Lecture Notes

Math 105A — Summer 2023
Real Analysis

University of California, Santa Cruz

Deewang Bhamidipati

(all errors introduced are my own)

Last Updated: Thursday 24th August, 2023

Contents

1 Some Basic Notions **3**
- 1.1 Logic of Quantifiers and Proof Methods . . . 3
- 1.2 Set Theory Preliminaries, Functions and Countability . . . 9

2 The Reals **14**
- 2.1 Rational Numbers . . . 14
- 2.2 Ordered Fields . . . 16
- 2.3 The Completeness Axiom and the Reals . . . 20
- 2.4 The Archimedean Principle . . . 22

3 Sequences **27**
- 3.1 Sequences & Convergent Sequences . . . 27
- 3.2 Limit Laws . . . 33
- 3.3 Subsequences . . . 36
- 3.4 Cauchy Sequences . . . 39

4 Series **41**
- 4.1 Series & Convergence Tests . . . 41
- 4.2 Absolute Convergence . . . 47

5 The Topology of Reals **52**
- 5.1 Open Sets . . . 52
- 5.2 Closed Sets . . . 56
- 5.3 Compact Sets . . . 58

6 Continuity **62**
- 6.1 Functional Limits and Properties . . . 62
- 6.2 Continuity . . . 66
- 6.3 Continuity on Compact Domains . . . 70
- 6.4 Uniform Continuity . . . 74

7 Differentiation **78**
- 7.1 The Derivative . . . 78
- 7.2 Affine Approximation . . . 82
- 7.3 Differentiability Rules . . . 83
- 7.4 Local Minimum and Maximums . . . 86
- 7.5 L'Hôpital's Rule . . . 89
- 7.6 Inverse Functions . . . 91

Lectures

- Lecture 1 (06/26), page 3
- Lecture 2 (06/28), page 7
- Lecture 3 (06/30), page 11
- Lecture 4 (07/03), page 13
- Lecture 5 (07/05), page 16
- Lecture 6 (07/07), page 20
- Lecture 7 (07/10), page 22
- Lecture 8 (07/12), page 27
- Lecture 9 (07/14), page 31
- Lecture 10 (07/17), page 33
- Lecture 11 (07/19), page 37
- Lecture 12 (07/21), page 39
- Lecture 13 (07/24), page 44
- Lecture 14 (07/26), page 47
- Lecture 15 (07/28), page 52
- Lecture 16 (07/31), page 56
- Lecture 17 (08/02), page 60
- Lecture 18 (08/04), page 62
- Lecture 19 (08/07), page 65
- Lecture 20 (08/09), page 68
- Lecture 21 (08/11), page 72
- Lecture 22 (08/14), page 74
- Lecture 23 (08/16), page 78
- Lecture 24 (08/18), page 83
- Lecture 25 (08/21), page 86
- Lecture 26 (08/23), page 88
- Lecture 27 (08/25), page 91

1. Some Basic Notions

Lecture 1

1.1. Logic of Quantifiers and Proof Methods

Definition 1.1.1 (Statement)**.** A Boolean expression is an expression admitting two possible values: `True` or `False`.

For us a statement will refer to Boolean expression, that is, declarative sentence which can be objectively determined to be either *true* or *false*.

Definition 1.1.2 (Logical Operators)**.** A logical operator (or a logical connective) is a mechanism to make Boolean expressions out of other boolean expressions. It can be viewed as a function $\{\mathbf{T}, \mathbf{F}\}^k \to \{\mathbf{T}, \mathbf{F}\}$, where k denotes the number of arguments.

The operators is unary if $k = 1$, and binary if $k = 2$.

Definition 1.1.3 (Negation)**.** The only really interesting unary logical operator (why are other unary operators uninteresting?) is the negation (denoted $\neg$) operator, with truth table

P	$\neg P$
T	**F**
F	**T**

For a statement P, we read $\neg P$ as `not` P.

