Introduction to Modern Analysis I

Andre Carneiro, Mark Ma

July 3, 2025

This is MATH S4061 Introduction to Modern Analysis I taught by Andre Carneiro in Summer 2025. The main goal is to cover chapters 1 - 6, while the subgoal is to learn proofs for the first time.

We use textbook Baby Rudin.

Any correction is appreciated!

Mark would like to thank people who provided hand-written notes for reference.

Contents

1 Construction of Real Numbers

The first question to think about: is $0.99\overline{9} = 1$? The difference looks like an infinitesimal. Unfortunately $<$ is not true. If you look at fractions, you multiply

$$0.99\overline{9} = 9 \times 0.11\overline{1} = 9 \times \frac{1}{9} = 1$$

It makes sense, right? We're going to unpack $0.99\overline{9}$ rigorously.

One way to interpret this is to think about rabbit racing a turtle. Assume rabbit starts at 0 m and turtle at 9 m. The speed of rabbit is $S_R = 10$ m/s while turtle is $S_T = 1$ m/s. Once rabbit is at 9 m, turtle is at 9.9 m. Once rabbit at 9.9 m, turtle is at 9.99 m... The Greeks thought this was a paradox as the rabbit never reaches the turtle. For

$$\begin{aligned} x_R &= 0 + 10t \\ x_T &= 9 + 1t \\ \text{align} \qquad x_R &= x_T = 10 \\ \implies t &= 1 \end{aligned}$$

The foundation always come from how one defines the real numbers. Let's begin with the construction of the real number system $\mathbb{R}$.

1.1 Natural Numbers and Integers

Definition 1.1 (Natural Numbers)**.** *Denote the natural numbers as*

$$\mathbb{N} := \{0, 1, 2, 3, \cdots\}$$

The cool property of $\mathbb{N}$ is that there exists a function s (successor) s.t.

$$\begin{aligned} s : \mathbb{N} &\to \mathbb{N} \\ n &\mapsto n + 1 \end{aligned}$$

One can take as an axiom the 'induction property'

Proposition 1.1 (Induction property)**.** *Let $A \subseteq \mathbb{N}$. If $0 \in A$ and for any $n \in A$, $s(n) = n + 1 \in A$. Then necessarily $A = \mathbb{N}$.*

The usefulness of this is that it allows one to do proofs by induction. For logical statement $p(n)$ where $n = 0, 1, 2, \cdots$. To prove $p(n)$ is true for any n, it suffices to prove that

1. $p(0)$ is true.
2. Assume $p(n)$ is true, then prove $p(n + 1)$ is true.

We're essentially using

$$A := \{n \in \mathbb{N} \mid p(n) \text{ is true}\}$$

and applying the Induction property.

Example 1.1 (Gauss)**.**

$$1 + \cdots + n = \frac{(n + 1)n}{2} \qquad \forall\ n \in \mathbb{N}$$

Proof. Note this is statement $p(n)$. We prove by induction.

1. Base Case. We prove $p(1)$. Indeed
$$1 = \frac{(1 + 1)1}{2}$$
2. Inductive Step. We prove $p(n) \implies p(n + 1)$. Assume that for some $n \in \mathbb{N}$, it holds that
$$1 + \cdots + n = \frac{(n + 1)n}{2}$$
We want to use this to prove $p(n + 1)$. We add what's missing to both sides.
$$\begin{aligned} 1 + \cdots + n + (n + 1) &= \frac{(n + 1)n}{2} + (n + 1) \\ &= \frac{n^2 + n + 2n + 2}{2} \\ &= \frac{n^2 + 3n + 2}{2} = \frac{[(n + 1) + 1](n + 1)}{2} \end{aligned}$$
And we're done.

$\square$

But there's no analysis that can be done for $\mathbb{N}$.
One can put two copies of $\mathbb{N}$ and define $\mathbb{Z}$

Definition 1.2 (Integers). *Denote the integers as*

$$\mathbb{Z} := \mathbb{N} \times \{+, -\}/\{+0 \sim -0\}$$

where we declare equivalence $\sim$ at 0 s.t. $+0 = -0$.

1.2 Rationals

We first introduce Equivalence Relation.

Definition 1.3. *$\sim$ is an equivalence relation on a set S if*

1. *Reflexive. $s \sim s$ for any $s \in S$.*
2. *Symmetric. $s \sim t$ implies $t \sim s$ for any $s, t \in S$.*
3. *Transitive. $s \sim t$ and $t \sim r$ implies $s \sim r$ for any $s, t, r \in S$.*

Definition 1.4 (Rationals). *We denote*

$$\mathbb{Q} := \{\frac{p}{q} \mid p, q \in \mathbb{Z}, \quad q \neq 0\}/\sim$$

where we declare equivalence $\sim$ s.t.

$$\frac{p}{q} \sim \frac{r}{s} \qquad \iff ps = rq \in \mathbb{Z}$$

Let's check $\sim$ imposed on $\mathbb{Q}$ is indeed an equivalence relation.

Proof. 1. Reflexive $\frac{p}{q} \sim \frac{p}{q}$.

2. Symmetric $\frac{p}{q} \sim \frac{r}{s}$ implies $\frac{r}{s} \sim \frac{p}{q}$.

3. Transitive $\frac{p}{q} \sim \frac{r}{s}$ and $\frac{r}{s} \sim \frac{m}{n}$ implies $\frac{p}{q} \sim \frac{m}{n}$.

$\square$

Usually we take irreducible fractions as the preferred representative of a class, i.e., we prefer $\frac{1}{2}$ over $\frac{2}{4}$ or $\frac{17}{34}$. This is procedure of set theory, where we partition the set into equivalence classes, and pick representative for each class. Hence $\mathbb{Q}$ is understood as a collection of representatives of equivalence classes with $\sim$ defined above.

Example 1.2. *We define $\mathbb{R}^2/\sim$ where*

$$(p, q) \sim (m, n) \iff p - m \in \mathbb{Z} \qquad q - n \in \mathbb{Z}$$

What are representatives? This is unit cube $[0, 1] \times [0, 1]$ but we glue into a cylinder, and then a torus.

What's good about $\mathbb{Q}$ is that it is an (totally) ordered field. Naively, ordered means there is notion of smaller, equal, bigger. A field is where +, and multiplication works.

Definition 1.5 (Field). *A field is a set $\mathbb{F}$ with two operations*

1. *Addition +*

$$\begin{aligned} + : \mathbb{F} \times \mathbb{F} &\to \mathbb{F} \\ (x, y) &\mapsto x + y \end{aligned} \tag{A1}$$

2. *Multiplication $\cdot$*

$$\begin{aligned} \cdot : \mathbb{F} \times \mathbb{F} &\to \mathbb{F} \\ (x, y) &\mapsto xy \end{aligned} \tag{M1}$$

s.t. for addition

(A2) $x + y = y + x$ *for every* $x, y \in \mathbb{F}$

(A3) $(x + y) + z = x + (y + z)$

(A4) there exists a special element $0 \in \mathbb{F}$ s.t $0 + x = x$

(A5) Given $x \in \mathbb{F}$, there exists an inverse $(-x) \in \mathbb{F}$ s.t $x + (-x) = 0$.

while for multiplication

(M2) $xy = yx$

(M3) $(xy)z = x(yz)$

(M4) there exists a special element $1 \in \mathbb{F}$ s.t. $1x = x$ for any $x \in \mathbb{F}$

(M5) Given $x \neq 0 \in \mathbb{F}$, there exists an inverse $\frac{1}{x} \in \mathbb{F}$ s.t. $x \cdot \frac{1}{x} = 1$.

They also need to satisfy Distributive property

$$x(y+z) = xy + xz \tag{D}$$

Definition 1.6 (Order)**.** *Let S be a set. A (total) order is a relation $<$ in S s.t. two properties hold*

1. *If $x, y \in S$, then exactly one of the following holds*
   - *(a) either $x < y$*
   - *(b) or $x = y$*
   - *(c) or $y < x$.*
2. *There is a transitivity property s.t. if $x, y, z \in S$ with $x < y$ and $y < z$, then $x < z$.*

Proposition 1.2. *$\mathbb{Q}$ is a field and $\mathbb{Q}$ has a total order $<$.*

Definition 1.7 (Ordered Field)**.** $(\mathbb{F}, <)$ *is called an order field if in addition, we have compatibility between the first and second property, i.e.,*

1. *Compatibility with Addition: if $x, y, z \in \mathbb{F}$ and $y < z$, then $x + y < x + z$.*
2. *Compatibility with Multiplication: if $x, y > 0$, then $xy > 0$.*

Proposition 1.3. *$\mathbb{Q}$ is an ordered field.*

However $\mathbb{Q}$ is not good enough for analysis.

Example 1.3. *$\sqrt{2}$ is not rational. One can ask two problems*

Problem 1: $x^2 - 2 = 0$ has no rational solutions, i.e., there does not exist $\frac{p}{q}$ s.t. $(\frac{p}{q})^2 = 2$. (This is usually called $\mathbb{Q}$ is not algebraically complete)

Problem 2: you can't measure hypotenuse of a right angle. (This is called $\mathbb{Q}$ is not metrically complete)

For π, similarly, one cannot measure the circumstance of a circle in $\mathbb{Q}$.

We need to complete $\mathbb{Q}$ by metric. But before that let's prove that $\sqrt{2}$ is not rational.

Proposition 1.4. *There does not exist $\frac{p}{q} \in \mathbb{Q}$ s.t. $(\frac{p}{q})^2 = 2$.*

Proof. Assume by contradiction. Assume there exists an irreducible fraction $\frac{p}{q}$ s.t.

$$(\frac{p}{q})^2 = 2$$

so that $\gcd(p, q) = 1$. Then

$$p^2 = 2q^2$$

thus p^2 is even. We'd love to conclude that p is even. We claim that p is indeed even. By contradiction, assume p is odd, i.e., $p = 2k + 1$ for some $k \in \mathbb{Z}$. Then $p^2 = (2k+1)^2 = 4k^2 + 4k + 1$ is odd, which contradicts our assumption that p^2 is even. Thus p is even, so assume

$$p = 2m$$

for some $m \in \mathbb{Z}$. We obtain that

$$p^2 = 4m^2 = 2q^2$$
$$q^2 = 2m^2$$

so q^2 is even, thus q is even. But both p and q are even, so $\gcd(p, q) \geq 2$, which contradicts our assumption that $\frac{p}{q}$ is irreducible. $\square$

1.3 Real Numbers

We would like to construct an order field that is 'complete'. Let's define a sequence in $\mathbb{Q}$ that should converge to $\sqrt{2}$. Let's start with $p_0 = 1$ (since $1^2 < 2$). To define the sequence recursively, we define

$$p_{n+1} := p_n - \frac{p_n^2 - 2}{p_n + 2} = \frac{2p_n + 2}{p_n + 2}$$

Then one can immediately check the property of the sequence

$$p_{n+1} > p_n$$
$$p_{n+1}^2 < 2$$

We're getting the numbers successively increasing but bounded above by $\sqrt{2}$. But in $\mathbb{Q}$ the limit does not converge.

Definition 1.8 (Upper Bound)**.** *Suppose S is a (totally) ordered set. Let $E \subseteq S$. If there exists $\beta \in S$ s.t.*

$$x \leq \beta \qquad \forall\, x \in E$$

Then we say E is bounded above by β, or that β is an upper bound for E.

Example 1.4. *The sequence $\{p_n^2\}$ is bounded above by* 2.

Now comes the first important definition for reals.

Definition 1.9 (Least Upper Bound)**.** *If β as an upper bound for E has the following additional property holds*

- *If $\alpha < \beta$, then α is NOT an upper bound of E.*

Then we call β the least upper bound of E. We denote

$$\beta \equiv \sup E$$

Similarly, we define the greatest lower bound.

Definition 1.10 (Greatest Lower Bound)**.** *Given an ordered set S. Let $E \subset S$. $\beta \in S$ is a lower bound for the set E if*

$$\beta \leq x \qquad \forall\, x \in E$$

If in addition, for any $\varepsilon > 0$, $\beta + \varepsilon$ is not a lower bound for E, then

$$\beta \equiv \inf E$$

is called the greatest lower bound of E.

Indeed $\sup\{p_n^2\} = 2$. But notice $\{p_n\}$ is bounded above, yet has no supremum in $\mathbb{Q}$. We take

$$E := \{q \in \mathbb{Q} \mid q^2 < 2\}$$

$\sup E$ does not exist in $\mathbb{Q}$. Hence $\mathbb{R}$ consists of adding the supremum of every bounded set in $\mathbb{Q}$, or equivalently, adding every supremum of a bounded monotonically increasing sequence in $\mathbb{Q}$.

What exactly is $\mathbb{R}$? How do we construct $\mathbb{R}$?

1.3.1 Sketch of the Construction of $\mathbb{R}$

One has three ways of doing this.

1. Let's be users of $\mathbb{R}$. Just accept it: There exists $\mathbb{R} \supseteq \mathbb{Q}$ that is an ordered field and s.t. every subset $E \subseteq \mathbb{R}$ bounded above has a supremum in $\mathbb{R}$.

2. In Rudin Chapter 1 Appendix. Define using 'cuts'.

   Definition 1.11. *A cut $\alpha \subseteq \mathbb{Q}$ is that*

   (a) $\alpha \neq \varnothing$ *and* $\alpha \neq \mathbb{Q}$

   (b) If $r \in \alpha$, then $s < r$ implies $s \in \alpha$.

   (c) Has no maximal element, i.e., if $r \in \alpha$, then there exists $s \in \alpha$ s.t. $s > r$.

We define for example

$$A := \{q \in \mathbb{Q} \mid q < 0\} \cup \{q \in \mathbb{Q} \mid q^2 < 2\}$$

This is a cut. We call this cut $\sqrt{2}$. We need to tell one how to square it and get 2. We try to define multiplication of the cuts. We define

$$A^2 := \{q < 2\}$$

Including exactly the irrational cuts gives the real number system. It is a nightmare to check each statement...

3. We do a completion procedure, called a Cauchy Completion. We define

$$\mathbb{R} := \{(q_n) \mid q_n \in \mathbb{Q},\ q_n \text{ is increasing and bounded above}\}$$

We call $\sqrt{2} := (p_n)$. There's a lot of redundancy here. We declare an equivalence relation s.t.

$$(p_n) \sim (r_n) \iff p_n - r_n \to 0 \qquad \text{formally}$$

Theorem 1.1. *There exists $\mathbb{R} \supseteq \mathbb{Q}$ s.t.*

1. *$\mathbb{R}$ is an order field*
2. *$\mathbb{R}$ has the 'Least Upper Bound Property', i.e. of $E \subset \mathbb{R}$ is bounded above, then there exists $\beta \in \mathbb{R}$ s.t. $\beta = \sup E$.*
3. *$\mathbb{Q}$ is dense in $\mathbb{R}$.*

1.3.2 Properties of $\mathbb{R}$

Theorem 1.2 (Archimedean Property)**.** *Let $x, y \in \mathbb{R}$ with $x > 0$. Then there exists a positive integers n s.t. $nx > y$.*

Proof. Define a set

$$A := \{nx \mid n \in \mathbb{N}\}$$

Assume by contradiction that $nx \leq y$ for all $n \in \mathbb{N}$. This is to say A is bounded above by y. Hence using Least Upper Bound Property, there exists $\alpha = \sup A \in \mathbb{R}$. Since $x > 0$, $\alpha - x < \alpha$, hence $\alpha - x$ is no longer an upper bound of A. Thus there exists $m \in \mathbb{N}$ s.t.

$$\alpha - x < mx$$

This implies

$$(m+1)x > \alpha$$

But $(m+1)x \in A$ and $\alpha = \sup A$. Hence we have a contradiction. Conclusion is that A is unbounded. □

Theorem 1.3 (Density Property)**.** *Let $x, y \in \mathbb{R}$ s.t. $x < y$. There exists $q \in \mathbb{Q}$ s.t.*

$$x < q < y$$

Proof. If $x < y$, then $y - x > 0$. We apply Theorem 1.2 to $y - x$. Hence there exists $n \in \mathbb{N}$ positive integer s.t.

$$n(y - x) > 1$$

Also by Theorem 1.2, we find m_1 and m_2 s.t.

$$m_1 > nx, \qquad m_2 > -nx$$

Now we have

$$-m_2 < nx < m_1$$

Hence there exists m s.t. $m - 1 < nx < m$. Thus

$$nx < m < 1 + nx < ny$$
$$x < \frac{m}{n} < y$$

□

2 Countability and Topology of Metric Space

2.1 Countability

In philosophy, one way to classify sets is based on 'how many elements' they contain. The simplest classification is whether a set contains finite or infinite element. However, the tricky thing here is: one can further divide 'infinite' into 'countable' and 'uncountable'.

Before we introduce the formal definition, we need to define a way to characterize 'how many elements' mappings input and output.

Definition 2.1 (Function)**.** *Recall a function $f : A \to B$ is a mapping s.t. for any $a \in A$, there exists a unique $b \in B$ s.t. $f(a) = b$.*

Definition 2.2. *Let $f : A \to B$ be a function. We say that f is*

1. *injective if $f(x) = f(y)$ implies $x = y$, i.e., f maps different elements of A to different elements of B.*
2. *surjective if for any $b \in B$, there exists $a \in A$ s.t. $f(a) = b$, i.e., the image of f is all of B.*
3. *bijective if f is both injective and surjective.*

Remark 2.1. *If $f : A \to B$ is bijective, for any $b \in B$, there exists a unique $a \in A$ s.t. $f(a) = b$, and vice versa. In this case we can define an inverse function*

$$f^{-1} : B \to A \qquad b \mapsto a$$

let us define a relation on set. We say

$$X \sim Y \qquad \text{there exists bijection } f : X \to Y$$

Lemma 2.1. *$\sim$ defined above is indeed an equivalence relation.*

Proof.

1. reflexive. $X \sim X$ since there exists $f : X \to X$ s.t. $f(x) = x$ for any $x \in X$, known as the identity map, which is indeed a bijection.
2. symmetric. If $X \sim Y$, then there exists bijection $f : X \to Y$. Thus we can build an inverse $f^{-1} : Y \to X$ which is also a bijection.
3. transitive. If $X \sim Y$ and $Y \sim Z$, then there exists

$$\begin{aligned} f : X \to Y &\qquad \text{bijection} \\ g : Y \to Z &\qquad \text{bijection} \end{aligned}$$

Hence the composition

$$g \circ f : X \to Z \qquad x \mapsto g(f(x))$$

is also a bijection.

□

Let's define countability. First denote $\mathbb{N}^\star := \mathbb{N} \setminus \{0\}$.

Definition 2.3. *We say a set X is*

1. *finite if $X \sim \{0, 1, 2, \cdots, n\}$ for some $n \in \mathbb{N}$.*
2. *countable if $X \sim \mathbb{N}^*$*
3. *uncountable if otherwise.*

Let's given some examples.

Example 2.1. $\mathbb{N}$ *is countable.*

$$f : \mathbb{N}^\star \to \mathbb{N} \qquad n \mapsto n - 1$$

is indeed a bijection.

Example 2.2. $\mathbb{Z}$ *is countable. Define*

$$f : \mathbb{N}^* \to \mathbb{Z} \qquad f(n) := \begin{cases} \frac{n}{2} & n \text{ even} \\ \frac{-n-1}{2} & n \text{ odd} \end{cases}$$

so that $f(1) = 0,\ f(2) = 1,\ f(3) = -1,\ f(4) = 2,\ f(5) = -2,\ \cdots$

Example 2.3. *We ask whether $\mathbb{Q}$ is countable. Indeed $\mathbb{Z} \times \mathbb{Z} := \{(p,q) \mid p,\, q \in \mathbb{Z}\}$ is countable as indicated in the following proposition.*

Proposition 2.1. *If A, B are countable, then $A \times B$ is countable.*

Proof. Since A and B are countable, we denote

$$A := \{a_i\}_i \qquad B := \{b_j\}_j$$

We rewrite

$$A \times B = \{(a_i, b_j) \mid a_i \in A,\, b_j \in B\} = \bigsqcup_{k=1} \{(a_i, b_j) \mid i + j = k\}$$

□

Corollary 2.1. *Countable unions of finite sets if at most countable.*

Now one can view $\mathbb{Q}$ as subset of $\mathbb{Z} \times \mathbb{Z}$ via

$$h : \mathbb{Q} \to \mathbb{Z} \times \mathbb{Z} \qquad \frac{p}{q} \mapsto (p,q) \qquad \text{where } \frac{p}{q} \text{ is of the irreducible form} \tag{1}$$

Thus

$$\mathbb{Q} \sim h(\mathbb{Q}) = \{(p,q) \in \mathbb{Z} \times \mathbb{Z} \mid \ \frac{p}{q} \text{ is of the irreducible form}\}$$

Hence the question whether $\mathbb{Q}$ is countable is equivalent to ask whether $h(\mathbb{Q})$ is countable. This can be achieved via the following proposition.

Proposition 2.2. *Infinite subset of countable sets are countable.*

Proof. Let our countable set be $A := \{a_i\}_{i \in \mathbb{N}^*}$. Let $E \subseteq A$ be an infinite subset. We define a sequence as follows

1. Let k_1 be the smallest index s.t. $a_{k_1} \in E$
2. Let k_i be the smallest index s.t. $k_i > k_{i-1}$ and $a_{k_i} \in E$, for any $i \geq 2$.

Thus

$$f(i) := a_{k_i}$$

is a bijection from $\mathbb{N}^\star$ to E. □

Corollary 2.2. *$\mathbb{Q}$ is countable.*

Proof. Since $\mathbb{Q}$ can be seen as a subset of $\mathbb{Z} \times \mathbb{Z}$ using (1), and $\mathbb{Q}$ is infinite, thus using Proposition 2.2, $h(\mathbb{Q})$ as an infinite subset of a countable set is itself countable. □

Example 2.4.

$$\mathbb{Q}^n := \{(p_1, \cdots, p_n) \mid p_i \in \mathbb{Q} \quad i = 1, \cdots, n\}$$

is countable from Proposition 2.1.

Now let's give a non-example.

Example 2.5. *Look at*

$$2^{\mathbb{N}} := \{\text{subsets of } \mathbb{N}\} = \{f : \mathbb{N} \to \{0,1\}\}$$

as the set of all sequences with entries zeros and ones.

Proposition 2.3. *$2^{\mathbb{N}}$ is uncountable.*

Proof. We argue using contradiction. Assume $2^{\mathbb{N}}$ is countable and enumerated, say, as

$$\begin{aligned} c_1 &= (c_1^1, c_1^2, c_1^3, \cdots) \\ c_2 &= (c_2^1, c_2^2, c_2^3, \cdots) \\ &\vdots \end{aligned}$$

where $c_i^j \in \{0,1\}$ for any $i,\, j \in \mathbb{N}^*$. Let's define a sequence of zeros and ones as

$$d := (1 - c_i^i)_{i \in \mathbb{N}^*} = (1 - c_1^1, 1 - c_2^2, 1 - c_3^3, \cdots)$$

which is essentially inverting each diagonal entry from $0 \to 1$ and $1 \to 0$. Then d does not lie in such enumeration. But d is indeed some sequence of zeros and ones, hence $2^{\mathbb{N}}$ is uncountable. □

Example 2.6. *$[0,1] \subset \mathbb{R}$ is also uncountable. Indeed*

$$f : 2^{\mathbb{N}} \to [0,1] \qquad c = (c^1, c^2, \cdots) \mapsto 0.c^1c^2 \cdots \qquad \text{as binary expansion}$$

defines a bijection. Since $2^{\mathbb{N}}$ is uncountable, $[0,1]$ is uncountable. Also, since Proposition 2.2 says subsets of countable sets are countable, we know $\mathbb{R}$ is uncountable.

2.2 Basic Topology of Metric Space

We want to rigorously rebuild calculus. Intuitively, limit/continuity needs 'open sets'; derivatives needs an underlying field $\mathbb{R}$; integrals need lengths of intervals $\sum_i f(x_i)\Delta x_i$.

We have now $\mathbb{R}^n := \{(x_1, x_2, \cdots, x_n) \mid x_i \in \mathbb{R}\}$. In fact we can build more structure on $\mathbb{R}^n$.

1. One can view $\mathbb{R}^n$ as vector space so we have vector addition and scalar multiplication.
2. One can define
   (a) inner product $\langle x, y\rangle := \sum_{i=1}^n x_i y_i$
   (b) norm $\|x\| := \sqrt{\langle x, y\rangle}$ as defined by the inner product. Notice from high school
   $$\langle x, y\rangle = \|x\|\,\|y\|\cos\theta \qquad \text{where } \theta \text{ is the angle formed by the vectors } x,\, y \text{ starting at a same origin}$$
   (c) metric $d(x, y) := \|x - y\|$ as defined by the norm. This denotes the distance between x and y.

Hence we study metric space. In particular, we study its topology, which is essentially 'what its open sets look like'.

Definition 2.4 (Metric Space)**.** *A metric space (X, d) is a set with a distance function*

$$d : X \times X \to \mathbb{R} \qquad (x, y) \mapsto d(x, y)$$

with the following properties

1. *$d(x, y) > 0$ for any $x \neq y$; $d(x, x) = 0$ for any $x \in X$.*
2. *$d(x, y) = d(y, x)$ for any $x,\, y \in X$.*
3. *The triangle inequality*
   $$d(x, y) \leq d(x, z) + d(z, y) \qquad \forall\ x,\, y,\, z \in X$$

Example 2.7. *There's multiple ways to define distance on $\mathbb{R}^n$.*

1. *$d_2(x, y) = \|x - y\|$ as before (Eulidean Distance).*
2. *$d_p(x, y) = \|x - y\|_p = ((x_1 - y_1)^p + \cdots + (x_n - y_n)^p)^{\frac{1}{p}}$.*
3. *$d_\infty(x, y) := \max\limits_{i=1,\cdots,n} |x_i - y_i|$.*

Example 2.8. *One can define a metric on a subset of a metric space (E, d) using the same metric. For example on $\mathbb{R}^2$ we define $d(x, y) := \|x - y\|$. Then on a subset*

$$(\{0\} \times [-1, 1]) \bigsqcup \{(x, \sin(\frac{1}{x})) \mid x \neq 0\} \subset \mathbb{R}^2$$

we can inherit the same metric.

Example 2.9. *For $\mathbb{S}^2 \subset \mathbb{R}^3$ where*

$$\mathbb{S}^2 := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$$

we can either define

$$d(x, y) := \|x - y\|$$

as inherited from $\mathbb{R}^3$ or

$$d_{round}(x, y) := \text{shortest arc of a great circle connecting } x \text{ and } y$$

2.2.1 Openness

Now we discuss topology!

