteq [a, b]$, observe

$$\sup_{t \in [x, y]} fg_n(t) \leq \sup_{t \in [x, y]} f(t) \sup_{t \in [x, y]} g_n(t)$$

Let $M_i^f, M_i^{g_n}, M_i^{fg_n}$ has the meaning of those in Rudin, and let $M = \sup_{t \in [a, b]} f(t)$. We have

$$\begin{aligned} U(P, fg_n, \phi) &= \sum_{i=1}^m M_i^{fg_n} \Delta\phi(x_i) \leq \sum_{i=1}^m (M_i^f M_i^{g_n}) \Delta\phi(x_i) \\ &\leq M \sum_{i=1}^m M_i^{g_n} \Delta\phi(x_i) \\ &= MU(P, g_n, \phi) \end{aligned}$$

This give us

$$\overline{\int_a^b fg_n d\phi} \leq M \overline{\int_a^b g_n d\phi}$$

Because

$$\lim_{n \rightarrow \infty} \int_a^b g_n d\phi = 0$$

We now have

$$\limsup_{n \rightarrow \infty} \int_a^b f g_n d\phi = 0$$

Because $\forall n$, $f g_n$ is non-negative and ϕ is increasing, we see

$$\forall n, \int_a^b f g_n d\phi \geq 0$$

This now give us

$$\lim_{n \rightarrow \infty} \int_a^b f g_n d\phi = 0$$

■

Question 85

Let ϕ be a strictly monotone increasing function on $[a, b]$.

- (a) Show that if f is Riemann-Stieltjes integrable with respect to ϕ , then the following quantities all exist and finite for $p > 0$:

$$\|f\|_\infty = \sup\{|f(x)| : x \in [a, b]\}$$

$$\|f\|_p = \left(\int_a^b |f|^p d\phi \right)^{1/p}$$

- (b) Prove that if f is continuous on $[a, b]$, then

$$\lim_{p \rightarrow \infty} \|f\|_p = \|f\|_\infty$$

Proof. Because f is integrable, we know f must be bounded. This tell us $\|f\|_\infty$ do exists as a real number.

Notice that $g, h : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$g(x) = |x| \text{ and } h(x) = x^p$$

are both continuous, then by Theorem 2.4.5, we know

$$h \circ g \text{ is continuous}$$

Then by Theorem 7.1.17, we know

$$h \circ g \circ f \in \mathcal{R}(\phi) \text{ on } [a, b]$$

This tell us

$$\int_a^b |f|^p d\phi \text{ exists}$$

Notice that the integrand $|f|$ is non-negative and ϕ increase, so we know the integral $\int_a^b |f|^p d\phi$ is non-negative. Then we know

$$\|f\|_p = (\int_a^b |f|^p d\phi)^{\frac{1}{p}} \text{ exists}$$

We now prove

$$\lim_{p \rightarrow \infty} \|f\|_p = \|f\|_\infty$$

Let $M := \|f\|_\infty = \sup_{x \in [a, b]} |f(x)|$. We have

$$\int_a^b |f|^p d\phi \leq \int_a^b M^p d\phi = M^p(\phi(b) - \phi(a))$$

This give us

$$\|f\|_p = (\int_a^b |f|^p d\phi)^{\frac{1}{p}} \leq M(\phi(b) - \phi(a))^{\frac{1}{p}}$$

Because $(\phi(b) - \phi(a))^{\frac{1}{p}} \rightarrow 1$ as $p \rightarrow \infty$, we now have

$$\limsup_{p \rightarrow \infty} \|f\|_p \leq M$$

Notice that $|f|$ is continuous and $[a, b]$ is compact. Then we know there must exists $y \in [a, b]$ such that

$$|f(y)| = M$$

Again because $|f|$ is continuous on $[a, b]$, for each $n \in \mathbb{N}$, we know there exists δ_n such that

$$\forall x \in [y - \delta_n, y + \delta_n], |f(x)| > M - \frac{1}{n}$$

Define

$$[c, d] := [y - \delta_n, y + \delta_n] \cap [a, b]$$

We now see

$$\int_c^d (M - \frac{1}{n})^n d\phi \leq \int_c^d |f|^n d\phi \leq \int_a^b |f|^n d\phi$$

Then we have

$$(M - \frac{1}{n})(\phi(d) - \phi(c)) = (\int_c^d (M - \frac{1}{n})^n d\phi)^{\frac{1}{n}} \leq (\int_a^b |f|^n d\phi)^{\frac{1}{n}} = \|f\|_n$$

Notice that

$$(M - \frac{1}{n})(\phi(d) - \phi(c)) \rightarrow M \text{ as } n \rightarrow \infty$$

Then we have

$$\liminf_{n \rightarrow \infty} \|f\|_n \geq M$$

Combining $\limsup_{p \rightarrow \infty} \|f\|_p \leq M$, we now have

$$\lim_{n \rightarrow \infty} \|f\|_n = M = \|f\|_\infty \text{ (done)}$$

■