Definition 1.1.4 (Important Binary Logical Operators)**.** The most relevant binary operators are:

conjunction (denoted $\wedge$),

disjunction (denoted $\vee$),

implication (denoted $\Rightarrow$), and

biconditional (denoted $\Leftrightarrow$),

with truth tables.

P	Q	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
T	**T**	**T**	**T**	**T**	**T**
T	**F**	**F**	**T**	**F**	**F**
F	**T**	**F**	**T**	**T**	**F**
F	**F**	**F**	**F**	**T**	**T**

For a statements P and Q, we read $P \wedge Q$ as P `and` Q, while we read $P \vee Q$ as P `or` Q.

There are multiple ways to express implication and biconditional in words, we list them below.

Discussion 1.1.5. The statement $P \Rightarrow Q$ is read in several different ways.

$$\left.\begin{array}{r} \textit{if } P\textit{, then } Q \\ P \textit{ implies } Q \\ Q \textit{ if } P \\ P \textit{ only if } Q \\ P \text{ is } \textit{sufficient} \text{ for } Q \\ Q \text{ is } \textit{necessary} \text{ for } P \end{array}\right\} \text{all mean } P \Rightarrow Q$$

While for $P \Leftrightarrow Q$ we say:

$$\begin{cases} P \text{ if and only if } Q \text{ (abbreviated as } P \text{ iff } Q) \\ P \text{ is equivalent to } Q \\ P \text{ is necessary } (Q \Rightarrow P) \text{ and sufficient } (P \Rightarrow Q) \text{ for } Q \end{cases}$$

The latter appearing as a consequence of Example 1.1.8 (d).

Definition 1.1.6 (Tautology and Contradiction)**.** A statement which is always `True` is called a tautology, such a statement is denoted as $\top$. A statement which is always `False` is called a contradiction is called a contradiction, such a statement is denoted as $\bot$.

Definition 1.1.7 (Logical Equivalence)**.** Two statements are logically equivalent if they have the same truth tables. We denote logical equivalence with $\equiv$.

Example 1.1.8. Let P, Q and R be statements.

(a) $(P \wedge Q) \equiv (Q \wedge P)$

(b) $(P \vee Q) \equiv (Q \vee P)$

(c) $(P \Rightarrow Q) \equiv (\neg P \vee Q)$

(d) $(P \Leftrightarrow Q) \equiv ((P \Rightarrow Q) \wedge (Q \Rightarrow P))$

(e) $\neg(P \wedge Q) \equiv (\neg P \vee \neg Q)$ (*De Morgan's Law A*)

(f) $\neg(P \vee Q) \equiv (\neg P \wedge \neg Q)$ (*De Morgan's Law B*)

(g) $(P \Rightarrow Q) \equiv (\neg Q \Rightarrow \neg P)$.
The latter statement is called the contrapositive of $P \Rightarrow Q$; when one, instead of giving a direct proof of $P \Rightarrow Q$, proves $\neg Q \Rightarrow \neg P$, then such a proof is called a *proof by contrapositive*.

(g) $R \equiv (\neg R \Rightarrow \bot)$.
When one, instead of giving a direct proof of R, proves $\neg R \Rightarrow \bot$, then such a proof is called a *proof by contradiction*.

(h) We have the following logical equivalences

$$\begin{aligned}((P \Leftrightarrow Q) \Leftrightarrow R) \\ &\equiv (P \Leftrightarrow (Q \Leftrightarrow R)) \\ &\equiv ((P \Leftrightarrow Q) \wedge (Q \Leftrightarrow R)) \\ &\equiv (((P \Rightarrow Q) \wedge (Q \Rightarrow R)) \wedge (R \Rightarrow P)) \\ &\equiv ((P \Rightarrow Q) \wedge ((Q \Rightarrow R) \wedge (R \Rightarrow P)))\end{aligned}$$

This will form the basis of proving the equivalence of more than two statements via a loop of implications.