Definition 2.5 (Open Ball, Interior Point, Open)**.** *Let (X, d) be a metric space. Let $E \subset X$.*

1. *The open ball of radius r around $p \in X$ is*

$$B_r(p) := \{q \in X \mid d(p, q) < r\}$$

   Remark 2.2. *Notice different metrics give rise to different-looking open balls. For example $(\mathbb{R}^2, d_2)$ gives open disc in $\mathbb{R}^2$, while $(\mathbb{R}, |\cdot|)$ gives open interval on the real line. If we choose $(\mathbb{R}^2, d_\infty)$ where*

$$d_\infty(x, y) := \max\{|x_1 - y_1|, |x_2 - y_2|\}$$

   then the open ball is open rectangle. If we choose $(\mathbb{R}^2, d_1)$ where

$$d_1(x, y) := |x_1 - y_1| + |x_2 - y_2|$$

   then the open ball is open diamond.

2. *$p \in E$ is called an interior point of E if $B_r(p) \subset E$ for some $r > 0$.*

3. *$E \subset X$ is open if all points of E are interior points.*

Remark 2.3. *Being open is a relative notion. The question open or not is what we ask of 'subsets'!*

1. *Consider an open interval on $\mathbb{R}$. It is indeed open in $(\mathbb{R}, |\cdot|)$. However it is not open in $(\mathbb{R}^2, d_2)$ when we identify $\mathbb{R}$ with $\mathbb{R} \times \{0\}$.*

2. *If we just ask if some set is open, then this question is itself not well-posed. One always needs to think of the underlying metric space. In the above example, one can interpret as either the interval in $(X, d) = (\mathbb{R}, |\cdot|)$ or in $(\mathbb{R}^2, d_2)$. This is essentially because the metrics in the two spaces are different, hence the definition for open balls differ.*

Proposition 2.4. *Let (X, d) be a metric space. $B_r(p)$ open ball is open for any $p \in X$ and any $r > 0$.*

Proof. We want to show that any point $q \in B_r(p)$ is an interior point, i.e., for each $q \in B_r(p)$, we want to find some $r_q > 0$ radius s.t. $B_{r_q}(q) \subset B_r(p)$. Indeed we define

$$r_q := r - d(p, q)$$

We claim that $B_{r_q}(q) \subset B_r(p)$. To show the claim, for any $x \in B_{r_q}(q)$, by definition of open ball we have

$$d(x, q) < r_q = r - d(p, q)$$

But then using definition of a metric we see

$$d(x, p) \le d(x, q) + d(q, p) < r$$

Thus $x \in B_r(p)$. But this works for any $x \in B_{r_q}(q)$, hence $B_{r_q}(q) \subset B_r(p)$. $\square$

Remark 2.4. *Following the above proof in Proposition 2.4, one has a recipe to prove if $E \subseteq (X, d)$ is open. The procedure is as follows*

1. *Take any $q \in E$*

2. *Construct a radius $r_q > 0$. One can always draw a picture to inspire what r_q might be.*

3. *Check that $B_{r_q}(q) \subseteq E$. How do we prove containment of sets? Take any point $x \in B_{r_q}(q)$ and show that $x \in E$.*

And we're done.

Let's prove some results for openness.

Theorem 2.1 ('Openness' closed under arbitrary union and finite intersection)**.** *Given (X, d) metric space. Let $E_\alpha \subseteq X$ be open subsets of (X, d) for every α. Then*

1. *$\bigcup_\alpha E_\alpha \subseteq X$ is open in (X, d). Notice this union may contain infinite, even uncountable terms.*

2. *Any finite intersection $E_{\alpha_1} \cap \cdots \cap E_{\alpha_n}$ is open in (X, d). Notice here intersection has to be finite.*

Proof. 1. We prove $\bigcup_\alpha E_\alpha \subseteq X$ is open. Take any $q \in \bigcup_\alpha E_\alpha$. What does it mean for q to belong to this union? This means

$$q \in E_{\alpha_0} \qquad \text{for some } \alpha_0$$

Now since E_{α_0} is an open subset, there exists $r_q > 0$ s.t. $B_{r_q}(q) \subseteq E_{\alpha_0}$. One needs to come up now a radius for which the ball is contained in $\bigcup_\alpha E_\alpha$. But this is trivial by taking the same r_q since

$$B_{r_q}(q) \subseteq E_{\alpha_0} \subseteq \bigcup_\alpha E_\alpha$$

Conclude using arbitrariness of q in $\bigcup_\alpha E_\alpha$.

2. Take a point $q \in E_{\alpha_1} \cap \cdots \cap E_{\alpha_n}$. What does it mean to be in the intersection? This means

$$q \in E_{\alpha_i} \qquad \forall\ i = 1, \cdots, n$$

Now since each E_{α_i} is open, there exists a radius $r_i > 0$ for $i = 1, \cdots, n$ s.t.

$$B_{r_i}(q) \subseteq E_{\alpha_i} \qquad \forall\ i = 1, \cdots, n$$

Here we're applying the 'openness' n times. Now what is the magical radius that works to fit the ball in $E_{\alpha_1} \cap \cdots \cap E_{\alpha_n}$? We take the minimum!

$$r := \min\{r_i \mid i = 1, \cdots, n\}$$

This is the place where finiteness comes in handy! If we're not looking at finitely many numbers then we couldn't take minimum. If there's infinitely many, the infimum could be 0, which is a risk. Now

$$B_r(q) \subseteq B_{r_i}(q) \subseteq E_{\alpha_i} \qquad \forall\ i = 1, \cdots, n$$

Hence

$$B_r(q) \subseteq E_{\alpha_1} \cap \cdots \cap E_{\alpha_n}$$

Conclude using arbitrariness of q in $E_{\alpha_1} \cap \cdots \cap E_{\alpha_n}$.

$\square$

Remark 2.5. *Is $\varnothing \subseteq (X, d)$ open? This is true because there's nothing to check (something related to logic). Or one can check the definition of a topological space, where $\varnothing$ is always contained in the topology.*

2.2.2 Closedness

Definition 2.6 (Limit Point; Isolated Point; Closed)**.** *Let (X, d) be metric space. Let $E \subseteq X$.*

1. *$p \in X$ is a limit point of E if every open ball around p contains a point $q \in E$ with $q \neq p$. (Notice p need not be in E). Or to use formal definition*

$$p \in X \quad \textit{is a limit point of } E \textit{ if} \qquad \forall\ r > 0,\ \exists\ q \in B_r(p) \cap E,\ \textit{s.t. } p \neq q$$

2. *If $p \in E$ is not a limit point of E, then we call it an isolated point of E.*

3. *We say $E \subseteq X$ is closed if every limit point of E belongs to E.*

Example 2.10. *Let $(X, d) = (\mathbb{R}, |\cdot|)$. Consider $E = \mathbb{Z} \subseteq \mathbb{R}$. Then for any integer $p \in E$,*

$$B_{\frac{1}{2}}(p) \cap E = \{p\}$$

Which means any point in E is not a limit point of E. Hence each point in E is an isolated point of E. In this case $\mathbb{Z} \supseteq \varnothing = \{$all limit points of $\mathbb{Z}\}$ hence $\mathbb{Z}$ is closed in $\mathbb{R}$.

Example 2.11. *Let*

$$E = \{\frac{1}{n} \mid n > 0,\, n \in \mathbb{N}\} \subseteq \mathbb{R} = (X, d)$$

This looks like a sequence converging to 0 (though we haven't defined yet). We would like 0 alone to be the only limit point of E. This is how the definition of limit point (contain another point $q \neq p$) is helping us!

Now we remark on closedness.

Remark 2.6. *To show a set E is not closed, one can argue using there exists $p \notin E$ s.t. p is a limit point of E.*

Remark 2.7. *WARNING: Closedness is also relative. One can construct $E \subseteq (X,d) \subseteq (Y,d)$ s.t. E is closed in X but E is not closed in Y.*

Remark 2.8. *WARNING: a subset can be both open and closed in a given metric space (X,d), and can also be neither open nor closed.*

Example 2.12 (Example for Remark 2.8)**.** *In $(\mathbb{R}, |\cdot|)$, intervals $[0,1)$ and $(0,1]$ are neither open nor closed.*

Example 2.13. *In $(\mathbb{R}, |\cdot|)$, $\varnothing$ and $\mathbb{R}$ are both open and closed (which we call clopen). In fact for $(\mathbb{R}, |\cdot|)$ they're the only clopen sets.*

Remark 2.9. *For any (X,d), X and $\varnothing$ are clopen. In fact one can construct (X,d) where there're more clopen subsets.*

2.2.3 Properties of Open and Closed

One has powerful duality operation that relates an open and closed set, which is set complement.

Proposition 2.5 (Duality)**.** *Let (X,d) be a metric space. $E \subseteq X$ is open iff $E^c \subseteq X$ is closed, where*

$$E^c \equiv X \setminus E := \{x \in X \mid x \notin E\} \text{ is the complement of } E \text{ in } X$$

Proof.

1. ($\Longrightarrow$) Given E open, we want to show E^c is closed. Take $p \in X$ a limit point of E^c, we want to show that $p \in E^c$. How do we do this? If $p \in X$ is a limit point of E^c, then for any $r > 0$, there exists $q \in B_r(p) \cap E^c$ s.t. $q \neq p$.
   (a) If $p \in E$, then p is not an interior point of E, because $B_r(p) \not\subseteq E$ due to the existence of q. This is contradiction since E is open by assumption.
   (b) If $p \notin E$ then $p \in E^c$ and we're happy.
2. ($\Longleftarrow$) Given E^c closed, we show E is open. Take any $p \in E$, we show that p is an interior point of E. Since $p \notin E^c$, and E^c is closed, then p is not a limit point of E^c. To violate the limit point definition, tehre exists $r > 0$ s.t.
   $$B_r(p) \cap E^c = \{p\} \text{ or } \varnothing$$
   But $\{p\}$ is not possible since $p \notin E^c$. Hence $B_r(p) \cap E^c = \varnothing$. But this means $B_r(p) \subseteq E$.

□

We introduce a lemma on set operations.

Lemma 2.2.

$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha E_\alpha^c$$
$$\left(\bigcap_\alpha E_\alpha\right)^c = \bigcup_\alpha E_\alpha^c$$

Proof. We only prove the first one. Let $x \in \left(\bigcup_\alpha E_\alpha\right)^c$ then $x \notin \bigcup_\alpha E_\alpha$, hence

$$x \notin E_\alpha \qquad \forall\ \alpha$$

Thus $x \in E_\alpha^c$ for any α. □

Corollary 2.3. *Let (X,d) be metric space. $F_\alpha \subseteq X$ closed for any α. Then*

1. *$\bigcap_\alpha F_\alpha$ is closed in X.*
2. *$F_{\alpha_1} \cup \cdots \cup F_{\alpha_n}$ is closed in X.*

Proof. Define $E_\alpha = F_\alpha^c$. Then using Proposition 2.5, E_α is open. Apply Lemma 2.2, then conclude using Theorem 2.1. More precisely

$$\left(\bigcap_\alpha F_\alpha\right)^c = \bigcup_\alpha F_\alpha^c = \bigcup_\alpha E_\alpha \qquad \text{open from Theorem 2.1}$$

hence $\bigcap_\alpha F_\alpha$ is closed by Proposition 2.5. □

Now we discuss subspace topology.

Proposition 2.6. *Let $(Y, d) \subseteq (X, d)$ with the same metric. E is open in Y iff*

$$E = Y \cap G \qquad \text{for some } G \text{ open in } X$$

Proof.

1. ($\implies$) E open in Y, so for any $p \in E$, there is some $r_p > 0$ s.t. $B^Y_{r_p}(p) \subseteq E$ where

$$B^Y_{r_p}(p) := \{y \in Y \mid d(y, p) \leq r_p\}$$

The question is : What is G? The way is to take the X ball

$$B^X_{r_p}(p) := \{x \in X \mid d(x, p) \leq r_p\}$$

and we define

$$G := \bigcup_{p \in E} B^X_{r_p}(p)$$

G is open using Theorem 2.1. Why is $E = Y \cap G$?

(a) If we take a point $q \in E$, then indeed $q \in Y$ because E is a subset of Y. Also $q \in G$ because $q \in B^X_{r_q}(q) \subseteq G$. Hence $q \in Y \cap G$.

(b) Same backwards. If $q \in Y \cap G$, then $q \in G$, i.e., there exists $p \in E$ s.t.

$$q \in B^X_{r_p}(p) \cap Y = B^Y_{r_p}(p) \subseteq E$$

2. ($\impliedby$) Assume $E = Y \cap G$ for some G open in X. Then let $p \in E$, $p \in Y$ and $p \in G$. Since G is open in X, there exists $r_p > 0$ s.t.

$$B^X_{r_p}(p) \subseteq G$$

Now we can interset both sides with Y and so

$$B^X_{r_p}(p) \cap Y = B^Y_{r_p}(p) \subseteq G \cap Y = E$$

Thus there exists open ball (in Y) around p contained in E.

□

We also look at an extreme example.

Example 2.14. *Look at (X, d) with*

$$d(x, y) := \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$

Then all singletons $\{p\} \subset X$ are open, then any subset $A \subseteq X$ is open. But also any subset $B \subseteq X$ is closed because B^c is always open in X. Thus everything in this metric space is clopen.

We have some insightful property of limit point.

Proposition 2.7. *Consider (X, d) where $E \subseteq X$. If p is a limit point of E, then every open ball around p contains infintely-many elements of E.*

Proof. Assume that there exists open ball $B_r(p)$ s.t.

$$B_r(p) \cap E = \{x_1, \cdots, x_n, p\} \qquad \text{is finite}$$

Then look at

$$r_i := d(p, x_i)$$

We take a new radius

$$p < \min\{r_i \mid i = 1, \cdots, n\}$$

Then look at

$$B_\rho(p) \cap E = \{p\} \text{ or } \varnothing$$

Thus p is not a limit point of E. □

2.3 Compactness

In calculus, one might have heard that a compact set is closed and bounded. This is OK for $\mathbb{R}^n$. The tragedy is merely closed and bounded, in general, does not imply what we like to be compact.

Consider rectangle in $\mathbb{R}^2$. Look at a sequence of points in the rectangle. We want some limit of the sequence. We divide up the square. Then there is some limit point if we keep dividing, up to certain choices of square.

Also, compactness itself is just beautiful :)

Definition 2.7 (Compactness)**.** *Let (X, d) be a metric space. $K \subseteq X$.*

1. *Let $G_\alpha \subseteq X$ be open subsets. We say that $\{G_\alpha\}$ is an open cover of K if*

$$K \subseteq \bigcup_\alpha G_\alpha$$

2. *We say K is compact if* ***every*** *open cover $\{G_\alpha\}$ of K admits a* ***finite*** *subcover, i.e., there exists $\alpha_1, \cdots, \alpha_n$ s.t.*

$$K \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$$

We first look at example of non-compactness.

Example 2.15. *It suffices to show there is some open cover that cannot have any finite subcover. Let*

$$E = (-1, 1)$$

We show this is not compact. Take

$$G_n := (-1 + \frac{1}{n}, 1 - \frac{1}{n})$$

Then

$$E \subseteq \bigcup_{n=1}^{\infty} G_n = (-1, 1)$$

Hence if we stop at finite union, we do not cover E. Thus it is not compact.

Remark 2.10. *Compactness is an absolute notion, i.e., it does not depend on the ambient metric space (X, d). Provisionally we call K compact in X and prove it is actually not the case.*

Proposition 2.8 (Compact is Absolute)**.** *Let $(Y, d) \subseteq (X, d)$ be metric space, and $K \subseteq Y \subseteq X$. Then K is compact in Y iff K is compact in X.*

Proof. We unravel the definition. The key is 'every' cover.

1. ($\Longrightarrow$) We start by taking any X-open cover $\{G_\alpha\}$ of K. Then taking intersections we define

$$E_\alpha := G_\alpha \cap Y \qquad \forall\ \alpha$$

Now E_α is Y-open for every α. Because

$$K \subseteq \bigcup_\alpha G_\alpha$$

we have

$$K \subseteq \bigcup_\alpha E_\alpha$$

Now we've produced a Y-open cover $\{E_\alpha\}$ of K. Using Y-compactness, there exists a subcover $\{E_{\alpha_i}\}_{i=1}^n$ s.t.

$$K \subseteq \bigcup_{i=1}^{n} E_{\alpha_i}$$

But then G_{α_i} are bigger. Hence

$$K \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$$

and this is a X-finite subcover. Thus K is X-compact.

2. ($\Longleftarrow$) We assume K is X-compact. Let $\{E_\alpha\}$ be any Y-open cover of K. By Proposition 2.6, each

$$E_\alpha = G_\alpha \cap Y$$

where G_α is X-open. Now note that G_α covers K, since $G_\alpha \supseteq E_\alpha$ for any α. Now we encountered a X-open cover of K. By X-compactness, there exists $\alpha_1, \cdots, \alpha_n$ s.t.

$$K \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$$

But since $K \subseteq Y$, one has

$$K \subseteq \bigcup_{i=1}^{n} E_{\alpha_i}$$

and this is Y-finite subcover. Thus K is Y-compact.

□

Remark 2.11 (Compact Metric Space)**.** *Equivalently, we could've started with a metric space (K, d), then say (K, d) itself as metric space is compact if every open cover $\{G_\alpha\}$*

$$K = \bigcup_{\alpha} G_\alpha$$

admits a finite subcover $\{G_{\alpha_i}\}_{i=1}^{n}$

$$K = \bigcup_{i=1}^{n} G_{\alpha_i}$$

Now these open cover are K-open.

2.3.1 Compactness and Closedness

We start by proving some simple proposition that relates compactness with closedness.

Proposition 2.9 (Compact implies Closed)**.** *Let (X, d) be metric space. If $K \subseteq X$ is compact, then K is closed in X.*

Proof. Let $p \notin K$. Then for every $q \in K$, let $r_q = \frac{1}{2} d(p, q)$. Note that

$$B_{r_q}(q) \cap B_{r_q}(p) = \varnothing \tag{2}$$

But the balls $\{B_{r_q}(q)\}_{q \in K}$ is indeed an open cover of K. By compactness, there exists $q_1, \cdots, q_n$ s.t.

$$K \subseteq \bigcup_{i=1}^{n} B_{r_{q_i}}(q_i)$$

We can now work with finitely-many balls! This is convenient since then we take

$$r := \min\{r_{q_i} \mid i = 1, \cdots, n\}$$

Then

$$B_r(p) \cap K = \varnothing$$

using (2). Hence p is an interior point of K^c. Thus K^c is open. This means K is closed in X via Proposition 2.5. □

Proposition 2.10 (Closed subsets of Compact sets are compact)**.** *If (K, d) is a compact metric space and $F \subseteq K$ is closed. Then F is also compact.*

Proof. For any $\{V_\alpha\}$ open cover of F, one wish to extract a finite subcover $\{V_{\alpha_i}\}_{i=1}^{n}$ of F. But how do we use compactness of K? The trick is to observe

$$F \cup F^c = K$$

Since F is closed, F^c is open, hence $\{V_\alpha\} \cup \{F^c\}$ is an open cover of K. By compactness of K there exists α_i for $i = 1, \cdots, n$ s.t.

$$K \subseteq V_{\alpha_1} \cup \cdots \cup V_{\alpha_n} \cup F^c$$

But $F \subseteq K$ by assumption, and notice $F \cap F^c = \varnothing$, so

$$F \subseteq \bigcup_{i=1}^{n} V_{\alpha_i}$$

is a finite open cover of F. □

2.3.2 Finite Intersection Property

One has powerful proposition that says nested compact sets has non-empty intersection.

Proposition 2.11 (Finite Intersection Property). *Let (X,d) be metric space. Let $\varnothing \neq K_\alpha \subseteq X$ be compact for any α. If every finite intersection is nonempty, i.e.*

$$K_{\alpha_1} \cap \cdots \cap K_{\alpha_n} \neq \varnothing \qquad \forall\ \alpha_1, \cdots, \alpha_n$$

Then

$$\bigcap_\alpha K_\alpha \neq \varnothing$$

Proof. This is quite clever argument. Assume for contradiction that

$$\bigcap_\alpha K_\alpha = \varnothing$$

Then indeed there exists some compact set in the collection, which we call K_1 s.t.

$$K_1 \cap \left(\bigcap_{\alpha,\, \alpha \neq 1} K_\alpha \right) = \bigcap_\alpha K_\alpha = \varnothing$$

In other words, no points in K_1 belong to all other K_α. Using Proposition 2.9 we know K_α are closed, hence K_α^c are open, and notice

$$K_1 \subseteq \left(\bigcap_{\alpha,\, \alpha \neq 1} K_\alpha \right)^c = \bigcup_{\alpha,\, \alpha \neq 1} K_\alpha^c$$

indeed forms an open cover. Using K_1 is compact, there exists finitely many $\alpha_1, \cdots, \alpha_n$ s.t.

$$K_1 \subseteq K_{\alpha_1}^c \cup \cdots \cup K_{\alpha_n}^c = \left(\bigcap_{i=1}^n K_{\alpha_i} \right)^c$$

But now

$$K_1 \cap \left(\bigcap_{i=1}^n K_{\alpha_i} \right) = \varnothing$$

This contradicts our assumption that finitely many intersections is always non-empty. □

Corollary 2.4 (Nested Interval Theorem (Generalized)). *If K_n is a nested collection of non-empty compact metric spaces, i.e.,*

$$K_n \supseteq K_{n+1} \supseteq \cdots$$

for any n. Then

$$\bigcap_n K_n \neq \varnothing$$

Proof. We need to verify why one can apply Finite Intersection Property 2.11. But this is done by noticing

$$K_1 \cap K_2 \cap \cdots \cap K_n = K_n \neq \varnothing \qquad \forall\ n$$

□

Case of $\mathbb{R}$ In particular we look at Finite Intersection Property on $\mathbb{R}$.

Proposition 2.12 (Nested Interval Theorem). *If $\{I_n\}$ is a nested sequence of closed and non-empty intervals of $\mathbb{R}$, i.e., $I_n \supseteq I_{n+1}$ for any n, then*

$$\bigcap_n I_n \neq \varnothing$$

Direct Proof without using Corollary 2.4. Consider $I_n = [a_n, b_n]$ with $a_n \leq b_n$ for any n. Due to the nested property, one has

$$a_n \leq a_\ell \leq b_\ell \leq b_n \qquad \forall\ \ell \geq n$$

Since all $\{a_n\}$ are bounded above by b_1, $x := \sup\{a_n\}$ exists in $\mathbb{R}$ by the Least Upper Bound Property. Also, since all b_n are upper bounds for the set $\{a_n\}$, by definition of x as the supremum (least upper bound), one obtain

$$x \leq b_n \qquad \forall\ n$$

Hence

$$a_n \leq x \leq b_n \qquad \forall\ n \implies x \in \bigcap_n I_n$$

□

Case of $\mathbb{R}^k$ We look at Finite Intersection Property on $\mathbb{R}^k$ as generalization.

Definition 2.8. *A k-cell is a subset of $\mathbb{R}^k$ of the type*

$$I := \{(x_1, \cdots, x_k) \in \mathbb{R}^k \mid a_i \leq x_i \leq b_i,\ i = 1, \cdots, k\}$$

for given

$$a_i \leq b_i \qquad i = 1, \cdots, k$$

In other words

$$I = [a_1, b_1] \times \cdots \times [a_k, b_k] = \prod_{k=1}^{n} [a_i, b_i]$$

Proposition 2.13 (Nested Interval Theorem ($\mathbb{R}^k$))**.** *If I_n is a nested sequence of non-empty k-cells, then*

$$\bigcap_n I_n \neq \varnothing$$

Proof. Write

$$I_n := [a_1^{(n)}, b_1^{(n)}] \times \cdots \times [a_k^{(n)}, b_k^{(n)}]$$

Then applying Proposition 2.12 to each coordinate we obtain existence of some

$$x_i^* := \sup\{a_i^{(n)}\} \in [a_i^{(n)}, b_i^{(n)}] \qquad \forall\ i = 1, \cdots, k$$

Define and notice

$$x^* := (x_1^*, \cdots, x_k^*) \in \bigcap_n I_n$$

□

To see Proposition 2.13 is indeed a special case of Corollary 2.4, we need to show that the k-cells are compact.

Lemma 2.3. *Every k-cell is compact.*

Proof. Take

$$I = [a_1, b_1] \times \cdots \times [a_k, b_k] \subset \mathbb{R}^k$$

and

$$\delta \equiv \|a - b\| := \sqrt{\sum_{i=1}^{k} (a_i - b_i)^2}$$

Then for any $x, y \in I$, $\|x - y\| < \delta$. Assume for contradiction that I is not compact, i.e., there exists an open cover $\{G_\alpha\}$ of I s.t. not finite subcover exists. But given I, one can bisect it into 2^k sub k-cells by defining

$$c_i := \frac{a_i + b_i}{2}$$

and constructing the new generation of k-cells using $[a_i, c_i]$, $[c_i, b_i]$ and so on, so that

$$\|x - y\| \leq \frac{\delta}{2} \qquad \text{in each sub } k\text{-cell}$$

Now there exists at least one sub k-cell, which we denote I_1 s.t. cannot be covered by finitely many $\{G_\alpha\}$. One do this procedure inductively to obtain a sequence k-cells $\{I_n\} \subseteq \mathbb{R}^k$ s.t.

1. $I \supseteq I_1 \supseteq I_2 \supseteq \cdots$
2. I_n cannot be covered by finitely many G_α for any n
3. For any $x, y \in I_n$, necessarily $\|x - y\| \leq 2^{-n}\delta$.

Now using Proposition 2.13, there exists $x^* \in \bigcap_n I_n$. Since

$$x^* \in \bigcap_n I_n \subseteq I \subseteq \bigcup_\alpha G_\alpha$$

there exists some α s.t. $x^* \in G_\alpha$. Using G_α is open, there exists some enough radius $r > 0$ s.t.

$$B_r(x^*) \subseteq G_\alpha$$

Now leveraging Theorem 1.2, one may choose N sufficiently large so that

$$2^{-N}\delta < r \implies I_N \subseteq B_r(x^*) \subseteq G_\alpha$$

Now I_N belongs to our sequence of k-cells but $\{G_\alpha\}$ is a finite subcover of I_N. This is contradiction. □

2.3.3 Sequential Compactness

One may naturally ask: What about Finite Intersection Property (Nested Interval Property) in ∞-dimensional spaces? In this case the k-cells are intervals around sequences, and non-empty intersection of nested intervals is equivalent to saying: there exists some limit of this sequence in the compact metric space.

Theorem 2.2 (Compactness implies Sequential Compactness)**.** *Let (K, d) be a compact metric space, and $E \subseteq K$ be an infinite set (with infinite elements). Then E has a limit point in K (in other words, for any sequence $\{x_n\} \subseteq E$, there exists $x \in K$ s.t. x is a limit point of $\{x_n\}$.*

Proof. Assume that every point in K is not a limit point of E, i.e., for any $q \in K$, there exists $r_q > 0$ s.t.

$$B_{r_q}(q) \cap E = \{q\} \text{ or } \varnothing$$

Thus $B_{r_q}(q)$ intersects E with at most one element for each $q \in K$. But $\{B_{r_q}(q)\}_{q\in K}$ is indeed open cover for K, hence using K compact, there exists finite subcover

$$K \subseteq \bigcup_{i=1}^{N} B_{r_{q_i}}(q_i)$$

But then intersecting with E yields

$$E \subseteq \bigcup_{i=1}^{N} B_{r_{q_i}}(q_i) \cap E \subseteq \{q_i \mid i = 1, \cdots, N\}$$

We deduce that E is subset of a finite set, contradiction to E being infinite set. $\square$

2.3.4 Heine-Borel

For compact sets $E \subseteq \mathbb{R}^k$ in finite dimensional spaces,we discuss its equivalence conditions.

Definition 2.9 (Bounded Set)**.** *Given metric space (X, d), a set $E \subseteq X$ is bounded if there exists $p \in X$ and $r > 0$ s.t.*

$$E \subseteq B_r(p)$$

Theorem 2.3 (Heine-Borel)**.** *If $E \subseteq \mathbb{R}^k$, the following are equivalent*

(a) E is closed and bounded.

(b) E is compact.

(c) Every infinite subset of E has a limit point in E.

Proof. We show equivalence in circular.

1. $(a) \implies (b)$. Using E is bounded, one can find a k-cell $I \subseteq \mathbb{R}^k$ s.t. $E \subseteq I$. Since I is compact by lemma 2.3, and since E is closed by assumption, using Proposition 2.10 we know that E is compact.

2. $(b) \implies (c)$. This is precisely Theorem 2.2.

3. $(c) \implies (a)$. By contradiction, assume E is either not closed or not bounded.

   (i) If E is not bounded, then one may choose a sequence $\{x_n\} \subseteq E$ s.t. $\|x_n\| > n$. Notice now $\{x_n\}$ is an infinite subset of E, and we claim it has no limit point in E. If so, denote the limit point as $y \in \mathbb{R}^k$, then consider the ball $B_1(y)$. Since y is limit point of $\{x_n\}$, using Proposition 2.7, for any ball around y, it necessarily intersects infinitely-many elements of $\{x_n\}$. But if $x_n \in B_1(y)$, then

   $$\|x_n\| \le \|x_n - y\| + \|y\| \le 1 + \|y\|$$

   Now for n sufficiently large, there exists some $n+1$ that violates the above due to $\|x_{n+1}\| > n + 1$. Thus E must be bounded.

   (ii) If E is not closed, then there exists $q \notin E$ s.t. q is a limit point of E. Using definition of a limit point, one can construct a sequence $\{x_n\} \subseteq E$ s.t.

   $$d(x_n, q) < \frac{1}{n} \qquad \forall\, n$$

Notice now $\{x_n\}$ is an infinite subset of E, and we claim it has no limit point in E. Indeed, for any $y \in E$ (so that $y \neq q$). Using triangle-inequality one has

$$\begin{aligned} \|x_n - q\| + \|x_n - y\| &\geq \|q - y\| \\ \|x_n - y\| &\geq \|q - y\| - \|x_n - q\| \\ &> \|q - y\| - \frac{1}{n} \end{aligned}$$

By Theorem 1.2 there exists $N \in \mathbb{N}$ sufficiently large s.t.

$$\|x_n - y\| \geq \frac{1}{2}\|y - q\| \qquad \forall\, n \geq N$$

Thus y cannot be limit point of $\{x_n\}$. But $y \in E$ is arbitrary, so $\{x_n\}$ has no limit point in E.

□

Corollary 2.5. *Every bounded infinite subset of $\mathbb{R}^k$ has a limit point.*

Proof. Let E be a bounded infinite subset of $\mathbb{R}^k$. Using E is bounded, one can fit it inside a k-cell $I \subseteq \mathbb{R}^k$. Using I is compact Lemma 2.3, one obtain from Theorem 2.3 that any infinite subset of I has a limit point in I, in particular, E does. □

2.4 Other topological concepts

We discuss another notion in topology known as connectedness. Before that we introduce certain notations

Definition 2.10 (Closure)**.** *Let (X, d) be a metric space. $A \subseteq X$. We denote*

$$\overline{A} := A \cup \{x \in X \mid x \text{ is a limit point of } A\}$$

as the closure of A in X.

Definition 2.11 (Connected)**.** *Let (X, d) be a metric space and $E \subseteq X$.*

1. *$A \subseteq X$ and $B \subseteq X$ are separated if both*

$$\overline{A} \cap B = \varnothing \qquad \text{and} \qquad A \cap \overline{B} = \varnothing$$

2. *$E \subseteq X$ is connected if it cannot be written as*

$$E = A \cup B$$

with A and B non-empty, and separated.

Example 2.16. *Connected subsets of $\mathbb{R}$ are intervals.*

We also define 'dense'. In analysis this is particular useful to simplify 'proofing for a statement in X' to 'proofing for a statement in a dense subset'.

Definition 2.12 (Dense)**.** *Let (X, d) be metric space. $E \subseteq X$ is dense in X if $\overline{E} = X$, i.e., every point of X is in E or is a limit point of E.*

Example 2.17. *$\mathbb{Q}$ is dense in $\mathbb{R}$. Indeed, for any $x \in \mathbb{R} \setminus \mathbb{Q}$, the ball*

$$B_r(x) = (x - r, x + r)$$

always contains some rational inside by Density Property Theorem 1.3.

We define a 'perfect set'.

Definition 2.13. *E is perfect if E is closed and every point of E is a limit point of E.*

3 Limits and Convergence of Sequences (in metric space)

We study rigorously what we mean by a limit of a sequence. Instead of $\mathbb{R}$, one can also work with definition using metric space.

Definition 3.1 (Sequence)**.** *Given metric space (X, d). A sequence is a function*

$$f : \mathbb{N} \to X \qquad n \mapsto f(n) \equiv x_n$$

3.1 Sequence Convergence

Definition 3.2 (Sequence Convergence in metric space)**.** *A sequence $\{p_n\}$ converges in (X, d) if there exists $p \in X$ (called the limit of the sequence) s.t. for any $\varepsilon > 0$, there exists a large enough $N = N(\varepsilon) \in \mathbb{N}$ s.t. for any $n \geq N(\varepsilon)$,*

$$d(p_n, p) < \varepsilon$$

In this case we say

$$p_n \to p \qquad or \qquad \lim_{n\to\infty} p_n = p$$

This is to say the terms p_n for $n \geq N$ enters the ball $B_\varepsilon(p)$ and never leaves. The name of the game is to trap the whole tail of the sequence inside the ball.

There is general procedure to prove convergence of a sequence.

1. First, one can guess (by observation) what the limit of $\{p_n\}$ should be. Call it p.
2. Let $\varepsilon > 0$ be arbitrary.
3. Construct $N(\varepsilon)$ s.t. $p_n \in B_\varepsilon(p)$ for any $n \geq N(\varepsilon)$.

Example 3.1. *Consider the sequence*

$$p_0 = 1 \qquad p_{n+1} := \frac{2p_n + 2}{p_n + 2}$$

In $(\mathbb{R}, |\cdot|)$, p_n converges to $\sqrt{2}$, while in $\mathbb{Q}$, p_n does not converge.

One has several consequences of convergence.

Theorem 3.1 (Properties of Convergence)**.** *Let $\{p_n\}$ be a sequence in (X, d).*

(a) $p_n \to p$ iff every open set U that contains p contains all but finitely many terms of p_n.

(b) If $p_n \to p$ and $p_n \to p'$, then $p = p'$, i.e., limits are unqiue in metric spaces.

(c) If $\{p_n\}$ converges, then $\{p_n\}$ is bounded.

(d) If $E \subseteq X$ and p is a limit point of E, then there is a sequence $\{p_n\}$ with $p_n \neq p$, $p_n \in E$ and $p_n \to p$.

Proof. (a) (i) ($\Longrightarrow$) Let $p_n \to p$. Let U be an open set in X, with $p \in U$. By openness, there exists $\varepsilon > 0$ s.t. $B_\varepsilon(p) \subseteq U$. Since $p_n \to p$, there exists $N = N(\varepsilon) \in \mathbb{N}$ s.t. for any $n \geq N$, $p_n \in B_\varepsilon(p) \subseteq U$. Thus U contains all of the p_n except $p_1, \cdots, p_{N(\varepsilon)-1}$ which are finitely many.

(ii) ($\Longleftarrow$) Every open set U s.t. $p \in U$ contains all but finitely many terms of the sequence. Let $\varepsilon > 0$, since $B_\varepsilon(p)$ is open, it contains all but finitely many of the p_n. Let

$$M := \max\{i \mid p_i \notin B_\varepsilon(p)\}$$

Then take $N := M + 1$, for any $n \geq N$, one has

$$p_n \in B_\varepsilon(p)$$

And this means $p_n \to p$.

(b) Given $\varepsilon > 0$, there exists $N(\varepsilon)$ s.t. $d(p_n, p) < \varepsilon$ for any $n \geq N(\varepsilon)$. Also there exists $M(\varepsilon)$ s.t. $d(p_n, p') < \varepsilon$ for any $n \geq M(\varepsilon)$. Hence for any

$$n \geq \max\{M, N\}$$

one has

$$d(p, p') \leq d(p, p_n) + d(p_n, p') \leq 2\varepsilon$$

But $\varepsilon > 0$ is arbitrary, hence we send $\varepsilon \to 0$. Thus

$$d(p', p) = 0 \iff p = p'$$

(c) Let $p_n \to p$. Take $\varepsilon = 1$. Then $B_1(p)$ contains all but finitely many terms $p_{n_1}, \cdots, p_{n_k}$. Take

$$r > \max(1, d(p, p_{n_1}), \cdots, d(p, p_{n_k})) > 0$$

One can choose this r because there's finitely many of the terms. It follows that $p_n \in B_r(p)$ for all n. This is to say, $\{p_n\}$ as a set is bounded.

(d) Since $E \subseteq X$ and p is a limit point of E, for any $n \in \mathbb{N}$, $B_{\frac{1}{n}}(p)$ contains points of E that are not equal to p. Choose $p_n \neq p$, $p_n \in B_{\frac{1}{n}}(p) \cap E$. Thus

$$d(p_n, p) < \frac{1}{n}$$

Indeed this implies convergence. To see this, for any $\varepsilon > 0$, choose $N(\varepsilon) > \frac{1}{\varepsilon}$ so that

$$d(p_n, p) < \frac{1}{n} \leq \frac{1}{N(\varepsilon)} \leq \varepsilon$$

for any $n \geq N(\varepsilon)$.

$\square$

3.1.1 Convergence in $\mathbb{R}$ or $\mathbb{R}^k$

Now we work in $\mathbb{R}$ or $\mathbb{C}$.

Proposition 3.1 (Rules of addition, scalar multiplication, multiplication, division). *If $\{s_n\}$ and $\{t_n\}$ are sequences of real numbers s.t.*

$$s_n \to s \in \mathbb{R} \qquad t_n \to t \in \mathbb{R}$$

Then it follows that

1. $\lim_{n\to\infty} s_n + t_n = s = t$
2. $\lim_{n\to\infty} cs_n = cs$ *for any* $c \in \mathbb{R}$ *constant.*
3. $\lim_{n\to\infty} s_n t_n = st$
4. $\lim_{n\to\infty} \frac{1}{s_n} = \frac{1}{s}$ *provided* $s \neq 0$.

Proof. 1. Let $\varepsilon > 0$. Since $s_n \to s$, there exists N_1 s.t.

$$|s_n - s| < \frac{\varepsilon}{2}$$

for any $n \geq N_1$. Since $t_n \to t$, there exists N_2 s.t.

$$|t_n - t| < \frac{\varepsilon}{2}$$

for any $n \geq N_2$. Thus

$$|s_n - s + t_n - t| \leq |s_n - s| + |t_n - t| < \varepsilon \qquad \forall\ n \geq \max(N_1, N_2)$$
$$|(s_n + t_n) - (s + t)| < \varepsilon \qquad \forall\ n \geq \max(N_1, N_2)$$

2. Let $\varepsilon > 0$, since $s_n \to s$, there exists $N = N(\varepsilon)$ s.t.

$$|s_n - s| < \frac{\varepsilon}{|c|} \qquad \forall\ n \geq N$$

Thus for any $n \geq N$

$$|cs_n - cs| = |c||s_n - s| < |c|\frac{\varepsilon}{|c|} = \varepsilon$$

3. Look at

$$\begin{aligned} |s_n t_n - st| &= |(s_n - s)(t_n - t) + s_n t + st_n - st - st| \\ &\leq |(s_n - s)(t_n - t)| + t|s_n - s| + s|t_n - t| \end{aligned}$$

We claim that $(s_n - s)(t_n - t) \to 0$. Let $\varepsilon > 0$, there exists N_1 s.t. for any $n \geq N_1$

$$|s_n - s| < \sqrt{\varepsilon}$$

and there exists N_2 s.t. for any $n \geq N_2$

$$|t_n - t| < \sqrt{\varepsilon}$$

Thus

$$|(s_n - s)(t_n - t)| < \varepsilon \qquad \forall\ n \geq \max(N_1, N_2)$$

4. Look at (WLOG let $s_n \neq 0$ for any n)

$$|\frac{1}{s_n} - \frac{1}{s}| = \frac{|s_n - s|}{s_n s}$$

We choose m s.t. $|s_n - s| < \frac{1}{2}|s|$ for any $n \geq m$. In particular, for any $n \geq m$

$$|s_n| \geq \frac{1}{2}|s|$$

Now given $\varepsilon > 0$, since $s_n \to s$, there exists $N \in \mathbb{N}$ s.t.

$$|s_n - s| < \frac{1}{2}|s|^2\varepsilon \qquad \forall\ n \geq N$$

Now for any $n \geq \max\{N, m\}$

$$\begin{aligned}|\frac{1}{s_n} - \frac{1}{s}| = |\frac{s_n - s}{s_n s}| &< \frac{2}{|s|^2}|s_n - s| \\ &< \frac{2}{|s|^2}\frac{1}{2}|s|^2\varepsilon = \varepsilon\end{aligned}$$

□

Proposition 3.2 (Convergence in $\mathbb{R}^k$ iff in coordinates)**.** *For sequence $\vec{x_n} \in \mathbb{R}^k$ with*

$$\vec{x_n} = (\alpha_{1,n}, \cdots, \alpha_{k,n})$$

One has

$$\vec{x_n} \to \vec{x} = (\alpha_1, \cdots, \alpha_n) \iff \lim_{n\to\infty} \alpha_{j,n} = \alpha_j \qquad \forall\ j = 1, \cdots, k$$

Notice the LHS is convergence in $(\mathbb{R}^k, \|\cdot\|)$ while the RHS is convergence in $(\mathbb{R}, |\cdot|)$.

Proof. 1. ($\implies$) Assume $\vec{x_n} \to \vec{x}$ in $\mathbb{R}^k$. Then for any $\varepsilon > 0$, there exists N s.t.

$$\|\vec{x_n} - \vec{x}\| < \varepsilon \qquad \forall\ n \geq N$$

Notice

$$\|\vec{x_n} - \vec{x}\| = \sqrt{\sum_{j=1}^{k} |\alpha_{j,n} - \alpha_j|^2}$$

Hence

$$|\alpha_{j,n} - \alpha_j| < \varepsilon \qquad \forall\ n \geq N \qquad j = 1, \cdots, k$$

Hence $\alpha_{j,n} \to \alpha_j$ for all $j = 1, \cdots, k$.

2. ($\impliedby$) Assume $\alpha_{n,j} \to \alpha_j$ for $j = 1, \cdots, k$. Let $\varepsilon > 0$, then there exists N_j s.t.

$$|\alpha_{n,j} - \alpha_j| < \frac{\varepsilon}{\sqrt{k}}$$

Now let

$$N := \max\{N_j \mid j = 1, \cdots, k\} > 0$$

So for any $n \geq N$,

$$\|\vec{x_n} - \vec{x}\| = \sqrt{\sum_{j=1}^{k} |\alpha_{j,n} - \alpha_j|^2} < \varepsilon$$

□

In fact the argument works for other norms

$$\|v\|_p = \left(\sum_i |v_i|^p\right)^{\frac{1}{p}}$$

$$\|v\|_\infty = \max_i |v_i|$$

These norms give rise to the same topology!

3.2 Subsequence Convergence

Definition 3.3 (Subsequence)**.** *Let $\{p_n\}$ be a sequence in (X, d). If we have integers*

$$n_1 < n_2 < \cdots < n_k < \cdots$$

Then $\{p_{n_k}\}$ is a subsequence of $\{p_n\}$.

Example 3.2. *Consider $x_k = (-1)^k$. For k even, $x_k = 1$ while for k odd, $x_k = -1$. The original sequence does not converges, but the subsequences*

$$x_{2k} \qquad \text{and} \qquad x_{2k+1}$$

converge.

Now we discuss the beauty of compactness! We have convergent subsequence :)

Theorem 3.2 (Sequential Compactness; Bolzano-Weierstrass)**.** *This is our well-known Bolzano-Weierstrass.*

1. *Assume (X, d) is compact. Then every sequence $\{p_n\}$ in X has a subsequence converging to a point in X.*
2. *Every bounded sequence $\{p_n\} \subseteq \mathbb{R}^k$ contains a convergent subsequence.*

Proof. 1. We distinguish two cases.

(a) For the first case, $\{x_n\}$ actually only contains finitely many distinct points of X. One of these points of X occurs infinitely many times, i.e., there exists

$$n_1 < n_2 < n_3 < \cdots < n_k < \cdots$$

s.t.

$$x_{n_1} = \cdots = x_{n_k} = \cdots = p$$

Then take

$$x_{n_k} \equiv p$$

as the constant sequence, which trivially converges to p.

(b) For the second case, $\{x_n\}$ contains infinitely many distinct points of X. By compactness of X Theorem 2.2, $\{x_n\}$ has a limit point in X. Let's be cautious. They cluster somewhere, but we need to put them in the correct order so that convergence works. We choose

$$n_1 \qquad \text{s.t.} \qquad d(p, x_{n_1}) < 1$$

We choose

$$n_2 \qquad \text{s.t.} \qquad d(p, x_{n_2}) < \frac{1}{2}, \text{ and } n_2 > n_1$$

Why can I choose $n_2 > n_1$? Remember p is a limit point. By Proposition 2.7, every ball around the limit points contains infinitely many elements of the set. So if there does not exist $n_2 > n_1$, there's only finitely many points around, contradiction. Now by induction, with

$$n_1, \cdots, n_{k-1} \qquad \text{chosen}$$

We want to choose $n_k > n_{k-1}$ s.t. $d(p, p_{n_k}) < \frac{1}{k}$. Hence we've constructed a subsequence p_{n_k} s.t.

$$p_{n_k} \to p$$

2. Every bounded sequence in $\mathbb{R}^k$ fits inside a k-cell I, which is compact. Now apply the first result.

□

We prove a lemma that will come in handy later.

Lemma 3.1. *The set $E^* \subseteq X$ of subsequential limits of a sequence $\{p_n\}$ is closed in X.*

Remark 3.1. *Consider enumeration of $\mathbb{Q} = \{q_n\}_{n \in \mathbb{N}}$. The set of its subsequential limits is in fact uncountable. However, one can show that it is closed.*

Proof. Let q be a limit point of E^*. We will create a subsequence of $\{p_n\}$ that converges to q. If we can do so, then $q \in E^*$. Now q as limit point of E^* is arbitrary, hence E^* contains all its limits, hence is closed.

Now choose n_1 s.t. $p_{n_1} \neq q$ (otherwise $p_n = q$ for any n, i.e., $p_n \equiv q$ and then $E^* = \{q\}$ so there's nothing to prove). We choose

$$\delta := d(q, p_{n_1})$$

Now assume $n_1, \cdots, n_{k-1}$ are chosen. Since q is a limit point of E^*, there exists $x \in E^*$ with

$$d(x, q) < 2^{-k}\delta$$

Since $x \in E^*$, there's a subsequence of $\{p_n\}$ converging to x. Now we choose $n_k > n_{k-1}$ s.t.

$$d(p_{n_k}, x) < 2^{-k}\delta$$

Thus

$$d(p_{n_k}, q) \leq d(p_{n_k}, x) + d(x, q) \leq 2^{-k+1}\delta$$

The trick is always working with half the distances. Notice here we have two sequences piling up with one another. Now

$$p_{n_k} \to q$$

i.e., p_{n_k} is a subsequential limit of p_n, but this is exactly what it means for $q \in E^*$. $\square$

3.3 Cauchy Sequence in Metric Space

Definition 3.4 (Cauchy Sequence). *Let (X, d) be metric space. $\{p_n\}$ is a Cauchy sequence if for every $\varepsilon > 0$, there exists $N(\varepsilon)$ s.t. for any m, $n \geq N$,*

$$d(p_n, p_m) < \varepsilon$$

Notice in the definition of Cauchy Sequence, there's no limit!

Remark 3.2. *If a sequence is convergent, indeed it is Cauchy, since for any $\varepsilon > 0$, there exists $N = N(\varepsilon) > 0$ s.t. for any n, $m \geq N$*

$$d(x_n, x) < \frac{\varepsilon}{2}, \qquad d(x_m, x) < \frac{\varepsilon}{2} \implies d(x_n, x_m) < \varepsilon$$

On the other hand, if a sequence is Cauchy, it may not be convergent. For example, a sequence $\{q_n\} \subseteq \mathbb{Q}$ may converge to some irrational number in $\mathbb{R}$, but the limit does not lie in $\mathbb{Q}$ hence $\{q_n\}$ is not convergent in $\mathbb{Q}$. But $\{q_n\}$ is Cauchy.

Remark 3.3. *One may think of Cauchy sequence more geometrically. For $E \subseteq X$, one may define its diameter*

$$\operatorname{diam}(E) := \sup\{d(x, y) \mid x,\, y \in E\}$$

Notice E is bounded iff $\operatorname{diam}(E) < \infty$.

Now take any $\{x_n\} \subseteq X$, define the tail terms as

$$E_N := \{x_N, x_{N+1}, \cdots\} \subseteq \{x_n\} \subseteq X$$

Lemma 3.2. *$\{x_n\}$ is a Cauchy sequence iff*

$$\lim_{N \to \infty} \operatorname{diam} E_N = 0$$

Proof. Let $\{x_n\}$ be Cauchy sequence. Then for any $\varepsilon > 0$, there exists $N = N(\varepsilon) > 0$ s.t. for any n, $m \geq N$, $d(x_n, x_m) < \varepsilon$, hence fix $n := N$ and let m vary

$$d(x_N, x_m) < \varepsilon \qquad \forall\, m \geq N \qquad \iff \qquad \sup\{d(x_N, x_m) \mid m \geq N\} = \operatorname{diam} E_N \leq \varepsilon$$

$\square$

In fact we have the following definition.

Definition 3.5 (Complete Metric Space). *A metric space (X, d) is complete if any Cauchy Sequence converges.*

We discuss a useful tool for later purpose, which basically says a nested sequence of non-empty compact sets whose diameter goes to zero necessarily has one point in the intersection.

Proposition 3.3. *Let (X, d) be metric space. Let $E \subseteq X$ and $\overline{E}$ denote closure of E. Then*

1. $\text{diam}E = \text{diam}\overline{E}$*.*

2. If $K_n \supseteq K_{n+1} \supseteq \cdots$ *is a sequence of non-empty compact spaces and*

$$\lim_{n\to\infty} \text{diam}K_n = 0$$

Then

$$\bigcap_{n=1}^{\infty} K_n \qquad \textit{contains exactly one point}$$

Proof. 1. Since $E \subseteq \overline{E}$, by definition $\text{diam}(E) \leq \text{diam}(\overline{E})$. On the other hand, for any $\varepsilon > 0$, and $p, q \in \overline{E}$, since they're limit points of E, there exists $p' \neq p$, $p' \in E$, and $q' \neq q$, $q' \in E$ s.t.

$$d(q, q') < \varepsilon \qquad d(p, p') < \varepsilon$$

Hence we obtain

$$d(p, q) \leq d(p, p') + d(p', q') + d(q', q) \leq d(p', q') + 2\varepsilon$$

We want to take supremum. But be careful of the order in which we do that! First notice $d(p', q') \leq \text{diam}E$, hence in the first step we obtain

$$d(p, q) \leq \text{diam}E + 2\varepsilon$$

Now the RHS is independent of points p, $q \in \overline{E}$, so we can take supremum on the LHS to obtain

$$\text{diam}(\overline{E}) \leq \text{diam}(E) + 2\varepsilon$$

But now LHS is independent of ε. We send $\varepsilon \to 0$ to obtain

$$\text{diam}(\overline{E}) \leq \text{diam}(E)$$

and this is the art of analysis :)

2. Let $K = \bigcap_{n=1}^{\infty} K_n$. By the Finite Intersection Property Proposition 2.11, one know that $K \neq \varnothing$. Assume that K has 2 or more points. Then necessarily $\text{diam}K > 0$ (due to definition of metric!) Since $K \subseteq K_n$ for any n, we know

$$\text{diam}(K) \leq \text{diam}(K_n) \qquad \forall\ n$$

Now sending $n \to \infty$ yields

$$\text{diam}(K) = 0$$

a contradiction.

□

Now we're ready to prove our big Theorem about how Cauchy Sequence and Convergent sequences are related, and how compactness helps us!

Theorem 3.3 (Cauchy and Convergence)**.** *1. In any metric space* (X, d)*, a convergent sequence is Cauchy.*

2. If X *is compact, then Cauchy implies convergence.*

3. If $X = \mathbb{R}^k$*, then Cauchy implies convergence.*

Proof. The first item is already proved.