Discussion 1.1.9. Logical operators for a large number of variables are less common, as the most we care about can be built from unary and binary ones. We mention two important examples.

(a) The $\wedge$ operators generalises to multiple inputs, namely

$$A_1 \wedge A_2 \wedge A_3 \wedge \cdots \wedge A_n = ((\cdots((A_1 \wedge A_2) \wedge A_3) \cdots) \wedge A_n)$$

This statement is `True` when all variables are true, and it is `False` as long as at least one variable is false.

(b) The $\vee$ operators generalises to multiple inputs, namely

$$A_1 \vee A_2 \vee A_3 \vee \cdots \vee A_n = ((\cdots((A_1 \vee A_2) \vee A_3) \cdots) \vee A_n)$$

This statement is `True` as long as at least one variable is true, and it is `False` when all variables are false.

The examples above may also be generalised to the case of infinitely inputs; *quantifiers* are the the gadgets that allow us to write this succinctly.

Definition 1.1.10 (Quantifiers)**.** Let S be a set, and $P(x)$ a statement for each $x \in S$. We produce quantified statements.

- *for all $x \in S$, $P(x)$ is true.*

  The phrase "for all" is referred to as the universal quantifier and is denoted by the symbol $\forall$. Other ways to express the universal quantifier are "for every", "for any" and "for each". Symbolically, we express the universally quantified statement as

  $$\forall x \in S,\ P(x)$$

  The statement is true if $P(x)$ is true for every $x \in S$.

  In particular, if $S = \{1, 2, 3, \ldots, n\}$ then

  $$\forall x \in S,\ P(x) \equiv P(1) \wedge P(2) \wedge P(3) \wedge \cdots \wedge P(n)$$

  In this way, the universal quantifier plays the role of an "and" operator over a potentially infinite set of boolean variables.

- *there exists $x \in S$ such that $P(x)$ is true.*

  The phrase "there exists" is referred to as the existential quantifier and is denoted by the symbol $\exists$. Symbolically, we express the existentially quantified statement as

  $$\exists x \in S,\ P(x)$$

  The statement is true if $P(x)$ is true for at least one $x \in S$.

  In particular, if $S = \{1, 2, 3, \ldots, n\}$ then

  $$\exists x \in S,\ P(x) \equiv P(1) \vee P(2) \vee P(3) \vee \cdots \vee P(n)$$

  In this way, the existential quantifier plays the role of an "or" operator over a potentially infinite set of boolean variables.

Discussion 1.1.11 (Negation of Quantified Statements)**.**

$$\begin{aligned}
\neg(\forall x \in S,\ P(x)) &= \text{it is not the case that for all } x \in S,\ P(x) \text{ is true.} \\
&= \text{there exists } x \in S \text{ such that } P(x) \text{ is false.} \\
&= \exists x \in S,\ \neg P(x)
\end{aligned}$$

$$\begin{aligned}
\neg(\exists x \in S,\ P(x)) &= \text{it is not the case that there exists } x \in S \text{ such that } P(x) \text{ is true.} \\
&= \text{for all } x \in S,\ P(x) \text{ is false.} \\
&= \forall x \in S,\ \neg P(x)
\end{aligned}$$

Summarising, we have

$$\text{under negation} \begin{cases} \wedge \leftrightarrow \vee \\ \forall \leftrightarrow \exists \\ P(x) \leftrightarrow \neg P(x) \end{cases}$$

Discussion 1.1.12 (Double Quantifiers)**.** Let $x \in S$ and $y \in T$ be variables. Consider the statement

$$\forall x \in S,\ \forall y \in T,\ P(x, y)$$

Then *the quantifiers, being of the same type, can interchanged*, i.e.

$$\forall x \in S,\ \forall y \in T,\ P(x, y) \equiv \forall y \in T,\ \forall x \in S,\ P(x, y)$$