1. Let $\{p_n\} \subseteq X$ be a Cauchy sequence in X a compact space. Denote

$$E_N := \{x_N, x_{N+1}, \cdots\}$$

Using Lemma 3.2, we know

$$\lim_{n\to\infty} \text{diam}E_n = 0$$

Now use Proposition 3.3 to $\overline{E_n}$ (compact because Proposition 2.10) one obtain

$$\bigcap_{n=1}^{\infty} \overline{E_n} = \{p\}$$

Our claim is $p_n \to p$. Why? Remember $\text{diam}E_N \to 0$. So for any $\varepsilon > 0$, there exists N_0 s.t. for any $N \geq N_0$

$$\text{diam}\overline{E_N} \leq \varepsilon \implies d(p_n, p) \leq \varepsilon \qquad \forall\ n \geq N \geq N_0$$

Why can we pick p? Because we're working with closure of E_N! And we know $p \in \overline{E_N}$ for any N.

2. If $X = \mathbb{R}^k$, let $\{p_n\} \subseteq \mathbb{R}^k$ be Cauchy. We claim $\{p_n\}$ is bounded. Indeed, using Lemma 3.2

$$\lim_{N\to\infty} \mathrm{diam} E_N = 0 \implies \mathrm{diam} E_N \le 1 \qquad \forall\, N \ge N_0 \qquad \text{for some } N_0$$

Thus

$$\mathrm{diam}\{p_n\} \le \mathrm{diam}\{p_1, \cdots, p_{N_1}\} + \mathrm{diam} E_N < \infty$$

Now $\overline{\{p_n\}}$ is closed and bounded in $\mathbb{R}^k$. Why bounded? Use Proposition 3.3 item one

$$\mathrm{diam}\overline{\{p_N\}} = \mathrm{diam}\{p_n\} < \infty$$

hence using Heine-Borel 2.3 $\overline{\{p_n\}}$ is compact. Now we use the previous item.

□

Let's summarize when we have completeness.

Corollary 3.1. *1. $\mathbb{R}^k$ is complete.*

2. Compact sets are complete.

3. Closed subsets of complete spaces are complete.

3.4 Cauchy Sequence in $\mathbb{R}$

Now we work with $X = \mathbb{R}$.

Definition 3.6. *A sequence $\{s_n\} \subseteq \mathbb{R}$ is*

1. *monotonically increasing if $s_n \le s_{n+1}$ for any n*
2. *monotonically decreasing if $s_n \ge s_{n+1}$ for any n*

We have a big convergence Theorem in $\mathbb{R}$!

Theorem 3.4 (Monotone Convergence Theorem)**.** *If $\{s_n\} \subseteq \mathbb{R}$ is monotonically increasing or decreasing, and bounded, then $s_n \to s$ for some $s \in \mathbb{R}$.*

Proof. WLOG let's assume $\{s_n\}$ is increasing. Since $\{s_n\}$ is bounded, denote

$$s := \sup_n \{s_n\}$$

We want to show $s_n \to s$. For any $\varepsilon > 0$, $s - \varepsilon$ is not an upper bound of $\{s_n\}$. Hence there exists N s.t.

$$s - \varepsilon < s_N \le s$$

But what do we know about the sequence $\{s_n\}$? For any $n \ge N$,

$$s - \varepsilon < s_n \le s \qquad \forall\, n \ge N$$

But this implies

$$|s_n - s| < \varepsilon \qquad \forall\, n \ge N$$

□

Next we need to introduce certain notations describing going to infinity.

1. We denote
$$\lim_{n\to\infty} x_n = \infty$$
if for any $M \in \mathbb{R}$, there exists $N \in \mathbb{N}$ s.t. for any $n \ge N$, $x_n \ge M$.

2. Similarly, we denote
$$\lim_{n\to\infty} x_n = -\infty$$
if for any $M \in \mathbb{R}$, there exists $N \in \mathbb{N}$ s.t. for any $n \ge N$, $x_n \le M$.

3.4.1 Limsup and Liminf

We introduce very important concepts! Limsup and Liminf. Given sequence $\{s_n\} \subseteq \mathbb{R}$

Definition 3.7. *We define*

$$\limsup_{n\to\infty} s_n := \inf_{n\in\mathbb{N}} \sup_{j\geq n} s_j$$

and

$$\liminf_{n\to\infty} s_n := \sup_{n\in\mathbb{N}} \inf_{j\geq n} s_j$$

The important thing about lim sup and lim inf is that, if we allow $\pm\infty$ to be consider, then for any sequence $\{s_n\} \subseteq \mathbb{R}$, its $\limsup_{n\to\infty} s_n$ and $\liminf_{n\to\infty} s_n$ always exists.

Lemma 3.3. *Given any $\{s_n\} \subseteq \mathbb{R}$, its $\limsup_{n\to\infty} s_n$ and $\liminf_{n\to\infty} s_n$ either exists in $\mathbb{R}$, or is $\pm\infty$.*

Proof. WLOG we prove for $\limsup_{n\to\infty} s_n$. Notice the sequence

$$a_n := \sup_{j\geq n} s_j$$

What properties does it have? If we enlarge n, then we have less choices to maximize s_j compared to the previous one, hence the sup is necessarily non-increasing. In other words

$$a_n = \sup_{j\geq n} s_j \geq \sup_{j\geq n+1} s_j = a_{n+1}$$

Now our sequence $\{a_n\}$ is decreasing!

1. If $\{a_n\}$ is bounded, then there by Monotone Convergence Theorem 3.4, there exists $s \in \mathbb{R}$ s.t.

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{j\geq n} s_j = s$$

But recall a_n is decreasing, so its limit is necessarily its infimum, so

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{j\geq n} s_j = \inf_{n\in\mathbb{N}} \sup_{j\geq n} s_j \equiv \limsup_{n\to\infty} s_n = s$$

Thus our Limsup exists.

2. If $\{a_n\}$ is not bounded, then

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{j\geq n} s_j = \inf_{n\in\mathbb{N}} \sup_{j\geq n} s_j \equiv \limsup_{n\to\infty} s_n = -\infty$$

□

Remark 3.4. *Alternatively, for $\{s_n\} \subseteq \mathbb{R}$ one can define*

$$E := \{\textit{all subsequential limits of } s_n, \textit{ possibly including } \pm\infty\}$$

and define

$$\limsup_{n\to\infty} s_n = \sup E$$
$$\liminf_{n\to\infty} s_n = \inf E$$

One can actually verify that they're equivalent definitions.

Example 3.3. *If $s_n \to s$, then $E = \{s\}$. Hence*

$$\limsup_{n\to\infty} s_n = \liminf_{n\to\infty} s_n = s$$

Example 3.4. *If $s_n = (-1)^n$, then $E = \{\pm 1\}$. In this case*

$$\limsup_{n\to\infty} s_n = 1$$
$$\liminf_{n\to\infty} s_n = -1$$

Example 3.5. *If $s_n = \frac{(-1)^n}{1-\frac{1}{n}}$ for any $n > 1$. Then $E = \{\pm 1\}$. In this case*

$$\limsup_{n\to\infty} s_n = 1$$
$$\liminf_{n\to\infty} s_n = -1$$

4 Series

At first glance, one might think a series is an infinite $a_1 + a_2 + \cdots$. That's true, but not a good intuition.

We would love to create(or reinterpret) some new function $f : \mathbb{R} \to \mathbb{R}$ s.t. $\frac{P}{Q}$ polynomial division, e^x, $\sin(x)$ etc, in the following form

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$
$$f(x) = \sum_n c_n x^n \qquad x \in \mathbb{R}$$
$$f(z) = \sum_n c_n z^n \qquad z \in \mathbb{C}$$

Now we rigorously introduce what series are.

Definition 4.1. *$\{a_n\}$ is a sequence of reals (or could be complex). Define its partial sum as*

$$S_k := \sum_{n=1}^{k} a_n$$

as sum of the first k terms. If the sequence $S_k \to S$ for some $S \in \mathbb{R}$, we denote

$$\sum_{n=1}^{\infty} a_n := S$$

In other words

$$\lim_{k\to\infty} \sum_{n=1}^{k} a_n = \lim_{k\to\infty} s_k$$

In this case we say the series $\sum_n a_n$ is convergent. If S does not exists in $\mathbb{R}$, we say series $\sum_n a_n$ is divergent.

Since we're working in $\mathbb{R}$, Cauchy characterizes Convergence.

Theorem 4.1 (Cauchy Criterion)**.** *By using definition of Cauchy Sequence, one directly obtain*

1. *A series $\sum_n a_n$ is convergent iff for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ s.t. for any $n > m \geq N$*

$$|\sum_{k=m}^{n} a_k| < \varepsilon$$

2. *Alternatively we could say S_k converges iff for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ s.t. for any $n > m \geq N$*

$$|S_m - S_n| < \varepsilon$$

An immediate and simple corollary.

Corollary 4.1. *If $\sum_n a_n$ is convergent, then $a_n \to 0$.*

Proof. For $\varepsilon > 0$ there exists $N \in \mathbb{N}$ s.t. for any $n > m \geq N$

$$|\sum_{k=m}^{n} a_k| < \varepsilon \implies |a_n| < \varepsilon$$

But $n \geq N$ is arbitrary. □

However, the reverse is false. Think about

$$a_n = \frac{1}{n} \to 0$$

Yet we'll see that

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

hence it is divergent.

Theorem 4.2 (Monotone Convergence Criterion)**.** *A series of non-negative terms converges if it is bounded, i.e., if $a_n \geq 0$ for any n, and define $S_n := \sum_{k=1}^{n} a_k$, then if $\{S_n\}_n$ is bounded, one has*

$$\sum_n a_n < \infty$$

Proof. Note

$$S_{n+1} = S_n + a_{n+1} \geq S_n \qquad \forall\, n$$

Hence S_n is increasing. Using $\{S_n\}$ bounded, one apply Theorem 3.4. □

4.1 Convergence Tests

Recall at this moment, we only know the series

$$\sum_{n=1} x^n = \frac{1}{1-x} \qquad 0 \le x < 1$$

We want to use this to craft some convergence tests for series.

Comparison Test

Theorem 4.3 (Comparison Tests). *We need powerful comparison tests.*

1. *If $|a_n| \le c_n$ for any $n \ge N_0$ with some $N_0 \in \mathbb{N}$ fixed, and $\sum_n c_n$ converges, then*

$$\sum_{n}^{\infty} a_n < \infty$$

2. *If $a_n \ge d_n \ge 0$ for any $n \in N$ and $\sum_n d_n$ diverges, then*

$$\sum_n a_n = \infty$$

Proof. 1. For any $\varepsilon > 0$, there exists $N \ge N_0$ s.t. for any $n > m \ge N$

$$\sum_{k=m}^{n} c_k < \varepsilon$$

using Cauchy Criterion Theorem 4.1. Thus

$$|\sum_{k=m}^{n} a_k| \le \sum_{k=m}^{n} |a_k| \le \sum_{k=m}^{n} c_k < \varepsilon$$

This implies $\sum_n a_n$ converges, again by Cauchy Criterion Theorem 4.1.

2. We use contrapositive. If $\sum_n a_n$ converges, then so does $\sum_n d_n$. Contradiction.

□

Example 4.1. *Why does $\sum_{n=1} \frac{1}{n}$ diverge? We compare with*

$$\begin{aligned} &1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \cdots \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots = \infty \end{aligned}$$

Geometric Series Test The second family is known as Geometric Series test. This is rather standard and should be our model case.

Theorem 4.4 (Geometric Series Test). *If $0 \le x < 1$, then*

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

If $x \ge 1$, then $\sum_{n=0}^{\infty} x_n$ diverges.

Proof. Let

$$S_n := \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1-x}$$

So for $0 \le x < 1$

$$\lim_{n\to\infty} S_n = \frac{1}{1-x}$$

□

Root and Ratio Tests The third family is Root and Ratio Tests. They're extremely powerful. The idea is as following: Given series $\sum_n a_n$

1. If $|a_n|^{\frac{1}{n}} \sim$ constant, then $|a_n| \sim (\text{constant})^{\frac{1}{n}}$ and this turns $\sum_n a_n$ into a geometric series.
2. If $\frac{a_{n+1}}{a_n} \sim$ constant, then $a_n \sim (\text{constant})^{\frac{1}{n}}$ and this again turns $\sum_n a_n$ into a geometric series.

Theorem 4.5 (Root Test)**.** *Given $\sum_n a_n$ series, denote*

$$\alpha := \lim_{n\to\infty} |a_n|^{\frac{1}{n}}$$

1. If $\alpha < 1$ the series converges.

2. If $\alpha > 1$ the series diverges.

3. $\alpha = 1$ is inconclusive.

Proof. 1. Since $\alpha < 1$, there exists $\beta \in (\alpha, 1)$. Use definition of

$$\alpha = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$$

there exists $N \in \mathbb{N}$ s.t. for any $n \geq N$

$$|a_n|^{\frac{1}{n}} < \beta \iff |a_n| < \beta^n$$

Now use Geometric Series Test Theorem 4.4 so that

$$\sum_{n=0}^{\infty} \beta^n < \infty$$

and then Comparison Test Theorem 4.3 so that

$$\sum_{n=0}^{\infty} a_n < \infty$$

2. If $\alpha > 1$, use definition of α as Limsup, there exists a subsequence n_k s.t.

$$|a_{n_k}|^{\frac{1}{n_k}} \to \alpha > 1 \qquad k \to \infty$$

Thus there exists $N \in \mathbb{N}$ s.t. for any $n_k \geq N$

$$|a_{n_k}|^{\frac{1}{n_k}} \geq 1 \iff |a_{n_k}| \geq 1 \implies |a_n| \not\to 0$$

Using Corollary 4.1 one obtain $\sum_n a_n$ is divergent.

□

Theorem 4.6 (Ratio Test)**.** *Consider series $\sum_n a_n$.*

1. $\sum_n a_n$ converges if

$$\limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}| < 1$$

2. $\sum_n a_n$ diverges if there exists N_0 s.t. for any $n \geq N_0$

$$|\frac{a_{n+1}}{a_n}| > 1$$

Proof. 1. Using definition of Limsup, there exists β and $N \in \mathbb{N}$ s.t. for any $n \geq N$

$$|\frac{a_{n+1}}{a_n}| < \beta < 1 \implies |a_{N+p}| < \beta|a_{N+p-1}| < \cdots < \beta^p |a_N|$$

Now renaming $n = N + p$, so $p = n - N$ one obtain

$$|a_n| < |a_N|\beta^{n-N} = |a_N|\beta^{-N}\beta^n$$

Since $\beta < 1$ and use Geometric Series Test Theorem 4.4

$$\sum_{n=0}^{\infty} |a_N|\beta^{-N}\beta^n < \infty \implies \sum_n a_n < \infty$$

2. If $|a_{n+1}| \geq |a_n|$ for any $n \geq N_0$ for some N_0 fixed, then $a_n \not\to 0$. Using Corollary 4.1 one obtain $\sum_n a_n$ is divergent.

$\square$

Remark 4.1. *We remark that usually Root test works better than Ratio Test. There are cases where Root works but Ratio fails. Look at*

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots$$

Then

$$\limsup_{n\to\infty} |a_n|^{\frac{1}{n}} = \lim_{n\to\infty} (\frac{1}{2^n})^{\frac{1}{2n}} = \frac{1}{\sqrt{2}} < 1$$

Hence apply Theorem 4.5 we know the series converges. However

$$\limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}| = \lim_{n\to\infty} \frac{1}{2}(\frac{3}{2})^n = \infty$$

and hence is inconclusive.

4.2 Power Series

Now we encode the variable x into the series. Look at series

$$\sum_n c_n x^n$$

with c_n coefficients, and $x \in \mathbb{R}$ or $\mathbb{C}$ as variable. We ask two questions

1. For which values of x does $\sum_n c_n x^n$ converge?
2. If $\sum_n c_n x^n$ converges for $x \in I \subseteq \mathbb{R}$ on an interval, if we denote

$$f(x) = \sum_n c_n x^n$$

   Then how good is this function?

We make sense using Root Test. Notice

$$|c_n x^n|^{\frac{1}{n}} = |c_n|^{\frac{1}{n}} |x|$$

Then

$$\limsup_{n\to\infty} |c_n x^n|^{\frac{1}{n}} = |x| \limsup_{n\to\infty} |c_n|^{\frac{1}{n}}$$

We denote $\alpha := \limsup_{n\to\infty} |c_n|^{\frac{1}{n}}$. Using Root Test 4.5

1. $\sum_n c_n x^n$ converges if

$$|x| \cdot \alpha < 1 \iff |x| < \frac{1}{\alpha} =: R$$

   where R is known as the radius of convergence for the power series $\sum_n c_n x^n$.

2. $\sum_n c_n x^n$ diverges if

$$|x| \cdot \alpha > 1 \iff |x| > R$$

Now if $R > 0$, one can define the function

$$f : (-R, R) \to \mathbb{R} \qquad x \mapsto \sum_n c_n x^n$$

Example 4.2. *Compute the Radius of convergence for the power series*

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

We use Ratio test so

$$|\frac{x^{n+1}/(n+1)!}{x^n/n!}| = |\frac{x}{n+1}| \to 0 \qquad n \to \infty$$

for any $x \in \mathbb{R}$. Hence the radius of convergence $R = \infty$. Thus

$$f : \mathbb{R} \to \mathbb{R} \qquad x \mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

is a well-defined function (in the sense that at each point $x \in \mathbb{R}$, $f(x) \in \mathbb{R}$). In fact this is

$$f(x) = e^x$$

Definition 4.2 (Analytic)**.** *A function defined by*

$$f(x) := \sum_{n=0}^{\infty} c_n x^n$$

on $(-R, R)$ is called real analytic C^w on $(-R, R)$. In $\mathbb{R}$ this is usually better than being infinitely differentiable. In $\mathbb{C}$, however, if f' exists then f is analytic.

4.3 Rearrangements

Some warnings about Series operation. They're usually easy to add

$$\sum_n (a_n + b_n) = \sum_n a_n + \sum_n b_n$$

What about multiplication?

$$(\sum_n a_n)(\sum_n b_n)$$

Definition 4.3. *A series $\sum_n a_n$ converges absolutely if $\sum_n |a_n| < \infty$.*

Lemma 4.1. *If $\sum_n a_n$ converges absolutely, then $\sum_n a_n$ converges.*

Let $\{k_n\} \subseteq \mathbb{N}$ be a sequence of natural numbers where every positive integer appears once and only once. Denote

$$a'_n = a_{k_n}$$

Then we say $\sum_n a'_n$ is an rearrangement of $\sum_n a_n$. Sequence of partial sums of $\sum_n a'_n$ and $\sum_n a_n$ could contain totally different numbers.

We ask: Under what conditions all rearrangements of a convergent series will converge and whether the sums are necessarily the same.

Theorem 4.7. *If $\sum_n a_n$ converges absolutely, then every arrangements $\sum_n a'_n$ converges, and*

$$\sum_n a'_n = \sum_n a_n$$

Proposition 4.1. *If $\sum_n a_n$ converges but not absolutely, then for any $\alpha < \beta$ (including $\pm\infty$, there exists an arrangement $\sum_n a'_n$ s.t.*

$$\limsup_{n \to \infty} S'_n = \beta$$
$$\liminf_{n \to \infty} S'_n = \alpha$$

where S'_n denotes partial sums of $\sum_n a'_n$. In particular, one can rearrange $\sum_n a_n$ and make it to have any sum that you want.

5 Continuity

Given two metric spaces (X, d) and (Y, d), we want to define what is means for a function

$$f : X \to Y$$

to be continuous at a point $x \in X$. There's in fact three equivalent definitions.

1. The original definition using $\varepsilon - \delta$
2. Definition using limits of sequences.
3. Definition using only open sets. This is the most general one.

We begin by defining what

$$\lim_{x \to p} f(x) = q$$

means.

Definition 5.1. *Given two metric spaces (X, d) and (Y, d). Let $E \subseteq X$ and x be a limit point of E. Define a function*

$$f : E \to Y$$

We say that $f(x) \to q$ as $x \to p$, or

$$\lim_{x \to p} f(x) = q$$

if for every $\varepsilon > 0$, there exists $\delta = \delta(p, \varepsilon) > 0$ s.t. for any x with

$$0 < d_X(x, p) < \delta$$

one has

$$d_Y(f(x), q) < \varepsilon$$

Remark 5.1. *In the above definition, p may not lie in E, hence $f(p)$ may not be define. Even if $f(p)$ is define, we don't really care about the value $f(p)$. Continuity is about what f behaves arbitrarily near p, but not at p.*

Proposition 5.1 (Sequential Characterisation)**.** *Given two metric spaces (X, d) and (Y, d). Let $E \subseteq X$ and x be a limit point of E. Define a function*

$$f : E \to Y$$

Then

$$\lim_{x \to p} f(x) = q$$

iff

$$\lim_{n \to \infty} f(x_n) = q$$

for every sequence $\{p_n\} \subseteq E$ s.t. $p_n \neq p$ for any n and $p_n \to p$.

Proof. 1. $\Longrightarrow$. Let $p_n \to p$ and $p_n \neq p$ for any n. Take any $\varepsilon > 0$. By definition of $\lim_{x \to p} f(x) = q$, there exists $\delta = \delta(\varepsilon, p) > 0$ s.t.

$$0 < d_X(x, p) < \delta \implies d_Y(f(x), q) < \varepsilon$$

Using $p_n \to p$, given $\delta > 0$ as above, there exists $N = N(\delta) \in \mathbb{N}$ s.t. for any $n \geq N$

$$0 < d_X(p_n, p) < \delta$$

Thus using

$$\lim_{x \to p} f(x) = q$$

one get

$$d_Y(f(p_n), q) < \varepsilon \qquad \forall\ n \geq N$$

Viewing $\{f(p_n)\}_{n \in \mathbb{N}}$ as a sequence, this is to say that

$$\lim_{n \to \infty} f(x_n) = q$$

2. $\Longleftarrow$. We argue using contradiction. Assume not, i.e.,

$$f(x) \not\to q \qquad x \to p$$

Then there exists $\varepsilon > 0$ s.t. for any $\delta > 0$, there exists a point x_δ s.t. $0 < d_X(x_\delta, p) < \delta$ yet

$$d_Y(f(x_\delta), q) \geq \varepsilon$$

But here δ is arbitrary, hence we're allowed to pick. We choose

$$\delta_n = \frac{1}{n}$$

and denote $x_n := x_{\delta_n}$. Then

$$0 < d_X(x_n, p) < \frac{1}{n} \implies x_n \to p$$

Note $x_n \neq p$ for any n, so by our assumption, necessarily $f(x_n) \to q$. However for some $\varepsilon > 0$

$$d_Y(f(x_n), q) \geq \varepsilon$$

for any n excludes this possibility, and we have a contradiction.

□

Corollary 5.1 (Uniqueness). *If*

$$\lim_{x \to p} f(x) = q$$

exists, then it is unique.

Corollary 5.2 (Operation Rules).

$$\lim_{x \to p} (f(x) + g(x)) = \lim_{x \to p} f(x) + \lim_{x \to p} g(x)$$

and similarly for $f \cdot g$, $\frac{f}{g}$, $\langle \vec{f}, \vec{g} \rangle$ etc.

5.1 Continuity and its Propositions

Now we define what continuity at a point means.

Definition 5.2 (Continuity at p). *f is continuous at a point $p \in E$ (not just a limit point) if for every $\varepsilon > 0$, there exists $\delta > 0$ s.t. for any*

$$d_X(x, p) < \delta$$

one has

$$d_Y(f(x), f(p)) < \varepsilon$$

In other words, f is continuous at p iff

$$\lim_{x \to p} f(x) = f(p)$$

Remark 5.2. *One has two cases depending on the relationship between p and E.*

1. *If $p \in E$ is isolated, then there exists $N \in \mathbb{N}$ large s.t. for any $n \geq N$*

$$f(p) = f(p_n) \implies d_Y(f(p), f(p_n)) = 0 < \varepsilon \qquad \forall\ n \geq N$$

2. *If $p \in E$ is limit point, then continuity means*

$$p_n \to p \implies f(p_n) \to f(p)$$

Definition 5.3 (Continuous Function). *We say f is continuous function on E if f is continuous at p for any $p \in E$.*

Proposition 5.2 (Composition). *Let $f : X \to Y$ be continuous at $x \in X$ and $g : Y \to Z$ be continuous at $f(x) \in Y$. Then*

$$g \circ f : X \to Z$$

is continuous at $x \in X$

Proof using Sequences. Let $x_n \to x$, continuity of f at x yields $f(x_n) \to f(x)$ while, continuity of g at $f(x)$ yields $g(f(x_n)) \to g(f(x))$. □

Proof using Definition. For any $\varepsilon > 0$, since g is continuous at $f(x)$, there exits $\delta_1 > 0$ s.t.

$$d_Y(y, f(x)) < \delta \implies d_Z(g(y), g(f(x))) < \varepsilon$$

But since f is continuous at x, there exists $\eta > 0$ s.t. for our given $\delta > 0$,

$$d_X(x, p) < \eta \implies d_Y(f(x), f(p)) < \delta$$

But now there exists such $\eta > 0$ (depending on ε) s.t. for any

$$d_X(x, p) < \eta \implies d_Y(f(x), f(p)) < \delta \implies d_Z(g(f(p)), g(f(x))) < \varepsilon$$

Hence $g \circ f$ is continuous at $x \in X$. □

In fact, one has characterisation of continuity merely using the definition of open sets.

Proposition 5.3 (Characterisation of Continuity using Open sets)**.** *A function $f : X \to Y$ is continuous iff $f^{-1}(V) \subseteq X$ is open in X for every $V \subseteq Y$ open in Y.*

Proof.

1. $\implies$. Let $V \subseteq Y$ be open. For any $x \in f^{-1}(V)$, i.e., $f(x) \in V$, since V is open, there exists $\varepsilon > 0$ s.t. $B_\varepsilon(f(x)) \subseteq V$. Now using continuity of f, there exists $\delta > 0$ s.t.

$$f(B_\delta(x)) \subseteq B_\varepsilon(f(x)) \subseteq V \implies B_\delta(x) \subseteq f^{-1}(V)$$

hence $f^{-1}(V)$ is open.

2. $\impliedby$. For any $x \in X$, for any $\varepsilon > 0$, the ball

$$B_\varepsilon(f(x)) \subseteq Y$$

is open. Hence using our assumption,

$$f^{-1}(B_\varepsilon(f(x))) \subseteq X$$

is open. Thus, since $x \in f^{-1}(B_\varepsilon(f(x)))$, x is an interior point, so there exists $\delta > 0$ s.t.

$$B_\delta(x) \subseteq f^{-1}(B_\varepsilon(f(x))) \implies f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$$

Thus f is continuous at x.

□

Proposition 5.4 (Composition)**.** *Let $f : X \to Y$ and $g : Y \to Z$ be continuous functions. Then*

$$g \circ f : X \to Z$$

is continuous.

Proof using Open sets. Let $V \subseteq Z$ be open, using g is continuous, $g^{-1}(V) \subseteq Y$ is open, using f is continuous,

$$f^{-1}(g^{-1}(V)) = (g \circ f)^{-1}(V) \subseteq X$$

is open. Since $V \subseteq Z$ open is arbitrary, one has $g \circ f$ continuous. □

In fact one has characterisation using closed sets as well.

Corollary 5.3. *$f : X \to Y$ is continuous iff $f^{-1}(C) \subseteq X$ is closed for any $C \subseteq Y$ closed.*

Let's look at a collection of examples for continuous functions.

Example 5.1.

1. *$f(x) = c$ constant, $f(x) = x$, $f(x) = x^n$ polynomials.*
2. *e^x, $\sin(x)$, $\cos(x)$.*
3. *Inner product*

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$$

Norm

$$\|\cdot\| : \mathbb{R}^n \to [0, \infty)$$

and metric

$$d : X \times X \to [0, \infty)$$

5.2 Continuity and Compactness

We first define what it means for a function to be bounded.

Definition 5.4. *$f : X \to \mathbb{R}^n$ is bounded if there exists $M > 0$ s.t.*

$$|f(x)| \leq M \qquad \forall\ x \in X$$

Proposition 5.5 (Continuity preserves Compactness)**.** *Let (X, d_X) be compact metric space and $f : X \to Y$ be continuous. Then $(f(X), d_Y)$ is a compact metric space.*

Proof. Let $\{V_\alpha\}_\alpha$ be an open cover of $f(X)$. Then using continuity of f

$$\{f^{-1}(V_\alpha)\}_\alpha$$

is open cover of X. Since X is compact, there exists $i = 1, \cdots, N$ s.t.

$$X \subseteq \bigcup_{i=1}^{N} f^{-1}(V_{\alpha_i})$$

But this means

$$f(X) \subseteq \bigcup_{i=1}^{N} V_{\alpha_i}$$

and thus we've extracted a finite subcover for $f(X)$. □

Corollary 5.4. *If $f : X \to \mathbb{R}^n$ is continuous and X is compact, then $f(X)$ is closed and bounded.*

Corollary 5.5 (Continuous Function over Compact set attains maximum and minimum)**.** *If $f : X \to \mathbb{R}$ is continuous and X is compact, then there exists $p, q \in X$ s.t.*

$$f(p) = \sup_{x \in X} f(x) < \infty \qquad f(q) = \inf_{x \in X} f(x) > -\infty$$

Proof. $f(X) \subseteq \mathbb{R}$ is bounded so by Least Upper Bound property, $\sup_{x \in X} f(x)$ and $\inf_{x \in X} f(x)$ exists in $\mathbb{R}$. Since they're limit points of $f(X)$, and using $f(X)$ is closed, indeed, $\sup_{x \in X} f(x)$ and $\inf_{x \in X} f(x)$ belongs to $f(X)$, i.e., there exists $p, q \in X$ s.t.

$$f(p) = \sup_{x \in X} f(x) \qquad f(q) = \inf_{x \in X} f(x)$$

□

Theorem 5.1. *Let $f : X \to Y$ be continuous and bijective. If X is compact, then*

$$f^{-1} : Y \to X \qquad f(x) \mapsto x$$

is continuous.

Proof. First note f bijection yields

$$f^{-1}(f(V)) = f(f^{-1}(V)) = V$$

To show f^{-1} is continuous, for any $V \subseteq X$, we want to show

$$(f^{-1})^{-1}(V) = f(V) \subseteq Y \qquad \text{is open}$$

where this uses the fact that f is bijective. Now $V \subseteq X$ open implies V^c closed in X. Since X is compact, using Proposition 2.10, $V^c \subseteq X$ is compact. Now since f is continuous, by Proposition 5.5, $f(V^c)$ is compact. But using Proposition 2.9 $f(V^c)$ is closed, hence

$$(f(V^c))^c = f(V)$$

is open, where we again used f is bijection. □

5.3 Uniform Continuity

Now we introduce uniform continuity, an important distinction with continuity.

Definition 5.5 (Uniformly Continuous)**.** *$f : X \to Y$ is uniformly continuous if for any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ independent of $x \in X$ s.t.*

$$d_X(p,q) < \delta \implies d_Y(f(p), f(q)) < \varepsilon$$

Remark 5.3. *The most important thing about uniform continuity is that our δ does not depend on $x \in X$, i.e., it should be a neighborhood that works uniformly for all points in X. The original definition that f is continuous at a point x somehow allow $\delta = \delta(\varepsilon, x) > 0$. Even if this holds for any point $x \in X$, this continuity is a weaker statement then f being uniformly continuous over X.*

Consider the example

$$f(x) = \frac{1}{x} \qquad \forall\, x \in (0, \infty)$$

Now for any $\varepsilon > 0$, and any $x > 0$, one can take

$$\delta = \delta(\varepsilon, x) > 0$$

s.t.

$$|\frac{1}{x} - \frac{1}{y}| = \frac{|x-y|}{xy} \leq \frac{\delta}{x(x-\delta)} < \varepsilon$$

where our choice of δ is hence

$$\varepsilon x^2 - \delta\varepsilon x - \delta > 0 \implies \textit{choosing} \qquad \delta = \frac{1}{2}\frac{\varepsilon x^2}{1+\varepsilon x} \qquad \textit{suffices}$$

Notice this δ depends both on ε and x. Indeed f over $(0,\infty)$ is continuous function, since certain $\delta(\varepsilon, x) > 0$ exists for each $x \in (0,\infty)$ we pick.

However, for $\varepsilon > 0$, there does not exist $\delta = \delta(\varepsilon) > 0$ that is independent of x that works. In particular, in our choice of $\delta(\varepsilon, x)$ as above, as one push $x \to 0$, $\delta \to 0$. This is key difference between continuity and uniform continuity.

Theorem 5.2 (Continuous Function over Compact set is Uniformly Continuous)**.** *Let $f : X \to Y$ be continuous and X is compact. Then f is uniformly continuous.*

Proof. For any $\varepsilon > 0$, since $f : X \to Y$ is continuous, for any $p \in X$, there exists $\delta(p) > 0$ s.t.

$$d_X(p,q) \leq \delta(p) \implies d_Y(f(p), f(q)) \leq \frac{\varepsilon}{2}$$

Now define

$$J(p) := \{q \in X \mid d_X(p,q) < \frac{1}{2}\delta(p)\}$$

and indeed this is an open cover of X

$$X \subseteq \bigcup_{p\in X} J(p)$$

Using X is compact, there exists p_i for $i = 1, \cdots, N$ s.t.

$$X \subseteq \bigcup_{i=1}^{N} J(p_i)$$

Define

$$\delta := \frac{1}{2}\min\{\delta(p_i) \mid i = 1, \cdots, N\} > 0$$

Then for any x, $y \in X$ s.t. $d(x,y) < \delta$, one has $x \in J(p_i)$ for some $i = 1, \cdots, N$, thus

$$d_X(p_i, x) < \frac{1}{2}\delta(p_i)$$

Immediately one obtain

$$d_Y(f(x), f(p_i)) \leq \frac{\varepsilon}{2}$$

Also, using $d(x,y) < \delta$, one obtain by triangle inequality

$$d(y, p_i) \leq d(y,x) + d(x, p_i) \leq \delta + \frac{1}{2}\delta(p_i) \leq \delta(p_i)$$

Thus
$$d_Y(f(y), f(p_i)) \le \frac{\varepsilon}{2}$$
Now we conclude using triangle inequality
$$d_Y(f(x), f(y)) \le d_Y(f(x), f(p_i)) + d_Y(f(p_i), f(y)) \le \varepsilon$$
□

5.4 Continuity and Connectedness

Proposition 5.6. *Let $f : X \to Y$ be continuous and X be connected. Then $f(X)$ is connected.*

Proof Sketch. If A and B were separated subsets s.t.
$$A \cup B = f(X)$$
Then
$$f^{-1}(A) \cup f^{-1}(B) = X$$
for $f^{-1}(A)$ and $f^{-1}(B)$ separated, hence X would be disconnected. Here continuity is used to understand how taking closure compares with taking preimage. □

Corollary 5.6 (Continuity preserves interval)**.** *For $f : I \to \mathbb{R}$ continuous and I is an interval, $f(I)$ is also an interval, i.e., for any $\alpha \in (f(p), f(q))$ for p, $q \in I$, there exists $r \in I$ s.t.*
$$f(r) = \alpha$$

5.5 Discontinuities

We first define one-sided limits.

Definition 5.6 (Left/Right Limits)**.** *Let $f : (a, b) \to \mathbb{R}$ and $x_0 \in (a, b)$. Then*
$$f(x_0^+) \equiv \lim_{x \to x_0^+} f(x)$$
$$f(x_0^-) \equiv \lim_{x \to x_0^-} f(x)$$
are Right and Left limits.

If f is discontinuous at x_0, necessarily
$$\lim_{x \to x_0} f(x) \neq f(x_0)$$
What could happen?

1. Type 1 (Simple Discontinuity). This happens when $\lim\limits_{x \to x_0^+} f(x)$ and $\lim\limits_{x \to x_0^-} f(x)$ both exists but are not equal.
2. Type 2. One or both of the one-side limits do not exist. E.g.
$$f(x) = \begin{cases} \sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Remark 5.4. *Monotone functions can only have jump discontinuities*

Proof Sketch. WLOG assume f is monotonically increasing, i.e., $a < b$ implies $f(a) \le f(b)$. Then let $x_0 \in I$ be a point of discontinuity for
$$f : I \to \mathbb{R}$$
Look at the set
$$A_{x_0} := \{f(x) \mid x < x_0\}$$
Then A_{x_0} is bounded from above by the value $f(x_0)$. Similarly
$$B_{x_0} := \{f(x) \mid x > x_0\}$$
is bounded from below by $f(x_0)$. Using Least Upper Bound property, both
$$\alpha := \sup A_{x_0} = \lim_{x \to x_0^-} f(x) \qquad \beta := \inf B_{x_0} = \lim_{x \to x_0^+} f(x)$$
exist. Hence x_0 is a jump discontinuity (type 1). □

Remark 5.5. *A monotone function only has at most countably many jump-discontinuities.*

Example 5.2. *Consider a function*

$$f : (a, b) \subseteq \mathbb{R} \to \mathbb{R}$$

and a countable subset

$$E = \{x_i\}_{i=1}^{\infty} \subseteq (a, b)$$

For each n, take positive number $c_n > 0$ s.t. $\sum_n c_n < \infty$. Now define

$$f(x) := \sum_{n \text{ s.t. } x_n < x} c_n$$

Then f has a jump of size c_n at $x = x_n$, and f is monotonically increasing.

6 Differentiability

Let
$$f : (a, b) \subseteq \mathbb{R} \to \mathbb{R}$$

Definition 6.1 (Differentiable at x_0)**.** *Fix $x_0 \in (a, b)$. We say f is differentiable at x_0 if the limit*
$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$
exists. If so, we denote
$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Definition 6.2 (Differentiable Function)**.** *If f is differentiable at every point $x \in (a, b)$, we say f is differentiable on (a, b). In this case we can define a new function*
$$f' : (a, b) \to \mathbb{R} \qquad x \mapsto f'(x)$$
In general, however, the domain of f' could shrink.

6.1 Basic Properties

Proposition 6.1 (Expansion of Differentiability)**.** *f is differentiable at $x = x_0$ iff there exists $c \in \mathbb{R}$ s.t.*
$$f(x_0 + h) = f(x_0) + ch + o(h)$$
where
$$\lim_{h \to 0} \frac{o(h)}{h} = 0$$
In fact, $c = f'(x_0)$.

Proof.

1. ($\Longrightarrow$) Define
$$o(h) = f(x_0 + h) - f(x_0) - f'(x_0)h$$
Then indeed
$$\lim_{h \to 0} \frac{o(h)}{h} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0) = 0$$

2. ($\Longleftarrow$) Compute
$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = c$$

$\square$

In this case we say $f(x_0)$ is the function value at base point x_0. $f'(x_0)h$ is the linear (first order) expansion of f at the point x_0. $o(h)$ is the error made by the linear approximation.

In fact, $f'(x_0)$ is the best linear approximation of f at $x = x_0$.

Remark 6.1. *If $\phi : V \to W$ where V and W are normed vector spaces, then we define ϕ to be differentiable at a point $x_0 \in V$ if*
$$\phi(x_0 + h) = \phi(x_0) + Ah + o(h)$$
where
$$A : V \to W$$
is a linear mapping and
$$\lim_{\|h\|_V \to 0} \frac{\|o(h)\|_W}{\|h\|_V} = 0$$
In this case we define
$$\phi'(x_0) := A$$

Proposition 6.2 (Differentiability implies Continuity)**.** *If f is differentiable at $x_0 \in (a, b)$, f is continuous at x_0.*

Proof. We compute

$$\begin{aligned}\lim_{x\to x_0} f(x) = \lim_{h\to 0} f(x_0+h) = \lim_{h\to 0}(f(x_0)+f'(x_0)h+o(h)) \\ = f(x_0)\end{aligned}$$

which proves continuity. □

Remark 6.2. *The converse is not true. If $f(x)=|x|$, then f is continuous at $x=0$, but f is not differentiable at 0.*

Proposition 6.3. *If f and g are differentiable at x, then*

$$\begin{aligned}(f\pm g)'(x) &= f'(x)\pm g'(x)\\ (f\cdot g)'(x) &= f'(x)g(x)+f(x)g'(x)\\ (\frac{f}{g})'(x) &= \frac{f'(x)g(x)-f(x)g'(x)}{g^2(x)}\end{aligned}$$

Proof. 1. Compute

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

2. Compute

$$\begin{aligned}(f\cdot g)(x+h) &= (f(x)+f'(x)h+o(h))\cdot(g(x)+g'(x)h+o(h))\\ &= f(x)g(x)+(f'(x)g(x)+f(x)g'(x))\cdot h+o(h)\\ \lim_{h\to 0}\frac{(f\cdot g)(x+h)-(f\cdot g)(x)}{h} &= f'(x)g(x)+f(x)g'(x)\end{aligned}$$

This is known as product rule.

3. It suffices to compute (assume $g(x)\neq 0$)

$$\begin{aligned}(\frac{1}{g(x)})' = \lim_{h\to 0}\frac{\frac{1}{g(x+h)}-\frac{1}{g(x)}}{h} = \lim_{h\to 0}\frac{\frac{g(x)-g(x+h)}{g(x+h)g(x)}}{h}\\ = -\frac{g'(x)}{g^2(x)}\end{aligned}$$

Then use product rule. This is known as quotient rule.

□

Let's see some examples.

Example 6.1. *For $f(x)=c$*

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

For $f(x)=x$

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=\lim_{h\to 0}\frac{x+h-x}{h}=1$$

For $f(x)=x^n$, by induction we assume

$$(x^n)'=nx^{n-1}\qquad n\in\mathbb{N}$$

Then for $n+1$

$$\begin{aligned}(x^{n+1})' &= (x\cdot x^n)' = x'x^n+x(x^n)'\\ &= x^n+xnx^{n-1}=(n+1)x^n\end{aligned}$$

For polynomials $f(x)=\sum_{n=0}^k a_nx^n$

$$f'(x)=\sum_{n=1}^k na_nx^{n-1}$$

Remark 6.3. *What is the derivative for $f(x) = \sum_{n=0}^{\infty} c_n x^n$ defined on $(-R, R)$? Is it true that one can exchange derivative and series summation? For power series, yes.*

Proposition 6.4 (Chain Rule)**.** *Assume f is differentiable at x and g is differentiable at $f(x)$, then*

$$(g \circ f)'(x) = g'(f(x))f'(x)$$

Proof.

$$\begin{aligned}
(g \circ f)(x+h) &= g(f(x+h)) = g(f(x) + f'(x)h + o(h)) \\
&= g(f(x)) + g'(f(x))(f'(x)h + o(h)) + o(f'(x)h + o(h)) \\
\lim_{h \to 0} \frac{(g \circ f)(x+h) - g(f(x))}{h} &= \lim_{h \to 0} g'(f(x))(f'(x) + \frac{o(h)}{h}) = g'(f(x))f'(x)
\end{aligned}$$

□

Example 6.2. *Consider*

$$f(x) = \begin{cases} \sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

then f is not continuous at $x = 0$ thus not differentiable at $x = 0$. But for

$$g(x) = \begin{cases} x\sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

We verify

1. *Note $|x\sin(\frac{1}{x})| \leq |x|$ Hence as $|x_n| \to 0$*

$$|g(x_n)| \leq |x_n| \to 0 \implies g(x_n) \to 0$$

This works for any sequence x_n going to 0, hence by sequential characterisation, g is continuous at $x = 0$.

2. *For any $x \neq 0$, g is indeed differentiable at x by chain rule*

$$g'(x) = \sin(\frac{1}{x}) + x\cos(\frac{1}{x}) \cdot (-\frac{1}{x^2})$$

But for $x = 0$

$$\frac{g(h) - g(0)}{h} = \frac{h\sin(\frac{1}{h})}{h} = \sin(\frac{1}{h})$$

as $h \to 0$ the RHS keeps oscillating, the limit does not exist. Hence g is not differentiable at 0.

Consider

$$h(x) = \begin{cases} x^2\sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Then indeed h is continuous at $x = 0$ and differentiable at any $x \neq 0$. For $x = 0$

$$\begin{aligned}
\frac{x^2\sin(\frac{1}{x})}{x} &= x\sin(\frac{1}{x}) \to 0 \qquad x \to 0 \\
h'(0) &= \lim_{x \to 0} \frac{x^2\sin(\frac{1}{x})}{x} = 0
\end{aligned}$$

Thus h is differentiable everywhere. In fact

$$h'(x) = \begin{cases} 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

and h' is not continuous at $x = 0$.

6.2 Mean Value Theorems

Definition 6.3 (local max)**.** *We say f has local maximum at x_0 if there exists $\delta > 0$ s.t.*

$$f(x) \leq f(x_0) \qquad |x - x_0| < \delta$$

Similarly one may define local min.

Proposition 6.5. *Let*

$$f : [a, b] \to \mathbb{R}$$

Suppose f is differentiable and has a local maximum at $c \in (a, b)$. Then $f'(c) = 0$.

Proof. Assume c is a local maximum. Then

$$\lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \leq 0$$
$$\lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \geq 0$$

But f is differentiable at $x = c$ hence

$$\lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'(c)$$

Thus $f'(c) = 0$. □

Corollary 6.1 (Rolle's Mean Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be differentiable over (a, b), continuous over $[a, b]$ and $f(a) = f(b)$. Then there exists $c \in (a, b)$ s.t.*

$$f'(c) = 0$$

Proof. Since $[a, b]$ is compact, and f is continuous, we know from Corollary 5.5 f attains a global max and global minimum. If both occur at endpoints, since $f(a) = f(b)$, f is constant over (a, b), thus $f' = 0$ throughout (a, b). If either maximum or minimum occurs in $c \in (a, b)$, then at that point $f'(c) = 0$. □

Corollary 6.2 (Lagrange Mean Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be differentiable over (a, b) and continuous over $[a, b]$. Then there exists $c \in (a, b)$ s.t.*

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Proof. We define function

$$L(x) := f(a) + \frac{f(b) - f(a)}{b - a}(x - a)$$

Then one check

$$L(a) = f(a)$$
$$L(b) = f(b)$$

Then define

$$g(x) := f(x) - L(x)$$

one obtain

$$g(a) = f(a) - L(a) = 0$$
$$g(b) = f(b) - L(b) = 0$$

Thus apply Corollary 6.1 to g one obtain there exists $c \in (a, b)$ s.t.

$$g'(c) = f'(c) - L'(c) = 0$$

Notice

$$L'(c) = \frac{f(b) - f(a)}{b - a}$$

Thus

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

□

What's good about MVT is that, one get first order derivative tests.

Proposition 6.6 (first order derivative test)**.** *f is differentiable over (a,b) and continuous over $[a,b]$.*

1. *If $f'(x) \geq 0$ on (a,b) then f is monotonically increasing.*
2. *If $f'(x) \leq 0$ on (a,b) then f is monotonically decreasing.*
3. *If $f'(x) = 0$ on (a,b), then f is constant.*

Proof. For any $a \leq x_1 < x_2 \leq b$, apply Corollary 6.2 so that there exists $c \in (x_1, x_2)$ s.t.

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$$

Hence monotonicity of f translates to sign of f'. □

Proposition 6.7 (Derivatives cannot have jump discontinuities)**.** *If $f'(a) < \lambda < f'(b)$, then there exists $c \in (a,b)$ s.t.*

$$f'(c) = \lambda$$

Proof. Let

$$g(x) = f(x) - \lambda x$$

Then

$$\begin{aligned} g'(a) &= f'(a) - \lambda < 0 \\ g'(b) &= f'(b) - \lambda > 0 \end{aligned}$$

Then neither a nor b is global minimum for g. Since $[a,b]$ is compact and g is continuous over $[a,b]$, the global minimum for g over $[a,b]$ however, must exist. Thus there exists $c \in (a,b)$ s.t.

$$g'(c) = 0 \implies f'(c) = \lambda$$

□

6.3 Taylor Series

Theorem 6.1 (Taylor's Theorem)**.** *Assume f is n times differentiable over $(a,b) \subseteq \mathbb{R}$. Let $\alpha \in (a,b)$. Write*

$$P_{n-1}(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k$$

Then for any $\beta \in (a,b)$ s.t. $\alpha \neq \beta$, there exists some t in between α and β s.t.

$$f(\beta) = P_{n-1}(\beta) + \frac{f^{(n)}(t)}{n!}(\beta - \alpha)^n \tag{3}$$

Remark 6.4. *The polynomial $P_{n-1}(x)$ is the only polynomial s.t.*

$$P_{n-1}^{(k)}(\alpha) = f^{(k)}(\alpha) \qquad \forall\ k = 0, \cdots, n-1$$

Proof of Theorem 6.1. Let

$$M = \frac{f(\beta) - P_{n-1}(\beta)}{(\beta - \alpha)^n}$$

This is constant. WLOG assume $\beta > \alpha$. Define

$$g(x) = f(x) - P_{n-1}(x) - M(x-\alpha)^n$$

Our task reduces to showing

$$n!M = f^{(n)}(t)$$

for some $t \in (\alpha, \beta)$. We compute

$$\begin{aligned} g(\alpha) &= 0 \\ g'(\alpha) &= 0 \\ &\vdots \\ g^{(n-1)}(\alpha) &= 0 \end{aligned}$$

On the other hand, note

$$g(\beta) = f(\beta) - P_{n-1}(\beta) - \frac{f(\beta) - P_{n-1}(\beta)}{(\beta - \alpha)^n}(\beta - \alpha)^n = 0$$

Using Corollary 6.1 there exists $x_1 \in (\alpha, \beta)$ s.t.

$$g'(x_1) = 0$$

Using $g'(\alpha) = 0$ and Corollary 6.1 there exists $x_2 \in (\alpha, x_1)$ s.t.

$$g^{(2)}(x_2) = 0$$

Keep iterating, there exists $t = x_n \in (\alpha, x_{n-1})$ s.t.

$$g^{(n)}(t) = 0$$

But what is $g^{(n)}$?

$$0 = g^{(n)}(t) = f^{(n)}(t) - n!M$$

Thus

$$\frac{f(\beta) - P_{n-1}(\beta)}{(\beta - \alpha)^n} = M = \frac{f^{(n)}(t)}{n!}$$

□

Remark 6.5. *If f has derivatives of all orders, then $f^{(k)}(x)$ exists for any $k \in \mathbb{N}$. We define*

$$P(x) := \sum_{k=0}^{\infty} \frac{f^{(k}(\alpha)}{k!}(x - \alpha)^k$$

as Taylor Series of f at α. In general $f(x) \neq P(x)$ even in the domain of radius of convergence. In fact such f are called analytic.

7 Riemann Integration

In Calculus one need

$$f : [a, b] \to \mathbb{R}$$

continuous, or piecewise continuous, to define derivatives. But in general, we only need f bounded and the length of intervals of $\mathbb{R}$ to define integrals.

Definition 7.1 (Partition)**.** *Let $[a, b] \subseteq \mathbb{R}$. A partition of $[a, b]$ is a finite set $x_0, \cdots, x_n$ in $[a, b]$ s.t.*

1. *$x_0 = a$ and $x_n = b$*
2. *$x_0 \leq x_1 \leq \cdots \leq x_n$. We denote*

$$\Delta x_i := x_i - x_{i-1} \qquad \forall\ i \in \{1, \cdots, n\}$$

as the length of the ith interval.

Definition 7.2 (Riemann Integral)**.** *Suppose*

$$f : [a, b] \to \mathbb{R}$$

be bounded. Given a partition P of $[a, b]$ one can define

$$M_i := \sup_{x_{i-1} \leq x \leq x_i} f(x) \qquad \forall\ i \in \{1, \cdots, N\}$$
$$m_i := \inf_{x_{i-1} \leq x \leq x_i} f(x) \qquad \forall\ i \in \{1, \cdots, N\}$$

Notice all M_i and m_i exist because f is bounded over $[a, b]$.

1. *We define Upper and Lower Riemann Sum of f over $[a, b]$ w.r.t. Partition P as*

$$U(P, f) = \sum_{i=1}^{N} M_i \Delta x_i$$
$$L(P, f) = \sum_{i=1}^{N} m_i \Delta x_i$$

2. *We define Upper and Lower Riemann Integral of f over $[a, b]$ as*

$$\overline{\int_a^b} f := \inf_{P \text{ partition of } [a,b]} U(P, f)$$
$$\underline{\int_a^b} f := \sup_{P \text{ partition of } [a,b]} L(P, f)$$

3. *If $\overline{\int_a^b} f = \underline{\int_a^b} f$, we define the Riemann Integral of f over $[a, b]$ as*

$$\int_a^b f = \overline{\int_a^b} f = \underline{\int_a^b} f$$

If so, we say $f \in R[a, b]$ Riemann Integrable over $[a, b]$.