One can also combine quantifiers of the same type, giving us that the statements above are equivalent to

$$\forall (x, y) \in S \times T,\ P(x, y) \equiv \forall (y, x) \in T \times S,\ P(x, y)$$

Similarly,

$$\exists x \in S,\ \exists y \in T,\ P(x, y) \equiv \exists y \in T,\ \exists x \in S,\ P(x, y)$$

and

$$\exists (x, y) \in S \times T,\ P(x, y) \equiv \exists (y, x) \in T \times S,\ P(x, y)$$

⚠ *Quantifiers of different types can never be interchanged*. That is, in general,

$$\forall x \in S,\ \exists y \in T,\ P(x,y) \not\equiv \exists y \in T,\ \forall x \in S,\ P(x,y)$$

To see this, consider the statement

$$\forall x \in \mathbb{Z},\ \exists y \in \mathbb{Z},\ y > x \tag{$*$}$$

which states "for every integer, there exists another integer that is larger". But if we were to interchange the quantifiers, one obtains the statement

$$\exists y \in \mathbb{Z},\ \forall x \in \mathbb{Z},\ y > x \tag{$\dagger$}$$

which states "there exists an integer that is larger than every integer". This is a distinct statement (that is false) from the earlier statement (which is true). The difference being that in ($*$) the y which comes after x *depends* on x, whereas in ($\dagger$) it does not.

Lecture 2

Discussion 1.1.13 (Proofs of Universally Quantified Statements)**.** Consider a universally quantified statement

$$\forall x \in S,\ P(x)$$

Is this true or false?

If you think it is true, *prove it* using various proof methods like direct proof, proof by contrapositive or contradiction.

If you think it is false, then its negation

$$\exists x \in S,\ \neg P(x)$$

is true. Therefore to show falsehood, we need to exhibit an (at least one) element $x_0 \in S$ such that $P(x_0)$ is false (that is, $\neg P(x_0)$ is true). x_0 is the called a counterexample to the given statement. Counterexamples prove the falsehood of, or disprove, universally quantified statements.

Discussion 1.1.14 (Proofs of Existentially Quantified Statements)**.** Consider a existentially quantified statement

$$\exists x \in S,\ P(x)$$

Methods to prove existence results include

(1) Exhibit $x \in S$ satisfying $P(x)$.

(2) Use other existence theorems (i.e., leverage other results).

(3) Use proof by contradiction. This is *the method of choice for proving existence results*.

We want to also mention a variation of the statement above. Recall that the above existentially quantified statement says *there exists an $x \in S$ such that $P(x)$ is true*. We focus on a statement when such an x is unique, that is, not only is $P(x)$ true for this x, x is the only element for which $P(x)$ is true. The statement is then *there exists a unique $x \in S$ such that $P(x)$ is true*, written symbolically as

$$(\exists x \in S,\ P(x)) \wedge (\forall y \in S,\ P(y) \Rightarrow y = x)$$

which is often shortened as

$$\exists! x \in S,\ P(x)$$

So, there are two steps to proving such a statement

Step 1. Existence proof. Prove such an x exists.

Step 2. Uniqueness proof. Prove such an x is unique.

We have already discussed Step 1. For Step 2., one typically does the following

(1) If $x, y \in S$ are such that $P(x)$ and $P(y)$ is true, prove that $y = x$.

(2) If $x, y \in S$ are such that $P(x)$ and $P(y)$ is true and $y \neq x$, then show that this leads to a contradiction.

Discussion 1.1.15 (Induction)**.** For any $b \in \mathbb{Z}$, let $\mathbb{N}_b := \{b, b+1, b+2, \ldots\} = \{k : k \geqslant b\} \subseteq \mathbb{Z}$. The method of proof known as the Principal of Mathematical Induction is a way to prove statements of the form

$$\forall n \in \mathbb{N}_b,\ P(n)$$

(often $b = 0$ or $b = 1$). This proof method appears in two flavours (which are logically equivalent).