Remark 7.1. *Why is $\underline{\int_a^b} f$ well-defined? Since f is bounded*

$$L(P, f) = \sum_{i=1}^{N} m_i \Delta x_i \leq \sup_{x \in [a,b]} f(x) \cdot (b - a) < \infty$$

Thus the set

$$\{L(P, f) \mid P \text{ partition of } [a, b]\}$$

has an upper bound. By least upper bound property

$$\underline{\int_a^b} f = \sup\{L(P, f) \mid P \text{ partition of } [a, b]\}$$

exists.

Example 7.1 (Non-example)**.** *Consider $f : [a, b] \to \mathbb{R}$ s.t.*

$$f(x) = \begin{cases} 0 & x \in \mathbb{Q} \\ 1 & x \notin \mathbb{Q} \end{cases}$$

Now take any Partition P, in each subinterval, there is always a rational and an irrational, so

$$M_i = 1 \qquad m_i = 0$$

and thus

$$\begin{aligned} U(P, f) &= b - a \\ L(P, f) &= 0 \end{aligned}$$

They never agree, so $f \notin R$.

Definition 7.3 (Refinement)**.** *Given P^* and P partitions of $[a, b]$. P^* refines P if $P^* \supsetneq P$.*

Remark 7.2. *Given P_1, P_2 partition of $[a, b]$, then $P_1 \cup P_2$ is a common refinement, i.e., it refines both P_1 and P_2.*

Proposition 7.1. *If P^* refines P, then*

$$\begin{aligned} L(P, f) &\leq L(P^*, f) \\ U(P, f) &\geq U(P^*, f) \end{aligned}$$

Proof. We first assume $P^* = P \cup \{x\}$, i.e., we add one point to P. Denote $P = \{x_i\}_{i=0}^N$ with corresponding m_i. Pick i s.t.

$$x_{i-1} < x < x_i$$

Then we define

$$w_1 := \inf_{[x_{i-1}, x]} f(x) \qquad w_2 := \inf_{[x, xi]} f(x)$$

One obtain (infimum over smaller set)

$$\begin{aligned} w_1 &\geq m_i \\ w_2 &\geq m_i \end{aligned}$$

Thus

$$\begin{aligned} w_1(x - x_{i-1}) + w_2(x_i - x) &\geq m_i(x - x_{i-1}) + m_i(x_i - x) \\ &= m_i \Delta x_i \\ L(P \cup \{x\}, f) &\geq L(P, f) \end{aligned}$$

Now we induct on adding k points to P^*. ☐

Remark 7.3. *For $P^* = P_1 \cup P_2$ one has*

$$L(P_2, f) \leq L(P^*, f) \leq U(P^*, f) \leq U(P_1, f)$$

7.1 Criterion for Riemann Integrability

Proposition 7.2 (Integrability Criterion)**.** *$f \in R[a, b]$ iff*

$$\text{for any } \varepsilon > 0, \text{ there exists a partition } P_\varepsilon \text{ s.t.} \qquad U(P_\varepsilon, f) - L(P_\varepsilon, f) < \varepsilon \tag{4}$$

Proof. 1. ($\Longleftarrow$) Since

$$L(P, \varepsilon) \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq U(P, f)$$

for any partition P, then for any $\varepsilon > 0$, there exists P_ε s.t.

$$\overline{\int_a^b} f - \underline{\int_a^b} f \leq U(P_\varepsilon, f) - L(P_\varepsilon, f) \leq \varepsilon$$

But LHS is independent of ε. Take $\varepsilon \to 0$ on RHS to conclude $\overline{\int_a^b} f = \underline{\int_a^b} f$.

2. ($\Longrightarrow$) Using definition of supremum and infimum, for any $\varepsilon > 0$, there exists P_1 and P_2 s.t.

$$U(P_1, f) - \overline{\int_a^b} f < \frac{\varepsilon}{2}$$
$$\underline{\int_a^b} f - L(P_2, f) < \frac{\varepsilon}{2}$$

Since $f \in R[a,b]$, $\overline{\int_a^b} f = \underline{\int_a^b} f$, thus the above translates to

$$U(P_1, f) - L(P_2, f) = U(P_1, f) - \overline{\int_a^b} f + \underline{\int_a^b} f - L(P_2, f) < \varepsilon$$

Take $P_\varepsilon = P_1 \cup P_2$ to conclude.

□

Proposition 7.3. *1. If (4) holds for some $\varepsilon > 0$ and P, then (4) holds for any refinement $P^* \supsetneq P$ with the same ε.*

2. If (4) holds for

$$P = \{x_0, \cdots, x_n\}$$

and s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$, then

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta x_i < \varepsilon \tag{5}$$

3. If $f \in R[a,b]$ and (5) holds, then

$$|\sum_{i=1}^{n} f(t_i)\Delta x_i - \int_a^b f| < \varepsilon$$

Proof. 1. Indeed

$$U(P^*, f) - L(P^*, f) \le U(P, f) - L(P, f) < \varepsilon$$

2. Since both $f(s_i)$, $f(t_i)$ are in $[m_i, M_i]$ one obtain

$$|f(s_i) - f(t_i)| \le M_i - m_i$$

Thus

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta x_i \le \sum_{i=1}^{n} (M_i - m_i)\Delta x_i = U(P, f) - L(P, f) < \varepsilon$$

3. Since $t_i \in [x_{i-1}, x_i]$ s.t. $f(t_i) \in [m_i, M_i]$ one has

$$L(P, f) \le \sum_{i=1}^{N} f(t_i)\Delta x_i \le U(P, f)$$

On the other hand, using $f \in R[a,b]$

$$L(P, f) \le \int_a^b f \le U(P, f)$$

one obtain

$$|\sum_{i=1}^{n} f(t_i)\Delta x_i - \int_a^b f| \le U(P, f) - L(P, f) < \varepsilon$$

□

7.1.1 Theorems for Riemann Integrability

Theorem 7.1 (Continuous implies Integrability). *If $f : [a, b] \to \mathbb{R}$ is continuous, then $f \in R[a, b]$.*

Proof. Since $[a, b]$ is compact, f is continuous, we know f is uniformly continuous over $[a, b]$ using Theorem 5.2. Thus for any $\varepsilon > 0$ one can choose $\delta = \delta(\varepsilon) > 0$ s.t.

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon}{b - a}$$

Now we choose a partition P s.t. $\Delta x_i < \delta$ for any i. One obtain

$$U(P, f) - L(P, f) = \sum_{i=1}^{N} (M_i - m_i)\Delta x_i \le \sum_{i=1}^{N} \frac{\varepsilon}{b - a}\Delta x_i \le \varepsilon$$

Using Proposition 7.2 one obtain $f \in R[a, b]$. □

Proposition 7.4 (Monotone implies Integrability). *If f is monotone on $[a, b]$, then $f \in R[a, b]$.*

Proof. For any $\varepsilon > 0$, let P with $\Delta x_i = \frac{b-a}{n}$ where n is the equal length for n subintervals. WLOG assume f is monotone increasing, then

$$m_i = f(x_{i-1}) \qquad M_i = f(x_i)$$

and so

$$\begin{aligned} U(P, f) - L(P, f) &= \frac{b - a}{n} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \\ &= \frac{b - a}{n}(f(b) - f(a)) \end{aligned}$$

Now one may choose n sufficiently large so that

$$\frac{b - a}{n}(f(b) - f(a)) < \varepsilon$$

Using Proposition 7.2 one obtain $f \in R[a, b]$. □

Theorem 7.2 (Finite Discontinuity implies Integrability). *If $f : [a, b] \to \mathbb{R}$ bounded, has finitely many discontinuities, then $f \in R[a, b]$.*

Proof. For any $\varepsilon > 0$, since f is bounded,

$$M = \sup_{x \in [a,b]} |f(x)| < \infty$$

Denote

$$E = \{x \in [a, b] \mid f \text{ has a discontinuity at } x\}$$

which is finite. Now one can over E with disjoint $[u_j, v_j]$ s.t.

$$\sum_{j=1}^{N} v_j - u_j < \varepsilon$$

for some $N = N(\varepsilon)$, where we require each point in $E \cap (a, b)$ to lie within (u_j, v_j) for some j. Consider the set

$$K := [a, b] \setminus \bigcup_{j=1}^{N} (u_j, v_j)$$

Since K is closed and bounded, by Heine-Borel, K is compact. Notice by definition, f is continuous over K, hence by Theorem 5.2 f is uniformly continuous over K. There exists $\delta = \delta(\varepsilon) > 0$ s.t. for any x, $y \in K$ s.t.

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

How can we form a partition of $[a, b]$? We choose a partition $P = \{x_i\}$ s.t.

1. u_j, $v_j \in P$ for any $j = 1, \cdots, N$
2. no points within (u_j, v_j) occur in P

3. If $x_i \in P$ is not any of the u_j, then we require $\Delta x_i < \delta$.

Now

$$
\begin{aligned}
U(P,f) - L(P,f) = \sum_i (M_i - m_i)\Delta x_i &= \sum_{j=1}^{N} \big(\sup_{[u_j,v_j]} f - \inf_{[u_j,v_j]} f\big)(v_j - u_j) + \sum_{\text{other subintervals}} (M_i - m_i)\Delta x_i \\
&\le 2M \sum_{j=1}^{N} (v_j - u_j) + \varepsilon(b-a) \\
&\le 2M\varepsilon + (b-a)\varepsilon
\end{aligned}
$$

Using arbitrariness of ε, and then Proposition 7.2 to conclude $f \in R[a,b]$. □

Theorem 7.3 (Integrable iff Measure Zero Discontinuities)**.** *Let $f : [a,b] \to \mathbb{R}$ be bounded. Then $f \in R[a,b]$ iff the set E of discontinuities of f has measure zero, i.e., for any $\varepsilon > 0$ one can choose intervals I_j s.t.*

$$
E \subseteq \bigcup_{j\in\mathbb{N}} I_j \qquad \text{and} \qquad \sum_{j\in\mathbb{N}} |I_j| < \varepsilon
$$

Example 7.2. *Consider*

$$
f(x) = \begin{cases} \frac{1}{q} & x = \frac{p}{q} \\ 0 & x \notin \mathbb{Q} \end{cases}
$$

Then $f : [a,b] \to \mathbb{R}$ is continuous over irrationals. We compute

$$
\int_a^b f = 0
$$

Example 7.3. *Consider the Cantor set $C = \bigcup_n I_n$ for a sequence of sets*

$$
I_n \supseteq C \qquad \forall\ n
$$

and $|I_n| \to 0$. Here C is uncountable. Consider function $f : [0,1] \to \mathbb{R}$ that is discontinuous at every $x \in C$ but continuous at every $x \in [0,1] \setminus C$. Then $f \in R[0,1]$ because C has zero measure.

Theorem 7.4. *If $f : [a,b] \to \mathbb{R}$ is integrable, $m \le f(x) \le M$, and*

$$
\phi : [m,M] \to \mathbb{R}
$$

is continuous, then

$$
h(x) = \phi(f(x)) \in R[a,b]
$$

7.2 Properties

Proposition 7.5 (Properties of Riemann Integral)**.** *Let $[a,b] \subseteq \mathbb{R}$.*

1. *f_1, $f_2 \in R[a,b]$ implies $f_1 + f_2 \in R[a,b]$, $cf_1 \in R[a,b]$ and*

$$
\int_a^b (f_1 + f_2) = \int_a^b f_1 + \int_a^b f_2
$$

$$
\int_a^b cf_1 = c\int_a^b f_1 \qquad \forall\ c \in \mathbb{R}
$$

2. *If $f_1(x) \le f_2(x)$ then*

$$
\int_a^b f_1 \le \int_a^b f_2
$$

3. *If $f \in R[a,b]$ and $c \in (a,b)$, then*

$$
\int_a^b f = \int_a^c f + \int_c^b f
$$

4. *If $f \in R[a,b]$ and $|f(x)| \le M$ then*

$$
|\int_a^b f| \le M(b-a)
$$

5. *If f, $g \in R[a,b]$ then $fg \in R[a,b]$.*

6. *If $f \in R[a,b]$, then $|f| \in R[a,b]$ and*

$$
|\int_a^b f| \le \int_a^b |f|
$$

7.3 Fundamental Theorem of Calculus

Theorem 7.5 (FTC)**.** *Let $f \in R[a,b]$ and set the function*

$$F(x) := \int_a^x f(t)dt$$

We call this F the anti-derivative of f. Then

1. *F is continuous over $[a,b]$*
2. *If f is continuous at some $x_0 \in [a,b]$, then F is differentiable at x_0 and*

$$F'(x_0) = f(x_0)$$

This F is leveling up the smoothness of f.

Proof. 1. Since $f \in R[a,b]$, we know that f is bounded. Thus

$$|f(x)| \leq M \qquad \forall\, x \in [a,b]$$

Now we consider the quantity

$$\begin{aligned}|F(x) - F(y)| &= |\int_a^x f(t)dt - \int_a^y f(t)dt| \\ &\leq |\int_x^y f(t)dt| \\ &\leq \int_x^y |f(t)|dt \leq M|y-x|\end{aligned}$$

Thus F is uniformly continuous over $[a,b]$. Why? Given $\varepsilon > 0$, one can now choose $\delta < \frac{\varepsilon}{M}$ s.t.

$$|x-y| \leq \delta \implies |F(x) - F(y)| \leq M|x-y| \leq \varepsilon$$

2. Now we assume that f is continuous at x_0. Given $\varepsilon > 0$, we choose $\delta > 0$ s.t.

$$|f(x) - f(x_0)| < \varepsilon \qquad \forall\, |x - x_0| < \delta$$

Now we look at

$$\begin{aligned}|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)| &= |\frac{1}{x-x_0}\int_{x_0}^x (f(t) - f(x_0))dt| \\ &\leq \frac{1}{|x-x_0|}\int_{x_0}^x |f(t) - f(x_0)|dt \\ &\leq \frac{1}{|x-x_0|}\varepsilon|x-x_0| = \varepsilon \qquad \forall\, |x-x_0| < \delta\end{aligned}$$

Hence we've shown that

$$F'(x_0) = \lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0)$$

□

Theorem 7.6 (FTC II)**.** *If $f \in R$ and there exists F s.t.*

$$F' = f$$

Then

$$\int_a^b f(x)dx = F(b) - F(a)$$

Proof. Let $\varepsilon > 0$. Choose P_ε s.t.

$$U(P,f) - L(P,f) < \varepsilon$$

using $f \in R$. By the Mean Value Theorem Corollary 6.2, there exists $t_i \in (x_{i-1}, x_i)$ s.t.

$$F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i$$

Now using the Riemann sum for $\int_a^b f$ one obtain

$$\sum_{i=1}^n f(t_i)\Delta x_i = \sum_{i=1}^n F(x_{i-1}) - F(x_i) = F(b) - F(a)$$

Thus one obtain

$$|\int_a^b f - \sum_{i=1}^n f(t_i)\Delta x_i| = |\int_a^b f - (F(b) - F(a))| \leq \varepsilon$$

But since this holds for any $\varepsilon > 0$ one has

$$F(b) - F(a) = \int_a^b f(t)dt$$

□

Corollary 7.1 (Integration By Parts)**.** *Suppose $F' = f \in R$ and $G' = g \in R$. Then*

$$\int_a^b Fg = F(b)G(b) - F(a)G(a) - \int_a^b fG$$

Also written as

$$\int_a^b FG' = F(b)G(b) - F(a)G(a) - \int_a^b F'G$$

Proof. Apply FTC Theorem 7.6 to

$$H(x) = F(x)G(x)$$

and recall

$$H'(x) = f(x)G(x) + F(x)g(x)$$

□

Corollary 7.2 (Change of Variables Formula)**.** *Let*

$$\varphi : [A, B] \to [a, b]$$

be differentiable and strictly increasing. And let $f : [a, b] \to \mathbb{R}$ be integrable. Then $f \circ \varphi \in R[A, B]$ and

$$\int_a^b f(x)dx = \int_A^B f(\varphi(y))\varphi'(y)dy$$

Remark 7.4. *This is saying one is looking at the real line $\mathbb{R}$ differently. If φ is nice as above, then one is relabeling. For Q a partition of $[A, B]$, one has a one-to-one correspondence (via φ) with a partition P of $[a, b]$.*

Proof. Let $\varepsilon > 0$. Choose P_ε a partition of $[a, b]$, $P = \{x_i\}_{i=0}^n$ s.t.

$$U(P, f) - L(P, f) < \varepsilon$$

Define the corresponding partition

$$Q = \{y_i = \varphi^{-1}(x_i)\}_{i=0}^n$$

For the original partition one has

$$\sum_{i=1}^n (M_i - m_i)\Delta x_i < \varepsilon$$

We want to compute the Riemann sum on the other side. Consider

$$m_i = \inf_{[x_{i-1}, x_i]} f(x) = \inf_{[y_{i-1}, y_i]} f(\varphi(y))$$
$$M_i = \sup_{[x_{i-1}, x_i]} f(x) = \sup_{[y_{i-1}, y_i]} f(\varphi(y))$$

Look at the Riemann sum

$$\sum_{i=1}^n M_i\Delta x_i = \sum_{i=1}^n M_i(x_i - x_{i-1}) = \sum_{i=1}^n M_i(\varphi(y_i) - \varphi(y_{i-1}))$$
$$= \sum_{i=1}^n M_i\varphi'(t_i)\Delta y_i \qquad t_i \in (y_{i-1}, y_i)$$

The LHS is Riemann sum for $\int_a^b f(x)dx$ while the RHS is Riemann sum for $\int_A^B f \circ \varphi \cdot \varphi'$. Hence

$$|U(f \circ \varphi, Q) - L(f \circ \varphi, Q)| < \varepsilon$$

□

Remark 7.5. *One has several generalization.*

1. *One can generalize to $\mathbb{R}^n$ as integration of differential forms. Stokes Theorem is stated as*

$$\int_M df = \int_{\partial M} f$$

2. *One can generalize to spaces not in $\mathbb{R}^n$.*

3. *One can generalize to Lebesgue Integral. The idea is to assign a length to subsets of $\mathbb{R}$ even for non-intervals. Lebesgue want*

   (a) $\ell([a, b]) := b - a$

   (b) *For intervals I_n disjoint countably many, one want*

$$\sum_n \ell(I_n) = \ell(\bigcup_n I_n)$$

   It follows from here that $\ell(\{a\}) = 0$ and $\ell(\mathbb{Q}) = 0$.

   Consider the function

$$f(x) = \begin{cases} 1 & \mathbb{Q} \\ 0 & \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

   Then $\int f = 0$ under the Lebesgue sense. Notice that Riemann integral doesn't work for this function.

A Assignments

In this section I discuss some problems in the Assignments which I believe to be helpful.

A.1 Assignment 1

Problem A.1 ([Rud76] Exercise 1.5)**.** *Let A be an nonempty set of real numbers which is bounded from below. Let $-A$ be the set of all numbers $-x$ where $x \in A$. Prove that*

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

Solution A.1. *Proof.*

1. We first prove

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

For any $x \in A$, since $\inf(A)$ is a lower bound of A, one has

$$x \geq \inf(A) \iff -x \leq -\inf(A)$$

But this holds for any $-x \in -A$, hence $-\inf(A)$ is an upper bound of the set $-A$. Now by definition of sup as the least upper bound, one has

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

2. We then prove

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

For any $x \in A$, since $\sup(-A)$ is an upper bound of $-A$, one has

$$-x \leq \sup(-A) \iff x \geq -\sup(-A) \qquad \forall\, x \in A$$

But now $-\sup(-A)$ is an lower bound for A. Thus by definition of inf as the greatest lower bound, one has

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

$\square$

Problem A.2 ([Rud76] Exercise 1.8)**.** *Prove that no order can be defined in a complex field $\mathbb{C}$ that turns it into an order field.*

Solution A.2. *Proof.* Assume there is a (total) order $<$ on $\mathbb{C}$. Then by definition of order one has to be able to compare any pair of elements in $\mathbb{C}$. Thus for 0 and i, either one of the following holds

$$i > 0 \qquad \text{or} \qquad i = 0 \qquad \text{or} \qquad i < 0$$

1. In the first case $i > 0$, assume $(\mathbb{C}, <)$ is an order field, then

$$i^2 = i \cdot i = -1 > 0$$

But using assumption $(\mathbb{C}, <)$ is an order field again yields

$$i \cdot (-1) = -i > 0$$

Now one add i on both sides and using $\mathbb{C}$ is a field to see

$$i + (-i) = 0 > i$$

which gives the contrary to $i > 0$.

2. In the second case $i = 0$ then

$$i^2 = -1 = 0$$

Adding 1 on both sides yields

$$0 = 1$$

which says the additive identity is equal to the multiplicative identity, and this leads to a contradiction to $\mathbb{C}$ as a field.

3. In the third case $i < 0$, i.e., $-i > 0$ then
$$(-i)^2 = -1 > 0$$
using assumption $(\mathbb{C}, <)$ is an order field again yields
$$-i \cdot (-1) = i > 0$$
and gives a contrary to $i < 0$.

$\square$

Problem A.3 ([Rud76] Exercise 2.2)**.** *A complex number $z \in \mathbb{C}$ is algebraic if there exist integers $a_0, \cdots, a_n$ (not all zero) s.t.*
$$a_n z^n + a_{n-1} z^{n-1} + a_{n-2} z^{n-2} + \cdots a_0 = 0 \tag{6}$$
Prove that the set of all algebraic numbers is countable.

Solution A.3. *Proof.* We first describe the set of algebraic numbers
$$\begin{aligned} A = \{\text{algebraic numbers}\} &:= \{z \in \mathbb{C} \mid \exists\, n \in \mathbb{N},\, a_0, \cdots, a_n \in \mathbb{Z} \text{ not all zero s.t. (6) holds}\} \\ &= \bigcup_{n=1}^{\infty} \{z \in \mathbb{C} \mid \exists\, a_0, \cdots, a_n \in \mathbb{Z} \text{ with } a_n \neq 0 \text{ the highest order non-zero coefficient s.t. (6) holds}\} \end{aligned}$$
In this case we write
$$A = \bigcup_{n=1}^{\infty} A_n$$
where
$$A_n := \{z \in \mathbb{C} \mid \exists\, a_0, \cdots, a_n \in \mathbb{Z} \text{ with } a_n \neq 0 \text{ the highest order non-zero coefficient s.t. (6) holds}\}$$
In view of [Rud76] Theorem 2.12, it suffices to prove A_n is countable for each n. But note
$$\begin{aligned} A_n &= \bigcup_{|a_0|, |a_1|, \cdots, |a_{n-1}|=0, |a_n|=1}^{\infty} \{z \in \mathbb{C} \mid \text{(6) holds}\} \\ &= \bigcup_{|a_0|, |a_1|, \cdots, |a_{n-1}|=0, |a_n|=1}^{\infty} A_{a_0, a_1, \cdots, a_n} \end{aligned}$$
where
$$A_{a_0, a_1, \cdots, a_n} := \{z \in \mathbb{C} \mid z \text{ is root of the polynomial } a_n z^n + \cdots a_0 = 0\}$$
By Fundamental Theorem of Algebra, any polynomial of order n has at most n roots in $\mathbb{C}$, hence $A_{a_0, a_1, \cdots, a_n}$ has at most n elements. Thus A_n as countable union of finite sets is countable. $\square$

A.2 Assignment 2

Problem A.4 ([Rud76] Exercise 2.5)**.** *Construct a bounded set of real numbers with exactly three limit points.*

Solution A.4. *Take three distinct points in* $\mathbb{R}$*, then construct a limiting sequence that approaches each point. For example, denote*

$$E_0 := \{0, 1, 2\}$$

and

$$E_1 := \{a + \frac{1}{n} \mid a \in E,\ n \in \mathbb{N},\ n \geq 2\}$$

Then E_1 *is bounded because we can fit it in a large enough ball.* E_0 *is the set of limit points of* E_1*. Why? For example choose* $a = 0$*, then for any* $(-\varepsilon, \varepsilon)$ *neighborhood around* 0*, there exists* $N = N(\varepsilon) \in \mathbb{N}$ *s.t.*

$$(-\frac{1}{N}, \frac{1}{N}) \subseteq (-\varepsilon, \varepsilon)$$

so for any $n \geq N$*,* $0 \neq \frac{1}{n} \in (-\varepsilon, \varepsilon)$*. Since* $\varepsilon > 0$ *is arbitrary,* 0 *is a limit point of* E_1*. Same for* 1 *and* 2*.*

Problem A.5 ([Rud76] Exercise 2.7)**.** *Let* $A_1, \cdots$ *be subsets of a metric space.*

1. *If* $B_n = \bigcup_{t=1}^n A_t$*, then*
$$\overline{B_n} = \bigcup_{t=1}^n \overline{A_t} \qquad \forall\ n \in \mathbb{N}$$
2. *If* $B = \bigcup_{t=1}^\infty A_t$*, then*
$$\overline{B} \supseteq \bigcup_{t=1}^\infty \overline{A_t}$$
Show that the inclusion can be proper.

Solution A.5. *Proof.* 1. (a) First prove

$$\overline{B_n} \subseteq \bigcup_{t=1}^n \overline{A_t}$$

Since $B_n = \bigcup_{t=1}^n A_t \subseteq \bigcup_{t=1}^n \overline{A_t}$, it suffices to prove for the set of limit points of B_n. Take any q limit point of $\bigcup_{t=1}^n A_t$, then for any $\varepsilon > 0$, there exists $p \in \bigcup_{t=1}^n A_t$ s.t. $d(p, q) < \varepsilon$. By definition of $\bigcup_{t=1}^n A_t$, there exists $t_0 \in \{1, \cdots, n\}$ s.t. $p \in A_{t_0}$. Can we now conclude $q \in \overline{A_{t_0}}$? No! Because the choice of t_0 depends on $p = p(\varepsilon) \in B_n$, which further depends on ε, and as one shrink ε, there is no guarantee that t_0 won't jump to another $t \in \{1, \cdots, n\} \setminus \{t_0\}$. To prove this inclusion, one wish to remove the dependence of t_0 on ε, and at this step we're stuck! Let's alternatively argue by contradiction. Assume

$$q \notin \bigcup_{t=1}^n \overline{A_t}$$

Then for any $t \in \{1, \cdots, n\}$, q is not a limit point of A_t, i.e., there exists $\varepsilon_t > 0$ s.t.

$$B_{\varepsilon_t}(q) \cap A_t = \varnothing \text{ or } \{q\}$$

Thus define

$$\varepsilon := \min\{\varepsilon_t \mid t = 1, \cdots, n\} > 0$$

Note $\varepsilon > 0$ makes use of finiteness! Otherwise taking inf could lead to $\varepsilon = 0$. We observe that

$$B_\varepsilon(q) \cap A_t = \varnothing \text{ or } \{q\} \qquad \forall\ t \implies B_\varepsilon(q) \cap \left(\bigcup_{t=1}^n A_t\right) = \varnothing \text{ or } \{q\}$$

But this is to say q is not limit point of $B_n = \bigcup_{t=1}^n A_t$, a contradiction to our assumption.

(b) Now we prove

$$\overline{B_n} \supseteq \bigcup_{t=1}^n \overline{A_t}$$

For any $q \in \bigcup_{t=1}^n \overline{A_t}$, there exists $t_0 \in \{1, \cdots, n\}$ s.t. $q \in \overline{A_{t_0}}$, hence for any $\varepsilon > 0$, there exists $p \in A_{t_0}$, $q \neq p$, s.t. $d(p, q) < \varepsilon$. But

$$p \in A_{t_0} \subseteq \bigcup_{t=1}^n A_t = B_n$$

Here we're good, as t_0 is fixed before $\varepsilon > 0$. By definition of limit point, this is to say $q \in \overline{B_n}$.

2. (a) We first prove the inclusion. The reasoning is exactly the same, since $q \in \bigcup_{t=1}^{\infty} \overline{A_t}$ yields there exists $t_0 \in \mathbb{N}^*$ s.t. $q \in \overline{A_{t_0}}$, hence for any $\varepsilon > 0$, there exists $p \in A_{t_0}$, $q \neq p$ s.t. $d(p,q) < \varepsilon$, but

$$p \in A_{t_0} \subseteq \bigcup_{t=1}^{\infty} A_t = B$$

This is to say $q \in \overline{B}$.

(b) For the reverse, we provide a counter-example. Consider

$$A_n := (-1 + \frac{1}{n}, 1 - \frac{1}{n}) \qquad \forall\, n \in \mathbb{N}^*$$

Then

$$\overline{A_n} = [-1 + \frac{1}{n}, 1 - \frac{1}{n}] \qquad \forall\, n \in \mathbb{N}^*$$

Thus

$$\bigcup_{n=1}^{\infty} \overline{A_n} = (-1, 1)$$

because the two endpoints ± 1 are never reached. Now on the other hand

$$B = \bigcup_{n=1}^{\infty} A_n = (-1, 1)$$

$$\overline{B} = [-1, 1] \supsetneq \bigcup_{n=1}^{\infty} \overline{A_n}$$

□

Problem A.6 ([Rud76] Exercise 2.8)**.** *Is every point of every open set $E \subseteq \mathbb{R}^2$ a limit point of E? What about closed sets?*

Solution A.6. *1. Yes, every point $q = (x_1, x_2) \in E \subseteq \mathbb{R}^2$ open is a limit point of E. Since E is open, (x_1, x_2) is an interior point, thus there exists $r_0 > 0$ s.t.*

$$B_{r_0}((x_1, x_2)) \subseteq E \iff \forall\ (y_1, y_2)\ \text{s.t.}\ \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} < r_0\ \text{belongs to}\ E$$

Is there such (y_1, y_2)? Indeed, take

$$(y_1, y_2) := (x_1 + \frac{r_0}{2}, x_2 + \frac{r_0}{2})$$

Then

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = \sqrt{(\frac{r_0}{2})^2 + (\frac{r_0}{2})^2} < r_0$$

But this works for any $r \leq r_0$ arbitrary. For $r > r_0$, we simply take its intersection with $B_{r_0}((x_1, x_2))$ and construct (y_1, y_2) as above. In either case (x_1, x_2) is a limit point of E.

2. No, not every point of a closed set is a limit point. (Notice closed sets are defined to include all its limit points, but not all its points are its limit points!). Consider the singleton set

$$\{(0, 0)\} \subseteq \mathbb{R}^2$$

Then $\{(0, 0)\}$ is closed as it does not have any limit point. In particular $(0, 0)$ is not a limit point of $\{(0, 0)\}$.

Problem A.7 ([Rud76] Exercise 2.9)**.** *Let E^o denote the set of interior points of E. Prove that*

(a) E^o is open.

(b) E is open iff $E^o = E$.

(c) If $G \subseteq E$ and G is open, then $G \subseteq E^o$.

(d) $(E^o)^c = \overline{E^c}$.

(e) Do E and $\overline{E}$ always have the same interior?

(f) *Do E and E^o always have the same closure?*

Solution A.7. *Proof.* (a) By definition E is open if all its points are interior points of E,i.e. $E \subseteq E^o$. To show E^o is open, one need to show $E^o \subseteq (E^o)^o$. This is done by showing

$$(E^o)^o = E^o$$

For $\subseteq$ direction, for any $x \in (E^o)^o$, there exists $r > 0$ s.t. $B_r(x) \subseteq E^o$, thus indeed $x \in E^o$. For $\supseteq$ direction, for any $x \in E^o$, there exists $r > 0$ s.t. $B_r(x) \subseteq E$. We claim that in fact,

$$B_r(x) \subseteq E^o$$

so that $x \in (E^o)^o$ follows. Now for any $y \in B_r(x)$, $d(x, y) < r$, so take the ball

$$B_{r-d(x,y)}(y) \subseteq B_r(x) \subseteq E$$

we find a ball centered at y with radius $r - d(x, y) > 0$ that locates in E, i.e., $y \in E^o$. But $y \in B_r(x)$ is arbitrary, hence $B_r(x) \subseteq E^o$.

(b) (i) ($\Longrightarrow$) If E is open, the set $E \subseteq E^o$. It suffices to prove $E^o \subseteq E$. But for any $x \in E^o$, there exists $r > 0$ s.t. $B_r(x) \subseteq E$, in particular, $x \in E$.

(ii) ($\Longleftarrow$) If $E = E^o$, indeed $E \subseteq E^o$, i.e., E is open.

(c) If G is open, for any $x \in G$, there exists $r > 0$ s.t. $B_r(x) \subseteq G \subseteq E$, hence $x \in E^o$. Since x is arbitrary, $G \subseteq E^o$.

(d) (i) For $\subseteq$ direction, for any $x \notin E^o$, for any $r > 0$, there exists $y \notin E \iff y \in E^c$ s.t. $d(x, y) < r$. If there exists $r > 0$ s.t. $y = x$, then indeed $x \in E^c \subseteq \overline{E^c}$. If $y \neq x$ for any $r > 0$, then $x \in \overline{E^c}$ by definition of a limit point.

(ii) For $\supseteq$ direction, for any $x \in \overline{E^c}$, i.e., for any $r > 0$, there exists $y \neq x$, $y \in E^c$ s.t. $d(y, x) < r$, but this is to say $B_r(x) \nsubseteq E$. Thus $x \notin E^o \iff x \in (E^o)^c$.

(e) No. Let $E = \mathbb{Q}$, then $\overline{E} = \mathbb{R}$. Unfortunately $E^o = \varnothing$ but $(\overline{E})^o = \mathbb{R}$.

(f) No. Again let $E = \mathbb{Q}$, then $E^o = \varnothing$, so $\overline{E} = \mathbb{R}$ but $\overline{E^o} = \varnothing$.

□

Problem A.8 ([Rud76] Exercise 2.10). *Let X be an infinite set. For $p, q \in X$ define*

$$d(p, q) := \begin{cases} 1 & p \neq q \\ 0 & p = q \end{cases}$$

Prove that this is a metric. Which subsets of this metric space are closed, which are open, and which are compact?

Solution A.8. *That d is a metric follows directly by verifying the three defining properties of a metric.*

1. *Let $x \in X$. Then for any $r < 1$,*

$$B_r(x) := \{y \in X \mid d(x, y) < r\} = \{x\} \subseteq \{x\}$$

Hence indeed a radius, say $r = \frac{1}{2}$ exists s.t. $B_{\frac{1}{2}}(x) \subseteq \{x\}$. Thus each singleton $\{x\} \subseteq X$ is open. Using Theorem 2.1, openness is invariant under arbitrary unions. Since any subset $Y \subseteq X$ can be written as union of singletons, i.e.,

$$Y = \bigcup_{y \in Y} \{y\}$$

All sets in X are open under this metric d. But then using Duality Proposition 2.5, the complements of all sets are closed. But they're the same collection of sets. Hence each set is both open and closed.

2. *Only finite subsets of X are compact. For infinite sets, each the covering to be the collection of all elements in the set, then this does not have any finite subcover, hence it is not compact. For any finite set, the worst open cover one can take is in fact the collection of all its elements, which is finite.*

Problem A.9 ([Rud76] Exercise 2.12). *Let*

$$K := \{0\} \cup \{\frac{1}{n} \mid n \in \mathbb{N}^*\} \subseteq \mathbb{R}$$

Prove K is compact directly from the definition.

Solution A.9. *Proof.* For any $\{G_\alpha\}$ open cover of K

$$K = \{0\} \cup \{\frac{1}{n} \mid n \geq 1, n \in \mathbb{N}\} \subseteq \bigcup_{\alpha} G_\alpha$$

Since the union covers 0, there exists G_{α_0} s.t.

$$0 \in G_{\alpha_0}$$

Now using G_{α_0} is open, there exists $\delta > 0$ s.t.

$$(-\delta, \delta) \subseteq G_{\alpha_0}$$

Now there exists $N = N(\delta) \in \mathbb{N}$ by Theorem 1.2 s.t.

$$\frac{1}{N} < \delta$$

and so for any $n \geq N$,

$$\frac{1}{n} \in (0, \frac{1}{N}) \subseteq (-\delta, \delta) \subseteq G_{\alpha_0}$$

Hence we can cover the part close to 0 using merely G_{α_0}

$$\{0\} \cup \{\frac{1}{n} \mid n \geq N\} \subseteq G_{\alpha_0}$$

Now for $1 \leq i \leq N-1$, take each open set which we label G_{α_i} s.t.

$$\frac{1}{i} \in G_{\alpha_i} \qquad \forall\, 1 \leq i \leq N-1$$

Thus we've constructed a finite subcover

$$K = \{0\} \cup \{\frac{1}{n} \mid n \geq N\} \cup \{\frac{1}{n} \mid 1 \leq n \leq N-1\} \subseteq G_{\alpha_0} \cup \bigcup_{i=1}^{N-1} G_{\alpha_i}$$

$\square$

Problem A.10 ([Rud76] Exercise 2.16)**.** *Regard $\mathbb{Q}$ the set of rational numbers as metric space with metric*

$$d(p, q) := |p - q|$$

Let

$$E := \{p \in \mathbb{Q} \mid 2 < p^2 < 3\}$$

Show that E is closed and bounded in $\mathbb{Q}$, but not compact. Is E open in $\mathbb{Q}$?

Solution A.10. *Proof.* 1. That E is bounded follows by fitting E inside a large enough ball. To prove E is closed in $\mathbb{Q}$, we prove $\mathbb{Q} \setminus E$ is open in $\mathbb{Q}$. For any $p \in \mathbb{Q} \setminus E$, we know either

$$p^2 \leq 2 \qquad \text{or} \qquad p^2 \geq 3$$

Since there is no $p \in \mathbb{Q}$ s.t. $p^2 = 2$ or $p^2 = 3$, $p \in \mathbb{Q} \setminus E$ is equivalent to

$$p^2 < 2 \qquad \text{or} \qquad p^2 > 3$$

Now fix $p^2 < 2$, and construct $r > 0$ s.t. (we cheat a bit by using irrationals)

$$\begin{aligned}
(p+r)^2 &< 2 \\
r^2 + 2rp + p^2 - 2 &< 0 \\
r &\in (\frac{-2p - \sqrt{4p^2 - 4p^2 + 8}}{2}, \frac{-2p + \sqrt{4p^2 - 4p^2 + 8}}{2}) = (-p - \sqrt{2}, -p + \sqrt{2}) \\
(p-r)^2 &< 2 \\
r &\in (p - \sqrt{2}, p + \sqrt{2})
\end{aligned}$$

Notice

$$\begin{aligned}
p + \sqrt{2} > -p - \sqrt{2} &\iff p > -\sqrt{2} \\
p - \sqrt{2} < -p - \sqrt{2} &\iff p < 0 \\
-p + \sqrt{2} < p + \sqrt{2} &\iff p > 0 \\
-p + \sqrt{2} > p - \sqrt{2} &\iff p < \sqrt{2}
\end{aligned}$$

Hence

$$(-p-\sqrt{2},-p+\sqrt{2})\cap(p-\sqrt{2},p+\sqrt{2})=\begin{cases}(p-\sqrt{2},-p+\sqrt{2}) & \text{if } p>0 \text{ and } p^2<2\\ (-p-\sqrt{2},p+\sqrt{2}) & \text{if } p<0 \text{ and } p^2<2\\ (-\sqrt{2},\sqrt{2}) & \text{if } p=0\end{cases}$$

For each p classified as above, pick $r\in\mathbb{Q}$ that lies within the interval $(-p-\sqrt{2},-p+\sqrt{2})\cap(p-\sqrt{2},p+\sqrt{2})$ to ensure

$$(p-r,p+r)\subseteq\mathbb{Q}\setminus E$$

Similarly for $p^2>3$. This shows that $\mathbb{Q}\setminus E$ is open in $\mathbb{Q}$, hence E is closed in $\mathbb{Q}$.

2. For the same reason as above, E is open in $\mathbb{Q}$. For any $p\in E$, we want to construct $r>0$ s.t. $(p-r,p+r)\subseteq E$, i.e. (WLOG we do for $p>0$)

$$\begin{aligned}(p-r)^2&>2\\ r^2-2pr+p^2-2&>0\\ r&\in(-\infty,\frac{2p-\sqrt{4p^2-4p^2+8}}{2})\cup(\frac{2p+\sqrt{4p^2-4p^2+8}}{2},\infty)=(-\infty,p-\sqrt{2})\cup(p+\sqrt{2},\infty)=:A_1\\ (p+r)^2&<3\\ r^2+2pr+p^2-3&<0\\ r&\in(\frac{-2p-\sqrt{4p^2-4p^2+12}}{2},\frac{-2p+\sqrt{4p^2-4p^2+12}}{2})=(-p-\sqrt{3},-p+\sqrt{3})=:A_2\end{aligned}$$

Then

$$A_1\cap A_2=\begin{cases}(-p-\sqrt{3},-p+\sqrt{3}) & \text{if } p^2>2,\ p^2<3 \text{ and } p>\frac{\sqrt{3}+\sqrt{2}}{2}\\ (-p-\sqrt{3},p-\sqrt{2}) & \text{if } p^2>2,\ p^2<3 \text{ and } p<\frac{\sqrt{3}+\sqrt{2}}{2}\\ (-\frac{3\sqrt{3}+\sqrt{2}}{2},\frac{\sqrt{3}-\sqrt{2}}{2}) & \text{if } p=\frac{\sqrt{3}+\sqrt{2}}{2}\end{cases}$$

Pick $r\in\mathbb{Q}\cap A_1\cap A_2$ for each classification of p. Do the same for $p<0$ yields E is open in $\mathbb{Q}$.

3. We show E is not compact. Indeed, consider open cover

$$E_n:=\{p\in\mathbb{Q}\mid 2+\frac{1}{n}<p^2<3-\frac{1}{n}\}$$

Then each E_n is open (same as above) and an open cover of E

$$E=\bigcup_{n=1}^{\infty}E_n$$

But not finite subcover exists for this covering. Hence E is not compact.

□

Problem A.11 ([Rud76] Exercise 2.22)**.** *A metric space is called separable if it contains a countable dense subset. Show that $\mathbb{R}^k$ is countable.*

Solution A.11. *Proof.* Notice we identify $(\mathbb{R}^k,d_2)$ with Euclidean metric. We claim that $\mathbb{Q}^k$ is the countable dense subset of $\mathbb{R}^k$. $\mathbb{Q}^k$ is countable follows from Proposition 2.1. To see $\mathbb{Q}^k$ is dense in $\mathbb{R}^k$, for any

$$x=(x_1,\cdots,x_k)\in\mathbb{R}^k$$

and for any $r_i>0$ with $i=1,\cdots,k$, by Density Property Theorem 1.3, there exists

$$y_i\in(x_i-r_i,x_i+r_i)\cap\mathbb{Q}\qquad\forall\ i=1,\cdots,k$$

Thus for any $r>0$, one can construct $(r_1,\cdots,r_k)$ and $y=(y_1,\cdots,y_k)\in\mathbb{Q}^k$ s.t.

$$d_2(x,y)=\sqrt{\sum_{i=1}^{k}(x_i-y_i)^2}\le\sqrt{\sum_{i=1}^{k}r_i^2}\le r$$

where the last inequality is due our choice of $(r_1,\cdots,r_k)$. For example we can pick

$$r_i:=\frac{r}{\sqrt{k}}\qquad\forall\ i=1,\cdots,k$$

Thus any point in $\mathbb{R}^k$ is a limit point of $\mathbb{Q}^k$, and since $\mathbb{Q}^k\subseteq\mathbb{R}^k$, one has $\overline{\mathbb{Q}^k}=\mathbb{R}^k$. □

Problem A.12 ([Rud76] Exercise 2.23)**.** *A collection $\{V_\alpha\}$ of open subset of X is a base for X if*

$$\text{for every } x \in X \text{ and every open set } G \subseteq X \text{ s.t. } x \in G, \text{ we have } x \in V_\alpha \subseteq G \text{ for some } \alpha$$

i.e. every open set in X is a union of subcollections of V_α. Prove that every separable metric space has a countable base.

Solution A.12. *Proof.* Since X is separable, it has a countable dense subset Y. Consider the collection of sets

$$\mathcal{A} := \{B_q(y) \mid y \in Y,\ q \in \mathbb{Q}\}$$

Then $\mathcal{A}$ is countable as indexed by $\mathbb{Q} \times Y$. We claim that $\mathcal{A}$ is a base for X. Now for any $x \in X$ and $G \subseteq X$ open s.t. $x \in G$, there exists $r > 0$ s.t. $x \in B_r(x) \subseteq G$.

1. If $x \in Y$ then one can take $q = r \in \mathbb{Q}$ small enough s.t. $x \in B_r(x) = B_q(y)$.
2. If $x \notin Y$, then using Y is dense subset in X, for $r > 0$ as above, there exists $y \in Y$ s.t.

   $$d(x, y) < \frac{r}{2}$$

   Thus by Density Theorem 1.3, there exists $q \in \mathbb{Q}$ and $d(x, y) < q < \frac{r}{2}$ so that

   $$d(x, y) < q \implies x \in B_q(y) \subseteq B_r(x)$$

Hence $\mathcal{A}$ is a base for X. □

Problem A.13 ([Rud76] Exercise 2.25)**.** *Prove that every compact metric space K has a countable base, and that K is therefore separable.*

Solution A.13. *Proof.*

1. We first prove K has a countable base. For any $n \in \mathbb{N}$, one has open covering of K

   $$K = \bigcup_{x \in K} B_{\frac{1}{n}}(x)$$

   Now using K is compact there exists finitely-many $N = N(n)$ subcover $\{B_{\frac{1}{n}}(x_{n_j})\}_{j=1}^{N}$ s.t.

   $$K = \bigcup_{j=1}^{N(n)} B_{\frac{1}{n}}(x_{n_j})$$

   Now the collection

   $$\mathcal{A} := \{B_{\frac{1}{n}}(x_{n_j}) \subseteq K \mid n \in \mathbb{N},\ j = 1, \cdots, N(n)\}$$

   is countable. To see $\mathcal{A}$ is indeed a base, for any $x \in K$ and every open set $G \subseteq K$ s.t. $x \in G$, there exists $r > 0$ s.t. $B_r(x) \subseteq G \subseteq K$. For n sufficiently large s.t.

   $$\frac{1}{n} < \frac{r}{2}$$

   there exists $N = N(n)$ and n_j for $j \in \{1, \cdots, N(n)\}$ s.t.

   $$x \in B_{\frac{1}{n}}(x_{n_j}) \iff d(x, x_{n_j}) < \frac{1}{n}$$

   Now for any $y \in B_{\frac{1}{n}}(x_{n_j})$

   $$d(y, x) \le d(y, x_{n_j}) + d(x_{n_j}, x) \le \frac{1}{n} + \frac{1}{n} < r \implies y \in B_r(x)$$

   Thus

   $$x \in B_{\frac{1}{n}}(x_{n_j}) \subseteq B_r(x) \subseteq G$$

   So $\mathcal{A}$ is a countable base for K.
2. Then we argue K is separable. Define

   $$\mathcal{B} := \{x_{n_j} \in K \mid n \in \mathbb{N},\ j = 1, \cdots, N(n)\}$$

   This is indeed countable. To see $\mathcal{B}$ is dense, we make use that $\mathcal{A}$ is a base for K. For any $x \in K$, and for any $r > 0$, there exists $n \in \mathbb{N}$ and x_{n_j} s.t.

   $$d(x, x_{n_j}) < \frac{1}{n}, \qquad \frac{1}{n} < r$$

   so that

   $$x \in B_r(x_{n_j})$$

   Hence $\mathcal{B}$ is dense subset of K.

□

Problem A.14 ([Rud76] Exercise 2.29; [Fol99] Proposition 0.21). *Prove that every open set in $\mathbb{R}$ is the union of an at most countable collection of disjoint segments.*

Solution A.14. *Proof.*

1. For any $U \subseteq \mathbb{R}$ open, for any $x \in U$, define the collection of all open intervals in U that contains x as
$$\mathcal{J}_x := \{I \text{ open interval} \subseteq U \mid x \in I\}$$
Define the maximal interval as the union of all such open intervals around x.
$$J_x := \bigcup_{I \in \mathcal{J}_x} I$$
This is indeed an open interval, and in fact the largest element in $\mathcal{J}_x$.

2. For any x, $y \in U$, either
$$J_x = J_y \qquad \text{or} \qquad J_x \cap J_y = \varnothing$$
Otherwise $J_x \cup J_y$ is a bigger open interval around both x and y, which contradicts maximality. Thus the collection
$$\mathcal{J} := \{J_x \mid x \in U\}$$
is a collection of disjoint open intervals. Moreover
$$U = \bigcup_{J \in \mathcal{J}} J$$

3. It suffices to prove $\mathcal{J}$ is at most countable. We define a clever map as follows.
$$f : \mathcal{J} \to \mathbb{Q} \qquad J \mapsto f(J) := \text{some rational number in } J$$
Since $f(J) \in J$ for any $J \in \mathcal{J}$, and all intervals J in $\mathcal{J}$ are disjoint, the map f is injective. Thus f is a bijection from $\mathcal{J}$ onto $f(\mathcal{J})$, and since
$$\mathcal{J} \sim f(\mathcal{J}) \subseteq \mathbb{Q}$$
by Proposition 2.2, $\mathcal{J}$ is at most countable.

$\square$

A.3 Assignment 4

Problem A.15 ([Rud76] Exercise 5.2)**.** *$f'(x) > 0$ in (a, b). Prove f is strictly increasing in (a, b). Denote g as inverse of f. Prove that g is differentiable and*

$$g'(f(x)) = \frac{1}{f'(x)} \qquad \forall\ a < x < b$$

Solution A.15. *Proof.*

1. For any $x \in (a, b)$, there exists $\varepsilon_x > 0$ s.t.

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \geq \varepsilon_x > 0$$

Hence there exists δ_x s.t. for $|h| < \delta_x$ one has

$$f(x+h) - f(x) \geq h\varepsilon_x > 0$$

hence f is strictly increasing over $(x, x + \delta_x)$. Since $x \in (a, b)$ is arbitrary, f is strictly increasing over (a, b).

2. We denote $f(x + k) = f(x) + h$. Using continuity of f at x, note as $h \to 0$, $k \to 0$. Thus

$$\frac{g(f(x)+h) - g(f(x))}{h} = \frac{g(f(x+k)) - g(f(x))}{f(x+k) - f(x)} = \frac{x+k-x}{f(x+k) - f(x)} = \frac{k}{f(x+k) - f(x)}$$
$$g'(f(x)) = \lim_{h \to 0} \frac{g(f(x)+h) - g(f(x))}{h} = \lim_{k \to 0} \frac{k}{f(x+k) - f(x)} = \frac{1}{f'(x)}$$

Since $f'(x) > 0$, g is differentiable over the range of f.

□

Problem A.16 ([Rud76] Exercise 5.3)**.** *Let g be real function on $\mathbb{R}$ with bounded derivative ($|g'(x)| \leq M$). For any $\varepsilon > 0$ define*

$$f(x) = x + \varepsilon g(x)$$

Show that f is one-to-one for ε small.

Solution A.16. *Proof.* Notice

$$f'(x) = 1 + \varepsilon g'(x) \geq 1 - \varepsilon |g'(x)| \geq 1 - \varepsilon M > 0$$

if we choose ε small so that $\varepsilon < \frac{1}{M}$. Now using Problem A.15, the inverse of f is well-defined over the range of f, and is differentiable. Using well-definedness of the inverse, and note f maps $\mathbb{R}$ to $\mathbb{R}$, we know its inverse is one-to-one from $\mathbb{R}$ to $\mathbb{R}$ as well. □

Problem A.17 ([Rud76] Exercise 5.4)**.** *If real constant $C_0, \cdots, C_n$ satisfies*

$$C_0 + \frac{C_1}{2} + \cdots + \frac{C_{n-1}}{n} + \frac{C_n}{n+1} = 0$$

Then the equation

$$C_0 + C_1 x + \cdots + C_{n-1}x^{n-1} + C_n x^n = 0$$

has at least one root between 0 and 1.

Solution A.17. *Proof.* We construct the polynomial

$$p(x) = C_0 x + \frac{C_1}{2}x^2 + \cdots + \frac{C_{n-1}}{n}x^n + \frac{C_n}{n+1}x^{n+1}$$

Immediately observe $p(0) = 0$. Using assumption we also know $p(1) = 0$. Now using Rolle's Mean Value Theorem Corollary 6.1 we know there exists $c \in (0, 1)$ s.t.

$$p'(c) = 0$$

But what is $p'(c)$? It is

$$C_0 + C_1 c + \cdots + C_{n-1}c^{n-1} + C_n c^n = 0$$

Hence $x = c$ is a root between 0 and 1. □

Problem A.18 ([Rud76] Exercise 5.5). *f is defined and differentiable for every $x > 0$ and $f'(x) \to 0$ as $x \to \infty$. Let $g(x) = f(x+1) - f(x)$. Then*

$$g(x) \to 0 \qquad x \to \infty$$

Solution A.18. *Proof.* Notice

$$|g'(x)| = |f'(x+1) - f'(x)| \leq |f'(x+1)| + |f'(x)| \to 0 \qquad x \to \infty$$

Thus every sequential limit $y_n \to \infty$ necessarily has $g'(y_n) \to 0$. We consider any sequence $x_n \to \infty$. Then using Lagrange Mean Value Theorem Corollary 6.2 there exists $y_n \in (x_n, x_n + 1)$ s.t.

$$g(x_n) = f(x_n + 1) - f(x_n) = f'(y_n)$$

Since $f'(y_n) \to 0$ as $n \to \infty$, necessarily $g(x_n) \to 0$. But x_n is any sequence, hence by sequential characterisation Proposition 5.1 one has $g(x) \to 0$ as $x \to \infty$. $\square$

Problem A.19 ([Rud76] Exercise 5.6). *Assume f continuous for $x \geq 0$, $f'(x)$ exists for $x > 0$, $f(0) = 0$ and f' is monotonically increasing. Then*

$$g(x) = \frac{f(x)}{x} \qquad x > 0$$

is monotonically increasing

Solution A.19. *Proof.* We compute

$$g'(x) = \frac{f'(x)x - f(x)}{x^2}$$

It suffices to show

$$f'(x)x - f(x) \geq 0 \qquad \forall\, x > 0$$

But by Lagrange Mean Value Theorem Corollary 6.2, for any $x > 0$, there exists $c_x \in (0, x)$ s.t.

$$f(x) - f(0) = f(x) = f'(c_x)x \leq f'(x)x$$

$\square$

Problem A.20 ([Rud76] Exercise 5.9). *Let f be continuous real function over $\mathbb{R}$. Assume $f'(x)$ exists for all $x \neq 0$ and that $f'(x) \to 3$ as $x \to 0$. Does it follow that $f'(0)$ exists?*

Solution A.20. *Yes. What is $f'(0)$ by definition? It is*

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h}$$

Now for any h, by Mean Value Theorem Corollary 6.2 there exists $c_h \in (0, h)$ s.t.

$$\frac{f(h) - f(0)}{h} = f'(c_h)$$

But notice $f'(x) \to 3$ as $x \to 0$ hence

$$\lim_{h \to 0} f'(c_h) = 3$$

Thus one has

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} f'(c_h) = 3$$

Problem A.21 ([Rud76] Exercise 5.12). *If $f(x) = |x|^3$. Compute $f'(x)$, $f''(x)$ for all $x \in \mathbb{R}$, and show $f^{(3)}(0)$ does not exist.*

Solution A.21. *For any $x > 0$*

$$\begin{aligned} f'(x) &= 3x^2 \\ f''(x) &= 6x \\ f^{(3)}(x) &= 6 \end{aligned}$$

For any $x < 0$

$$\begin{aligned} f'(x) &= -3x^2 \\ f''(x) &= -6x \\ f^{(3)}(x) &= -6 \end{aligned}$$

Now we use Problem A.20 upon observing

$$f'(x),\ f''(x) \to 0 \qquad x \to 0$$

and indeed f, f' are continuous reals over $\mathbb{R}$. Thus

$$f'(0) = f''(0) = 0$$

But $f^{(3)}(x)$ as $x \to 0$ as a jump discontinuity, hence $f^{(3)}(0)$ does not exist.

Problem A.22 ([Rud76] Exercise 5.15)**.** *Let $a \in \mathbb{R}$, f be twice differentiable real-valued function over (a, ∞), and define*

$$M_0 := \sup_{(a,\infty)} |f(x)| \qquad M_1 := \sup_{(a,\infty)} |f'(x)| \qquad M_2 := \sup_{(a,\infty)} |f''(x)|$$

Show that

$$M_1^2 \leq 4M_0M_2$$

Solution A.22. *Proof.* It suffices to prove for both M_0 and M_2 finite, and that

$$|f'(x)| \leq 2\sqrt{M_0M_2} \qquad \forall\ x \in (a, \infty)$$

Also note if $M_0 = 0$ then everything boils down to 0. On the other hand if $M_2 = 0$ then f' is constant, say $f'(x) = c$ for any $x \in (a, \infty)$. Then necessarily $f(x) = cx + d$ for some $d \in \mathbb{R}$. But by our assumption $M_0 < \infty$, then necessarily $c = 0$, and again everything boils down to 0.

Hence it suffices to prove for both $0 < M_0, M_2 < \infty$.

For any $x > a$, apply Taylor's Theorem (3) to $(x, x + 2h)$ so that there exists $\xi \in (x, x + 2h)$ s.t.

$$f(x + 2h) = f(x) + f'(x)(2h) + \frac{f''(\xi)}{2}4h^2$$

Thus

$$\begin{aligned} f'(x) &= \frac{f(x + 2h) - f(x)}{2h} - hf''(\xi) \\ &\leq \frac{1}{2h}2M_0 + hM_2 = \frac{M_0}{h} + hM_2 \end{aligned}$$

Notice h is free for us to choose. We pick

$$h = \sqrt{\frac{M_0}{M_2}}$$

so that

$$\frac{M_0}{h} + hM_2 = 2\sqrt{M_0M_2}$$

and our result follows.

Now to see $M_1^2 = 4M_0M_2$ can actually happen, consider the example for $a = -1$

$$f(x) = \begin{cases} 2x^2 - 1 & -1 < x < 0 \\ \frac{x^2-1}{x^2+1} & 0 \leq x \end{cases}$$

and compute that $M_0 = 1$, $M_1 = 4$, $M_2 = 4$. □

Problem A.23 ([Rud76] Exercise 5.17)**.** *Suppose f is real, three-times differentiable function on $[-1, 1]$ s.t.*

$$f(-1) = 0 \qquad f(0) = 0 \qquad f(1) = 1 \qquad f'(0) = 0$$

Show that there exists some $x \in (-1, 1)$ s.t.

$$f^{(3)}(x) \geq 3$$

Solution A.23. *Proof.* We apply Taylor (3) to $(0, 1)$ and $(-1, 0)$ at 0 respectively so that there exists $s \in (0, 1)$ and $t \in (-1, 0)$ s.t.

$$\begin{aligned} f(1) &= f(0) + f'(0) + \frac{f''(0)}{2} + \frac{f^{(3)}(s)}{6} \\ f(-1) &= f(0) - f'(0) + \frac{f''(0)}{2} - \frac{f^{(3)}(t)}{6} \end{aligned}$$

so that

$$\begin{aligned}
1 &= 0 + 0 + \frac{f''(0)}{2} + \frac{f^{(3)}(s)}{6} \\
0 &= 0 + 0 + \frac{f''(0)}{2} - \frac{f^{(3)}(t)}{6} \\
1 &= \frac{1}{6}(f^{(3)}(s) + f^{(3)}(t)) \\
6 &= f^{(3)}(s) + f^{(3)}(t)
\end{aligned}$$

Hence either s or t satisfies $f^{(3)} \geq 3$ otherwise the equation fails. ☐

Problem A.24 ([Rud76] Exercise 5.22)**.** *Let f be real-valued function on $(-\infty, \infty)$. x is a fixed point of f if $f(x) = x$.*

1. *If f is differentiable and $f'(t) \neq 1$ for every t then f has at most one fixed point.*
2. *The function*
$$f(t) = t + (1 + e^t)^{-1}$$
has no fixed point.
3. *If there is a constant $A < 1$ s.t.*
$$|f'(t)| \leq A \qquad \forall\ t \in \mathbb{R}$$
Show that a fixed point x of f exists, and that
$$\lim_{n\to\infty} x_n = x$$
where x_1 is arbitrary real number and
$$x_{n+1} = f(x_n) \qquad \forall\ n \geq 1$$

Solution A.24. *Proof.*

1. Suppose f has two fixed points $x < y$. Then by Lagrange Mean Value Theorem Corollary 6.2 there exists $a \in (x, y)$ s.t.
$$f'(a) = \frac{f(y) - f(x)}{y - x} = \frac{y - x}{y - x} = 1$$
But we know $f'(t) \neq 1$ for any t by assumption. Hence this is a contradiction.
2. If there exists t a fixed point for f, then
$$f(t) = t + (1 + e^t)^{-1} = t \implies (1 + e^t)^{-1} = 0$$
But the function $(1 + e^t)^{-1}$ never achieves 0 for any $t \in \mathbb{R}$.
3. The point is, how can we define the fixed point? Since $|f'(t)| \leq A$ for any $t \in \mathbb{R}$, we know f is uniformly continuous. This is because for any $\varepsilon > 0$, one can safely pick $\delta = \frac{\varepsilon}{A}$ which is independent of points $x \in \mathbb{R}$ s.t.
$$|x - y| < \delta \implies |f(x) - f(y)| = |f'(t)||x - y| \leq A\delta \leq \varepsilon \qquad \text{where } t \in (x, y) \text{ exists by Lagrange MVT}$$
Then we notice that $\{x_n\}$ in fact defines a Cauchy Sequence.
   (a) We first show
   $$|x_{n+1} - x_n| \leq A^{n-1}|x_2 - x_1|$$
   At the base case $n = 1$ this holds trivially. Now assume for n, i.e.
   $$|x_{n+1} - x_n| \leq A^{n-1}|x_2 - x_1|$$
   We prove that
   $$\begin{aligned}
   |x_{n+2} - x_{n+1}| &= |f(x_{n+1}) - f(x_n)| \leq |f'(t)||x_{n+1} - x_n| \qquad \text{for some } t \text{ between } x_n \text{ and } x_{n+1} \\
   &\leq A|x_{n+1} - x_n| \leq AA^{n-1}|x_2 - x_1| = A^n|x_2 - x_1|
   \end{aligned}$$

(b) Then for any $n > m > N$, we compute

$$\begin{aligned}
|x_n - x_m| &= |x_n - x_{n-1}| + \cdots + |x_{m+1} - x_m| \\
&\leq A^{n-2}|x_2 - x_1| + \cdots + A^{m-1}|x_2 - x_1| \\
&= |x_2 - x_1|A^{m-1}(A^{n-m-1} + \cdots + 1) \\
&\leq |x_2 - x_1|A^{m-1}\frac{1}{1-A} \qquad \text{using geometric series and } A < 1 \\
&\leq |x_2 - x_1|A^{N}\frac{1}{1-A} \qquad \text{using } A < 1
\end{aligned}$$

Now the bound only depends on N. Hence for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ sufficiently large s.t. for any $n > m > N$

$$|x_n - x_m| \leq |x_2 - x_1|A^N\frac{1}{1-A} \leq \varepsilon$$

And this means $\{x_n\}$ is a Cauchy Sequence.

Now what is good about a Cauchy Sequence? In $\mathbb{R}$, which is complete, all Cauchy sequence converges. Hence there exists $x \in \mathbb{R}$ s.t.

$$x = \lim_{n\to\infty} x_n \in \mathbb{R}$$

It suffices to prove that x is a fixed point. Indeed, using continuity of f

$$f(x) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x$$

$\square$

A.4 Assignment 5

Problem A.25 ([Rud76] Exercise 6.1)**.** *Let α increase on $[a, b]$. At some $a \leq x_0 \leq b$, α is continuous at x_0, $f(x_0) = 1$ and*

$$f(x) = 0 \qquad \forall\ x \neq x_0$$

Show that $f \in R(\alpha)$ and

$$\int f d\alpha = 0$$

Solution A.25. *Proof.* For any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ s.t.

$$|x_0 - x| \leq \delta \implies |\alpha(x) - \alpha(x_0)| < \frac{\varepsilon}{2}$$

Now one may choose partition $P = \{x_i\}_{i=0}^n$ of $[a, b]$ s.t.

$$\Delta x_i = x_i - x_{i-1} \leq \delta \qquad \forall\ i = 1, \cdots, N$$

There necessarily exists some i_0 s.t.

$$x_{i_0-1} \leq x_0 \leq x_{i_0} \qquad i_0 \in \{1, \cdots, n\}$$

and thus

$$\begin{aligned}\sum_{i=1}^n M_i(\alpha(x_i) - \alpha(x_{i-1})) &= 0 + 1 \cdot (\alpha(x_{i_0}) - \alpha(x_{i_0-1})) \\ &= \alpha(x_{i_0}) - \alpha(x_0) + \alpha(x_0) - \alpha(x_{i_0-1}) \leq \varepsilon\end{aligned}$$

Because

$$|x_{i_0} - x_0| \leq \delta \qquad |x_{i_0-1} - x_0| \leq \delta$$

Hence

$$U(P, f) - L(P, f) = \sum_{i=1}^n M_i(\alpha(x_i) - \alpha(x_{i-1})) \leq \varepsilon$$

and conclude using Proposition 7.2. Now indeed

$$0 \leq \int f d\alpha \leq U(P, f) \leq \varepsilon$$

Let $\varepsilon \to 0$. □

Problem A.26 ([Rud76] Exercise 6.2)**.** *Suppose $f \geq 0$, f continuous on $[a, b]$ and*

$$\int_a^b f(x)dx = 0$$

Then

$$f(x) = 0 \qquad \forall\ x \in [a, b]$$

Solution A.26. *Proof.* Suppose there exists $x \in [a, b]$ s.t. $f(x) > 0$. Then using continuity of f, there exists $\delta = \delta(x) > 0$ s.t.

$$f(y) \geq \frac{f(x)}{2} > 0 \qquad \forall\ y \in (x - \delta, x + \delta) \cap [a, b]$$

WLOG assume $x \in (a, b)$ and hence $a < x - \delta < x + \delta < b$. We consider a partition on $(x - \delta, x + \delta)$, call it $P = \{x_i\}_{i=0}^n$. Then

$$\underline{\int_{x-\delta}^{x+\delta}} f dx \geq L(P, f) = \sum_{i=1}^n m_i \Delta x_i \geq \frac{f(x)}{2} 2\delta = \delta f(x) > 0$$

But using $f \geq 0$ necessarily

$$\int_a^b f \geq \underline{\int_{x-\delta}^{x+\delta}} f dx > 0$$

contradicting $\int_a^b f = 0$. □

Problem A.27 ([Rud76] Exercise 6.4)**.** *If*

$$f(x) = \begin{cases} 0 & x \in \mathbb{R} \setminus \mathbb{Q} \\ 1 & x \in \mathbb{Q} \end{cases} \tag{7}$$

Then $f \notin R[a, b]$ for any $a < b$.

Solution A.27. *Proof.* For any partition $P = \{x_i\}_{i=0}^n$ we pick for $[a, b]$, we examine what are M_i and m_i. Notice for any $[x_{i-1}, x_i]$, by density property of rationals 1.3, there exists $q_i \in (x_{i-1}, x_i)$ hence

$$M_i = 1 \qquad \forall\ i$$

Also there exists an irrational $r_i \in (x_{i-1}, x_i)$ hence

$$m_i = 0 \qquad \forall\ i$$

Thus

$$U(P, f) = \sum_{i=1}^n M_i \Delta x_i = \sum_{i=1}^n 1 \Delta x_i = b - a$$
$$L(P, f) = \sum_{i=1}^n m_i \Delta x_i = \sum_{i=1}^n 0 \cdot \Delta x_i = 0$$

for any arbitrary partition we pick. Thus the criterion Proposition 7.2 fails, hence $f \notin R[a, b]$ for any $a < b$. □

Problem A.28 ([Rud76] Exercise 6.5)**.** *Suppose f is real-valued and bounded function on $[a, b]$, and suppose $f^2 \in R[a, b]$. Does $f \in R[a, b]$? What about $f^3 \in R[a, b]$?*

Solution A.28. *Consider the example (7), then*

$$f^2 = 1 \qquad \text{over} \qquad \mathbb{R}$$

And hence for any $[a, b]$, $f^2 \in R[a, b]$. But as we've just shown in Problem A.27 $f \notin R[a, b]$.

But if we assume $f^3 \in R[a, b]$, then notice the function

$$\phi(u) := u^{\frac{1}{3}} \qquad \forall\ u \in f^3([a, b])$$

is continuous function. Thus using Theorem 7.4

$$\phi(f^3(x)) = (f^3)^{\frac{1}{3}}(x) = f(x) \in R[a, b]$$

Problem A.29 ([Rud76] Exercise 6.7)**.** *Suppose f is a real-valued function over $(0, 1]$ and that $f \in R[c, 1]$ for every $c > 0$. We define*

$$\int_0^1 f(x)dx := \lim_{c \to 0} \int_c^1 f(x)dx$$

if the limit exists and is finite. Then

1. *If $f \in R[0, 1]$ show that this definition agrees with the original one.*
2. *Construct f s.t. the above limit exists, but $\int_0^1 |f|$ fail to exist.*

Solution A.29. 1. *Since $f \in R[0, 1]$, we know that in the original definition*

$$\int_0^1 f(x)dx = \sup_{P \text{ partition of } [0,1]} L(P, f) = \inf_{P \text{ partition of } [0,1]} U(P, f)$$

We want to show that for any $\varepsilon > 0$, there exists $c_0 = c_0(\varepsilon) > 0$ small so that for any $0 < c < c_0$ one has

$$|\int_0^1 f(x)dx - \int_c^1 f(x)dx| < \varepsilon$$

But what is $\int_c^1 f(x)dx$? Again this is

$$\int_c^1 f(x)dx = \sup_{P \text{ partition of } [c,1]} L(P, f) = \inf_{P \text{ partition of } [c,1]} U(P, f)$$

First of all, since $f \in R[0,1]$, necessarily f is bounded over $[0,1]$, i.e.

$$M := \sup_{x\in[0,1]} |f(x)| < \infty$$

We fix $c_0 = \frac{\varepsilon}{2M}$ and consider any $c < c_0$. Now we take any partition $P = \{x_i\}_{i=1}^n$ that includes the point c, say $x_k = c > 0$, and s.t.

$$\sum_{i=1}^N M_i \Delta x_i - \frac{\varepsilon}{4} \leq \int_0^1 f(x)dx \leq \sum_{i=1}^N m_i \Delta x_i + \frac{\varepsilon}{4}$$
$$\sum_{i=k}^N M_i \Delta x_i - \frac{\varepsilon}{4} \leq \int_c^1 f(x)dx \leq \sum_{i=k}^N m_i \Delta x_i + \frac{\varepsilon}{4}$$

Now we compute

$$\begin{aligned}|\int_0^1 f(x)dx - \int_c^1 f(x)dx| &= |\int_0^1 f(x)dx - \sum_{i=1}^N M_i\Delta x_i + \sum_{i=1}^N M_i\Delta x_i - \sum_{i=k}^N M_i\Delta x_i + \sum_{i=k}^N M_i\Delta x_i - \int_c^1 f(x)dx| \\ &\leq |\int_0^1 f(x)dx - \sum_{i=1}^N M_i\Delta x_i| + |\sum_{i=1}^k M_i\Delta x_i| + |\sum_{i=k}^N M_i\Delta x_i - \int_c^1 f(x)dx| \\ &\leq \frac{\varepsilon}{4} + M\frac{\varepsilon}{2M} + \frac{\varepsilon}{4} = \varepsilon\end{aligned}$$

But the partition can be chosen for any $c < \frac{\varepsilon}{2M}$. Thus we've shown the equivalence.

2. *One shall expect f to oscillate crazily near 0. Define*

$$f(x) = (-1)^n(n+1) \qquad \forall\ x \in (\frac{1}{n+1}, \frac{1}{n}] \qquad \forall\ n \in \mathbb{N}^*$$

Now

$$\int_0^1 f = \sum_{n=1}^\infty \int_{\frac{1}{n+1}}^{\frac{1}{n}} (-1)^n(n+1) = \sum_{n=1}^\infty (-1)^n \frac{n+1}{n(n+1)} = \sum_{n-1}^\infty (-1)^n \frac{1}{n}$$

By Alternating Series Test this converges. On the other hand

$$\int_{\frac{1}{N+1}}^1 |f| = \sum_{n=1}^N \int_{\frac{1}{n+1}}^{\frac{1}{n}} (n+1) = \sum_{n=1}^N \frac{1}{n} \to \infty \qquad N \to \infty$$

Thus $\int_0^1 |f|$ does not exist.

Problem A.30 ([Rud76] Exercise 6.8)**.** *Suppose $f \in R[a,b]$ for any $a < b$. Define*

$$\int_a^\infty f(x)dx := \lim_{b\to\infty} \int_a^b f(x)dx$$

if the limit exists. In this case we say the integral converges.

Assume that $f(x) \geq 0$ and f decreases monotonically on $[1,\infty)$. Prove that

$$\int_1^\infty f(x)dx \qquad \text{converges}$$

iff

$$\sum_{n=1}^\infty f(n) \qquad \text{converges}$$

Solution A.30. *Proof.* 1. Assume $\int_1^\infty f(x) < \infty$, then since $f \geq 0$

$$S_k := \sum_{n=1}^k f(n) \leq \int_1^k f(x) \leq \int_1^\infty f(x) < \infty \qquad \forall\ k \in \mathbb{N}$$

Hence the partial sums $\{S_k\}$ is a bounded sequence. Notice $\{S_k\}$ is an increasing sequence, hence by Monotone Convergence Theorem 3.4, there exists $S \in \mathbb{R}$ s.t.

$$S_k = \sum_{n=1}^k f(n) \to S$$

Thus $\sum_{n=1}^\infty f(n)$ converges.

2. On the other hand assume $\sum_{n=1}^{\infty} f(n)$ converges. Suppose for contradiction that

$$\int_1^{\infty} f(x) = \infty$$

Then using that f is monotonically decreasing

$$f(n) = \sup_{x\in[n,n+1]} f(x) \qquad \forall\, n \in \mathbb{N}^*$$

Then the upper Riemann sum necessarily blows up, i.e.

$$S_k = \sum_{n=1}^{k} f(n) \geq \sum_{n=1}^{k} \sup_{x\in[n,n+1]} f(x)\cdot(n+1-n) \geq \int_1^k f(x)dx \to \infty \qquad k\to\infty$$

Thus

$$\sum_{n=1}^{k} f(n) \to \infty$$

contradicting our assumption.

□

Problem A.31 ([Rud76] Exercise 6.12)**.** *Let $f \in R[a,b]$ and $\varepsilon > 0$. Prove that there exists a continuous function g over $[a,b]$ s.t.*

$$\|g-f\|_2 := \left(\int_a^b |g(x)-f(x)|^2 dx\right)^{\frac{1}{2}} < \varepsilon$$

Solution A.31. *Proof.* For $P = \{x_i\}_{i=0}^n$ partition to be chosen, define a function

$$g(t) := \frac{x_i - t}{\Delta x_i} f(x_{i-1}) + \frac{t-x_{i-1}}{\Delta x_i} f(x_i) \qquad \forall\, x_{i-1} \leq t \leq x_i$$

as linear function in t, g is piecewise continuous over each interval $[x_{i-1}, x_i]$, and since at the endpoints

$$g(x_i) = f(x_i) \qquad \forall\, i \in \{0,\cdots,n\}$$

This g is indeed a continuous function. Notice

$$\begin{aligned}
\int_a^b |g-f|^2 &= \sum_{i=1}^{n}\int_{x_{i-1}}^{x_i} |\frac{x_i-t}{\Delta x_i} f(x_{i-1}) + \frac{t-x_{i-1}}{\Delta x_i} f(x_i) - f(t)|^2 dt \\
&= \sum_{i=1}^{n}\int_{x_{i-1}}^{x_i} |\frac{x_i-t}{\Delta x_i} f(x_{i-1}) + \frac{t-x_{i-1}}{\Delta x_i} f(x_i) - \frac{x_i - t + t - x_{i-1}}{\Delta x_i} f(t)|^2 dt \\
&= \sum_{i=1}^{n}\int_{x_{i-1}}^{x_i} |\frac{x_i-t}{\Delta x_i}(f(x_{i-1}) - f(t)) + \frac{t-x_{i-1}}{\Delta x_i}(f(x_i) - f(t))|^2 dt \\
&\leq 2\sum_{i=1}^{n}\int_{x_{i-1}}^{x_i} \left(|f(x_{i-1}) - f(t)|^2 + |f(x_i) - f(t)|^2\right) dt \\
&\leq 4\sum_{i=1}^{n} |M_i - m_i|^2 \Delta x_i \leq 8M \sum_{i=1}^{n} |M_i - m_i|\Delta x_i < \varepsilon^2
\end{aligned}$$

where we defined

$$M := \sup_{x\in[a,b]} |f(x)| < \infty$$

and picked partition s.t.

$$\sum_{i=1}^{n} |M_i - m_i|\Delta x_i = U(P,f) - L(P,f) \leq \frac{\varepsilon^2}{8M}$$

□

Problem A.32 ([Rud76] Exercise 6.15)**.** *Let f be real-valued, continuously differentiable function over $[a,b]$ s.t. $f(a) = f(b) = 0$ and*

$$\int_a^b f^2(x) = 1$$

Then

$$\int_a^b xf(x)f'(x) = -\frac{1}{2}$$
$$\int_a^b (f'(x))^2 \cdot \int_a^b x^2 f^2(x) > \frac{1}{4}$$

Solution A.32. *Proof.* Side remark: This is well-known as the Uncertainly principle from Quantum Mechanics. By Integration by Parts, one can

$$\begin{aligned}
1 &= \int_a^b f^2(x) = \int_a^b (x)' f^2(x)dx \\
&= bf^2(b) - af^2(a) - \int_a^b x(f^2)'(x)dx \\
&= 0 - \int_a^b x(f^2)'(x)dx \\
&= -2\int_a^b xf \cdot f'(x)dx \\
-\frac{1}{2} &= \int_a^b xf \cdot f'(x)dx
\end{aligned}$$

Now by Cauchy-Schwarz one obtain

$$\frac{1}{4} = \left(\int_a^b xf \cdot f'(x)dx\right)^2 \leq \int_a^b (f'(x))^2 \cdot \int_a^b x^2 f^2(x)$$

Notice the equality can be excluded. If it is equal, necessarily (up to some constant)

$$\int_a^b (f'(x))^2 = K\int_a^b x^2 f^2(x)$$

and thus

$$K(\int_a^b x^2 f^2(x))^2 = \frac{1}{4} \implies \int_a^b x^2 f^2(x) = \frac{1}{2\sqrt{K}}$$

But now

$$\begin{aligned}
1 &= 2\int_a^b \sqrt{K} x^2 f^2(x) \\
&= \int_a^b f^2(x)dx
\end{aligned}$$

Hence

$$2\sqrt{K}x^2 - 1 = 0 \qquad \text{over non-measure zero portion in } [a, b]$$

since otherwise $\int_a^b f^2 = 1$ fails. But

$$2\sqrt{K}x^2 - 1 = 0$$

at most two points, which is of measure zero. ☐

References

[Fol99] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1999.

[Rud76] W. Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1976.