Weak Principle of Mathematical Induction. The steps in this method are.

Step 1. (*Base Step*, or *Initial Step*) Prove that $P(b)$ is true.

Step 2. (*Inductive Step*) Prove that for all $k \geqslant b$, $P(k)$ implies $P(k+1)$. That is, assume $P(k)$ is true (*inductive hypothesis*) and conclude $P(k+1)$ is true, for any $k \geqslant b$.

Strong Principle of Mathematical Induction. The steps in this method are.

Step 1. (*Base Step*, or *Initial Step*) Prove that $P(b)$ is true.

Step 2. (*Inductive Step*) Prove that for all $k \geqslant b$, $P(b)$ and $P(b+1)$ and $\cdots$ and $P(k)$ together imply $P(k+1)$. That is, assume $P(i)$ is true (*inductive hypothesis*) for all $b \leqslant i \leqslant k$, and conclude $P(k+1)$, is true for any $k \geqslant b$.

We conclude this section with some etiquettes and guidelines for proof writing.

Remark 1.1.16 (Proof Writing Etiquettes and Guidelines)**.**

- Always begin your proof with carefully restating, in words, the assumptions and the result(s) that needs to be proved.
- A written proof is the presentation of your mathematical ideas. Since it's a presentation, it needs to be narrative, so always replace logical symbols with their English counterparts (e.g. instead of $\exists$ write 'there exists', instead of $\vee$ write 'or', instead of $\Rightarrow$ write 'implies' etc.)
- Since a written proof is a presentation, the reader is a part of this; therefore one should use the pronoun 'we'.
- Always write the proof as if it's going to be read by a peer; that's your audience.
- Within your proof, always highlight or display (that is, center) important equations or calculations and mathematical expressions.
- Clearly indicate the method of proof.
- Clearly indicate the various cases you may have to consider in a proof.

- Indicate the completion of your proof. You can do this by using the words 'Q.E.D.', or 'hence proved' or the *tombstone* □. If you are using LaTeX, the tombstone is automatically created within the `proof` environment.
- Most importantly: examples do not constitute a proof. Unless you're disproving a statement and you're doing so with a counterexample.

1.2. Set Theory Preliminaries, Functions and Countability

We assume the reader is comfortable with the notion of a set, but to motivate later concepts, we revisit some set theory preliminaries.

Discussion 1.2.1. A set X will be described by listing all its elements and enclosing it within `{ }`, e.g. $X = \{x, y, z, \ldots\}$.

- We use $\in$ to denote membership of, or belongingness to, a set

$$x \in X : x \textit{ is an element of a set } X \qquad x \notin X : x \textit{ is} \text{ not } \textit{an element of a set } X$$

- We will also often use the *set-builder notation* to describe sets, where we first introduce a general element and the describe the relevant property that defines the set. The recipe is

$$\{\text{typical element} \ : \ \text{properties of that element}\}$$

  The vertical bar is read *such that*; the notation as a whole is to be read as "those elements such that they have this property". For example,

$$\mathbb{Q} = \left\{ \frac{p}{q} \ : \ p, q \in \mathbb{Z}, \ q \neq 0 \right\}$$

- There is a set with *no elements*, denoted $\varnothing$; we call it the empty set.

  ⚠ While $\varnothing$ has no elements, the set $X = \{\varnothing\}$ is not the empty set since it contains a single element $\varnothing \in X$.

- Given two sets S and X, we say S is a subset of X if every element of S is also an element of X, and denote it as $S \subseteq X$. That is, S is the set of *some* elements of X.Any set X always has two subsets: itself and the empty set.

  To prove $S \subseteq X$, you need to prove the statement $x \in S \Rightarrow x \in X$.

- We say two sets A and B are equal if $A \subseteq B$ and $B \subseteq A$, and we denote it $A = B$.
- A set S is called a proper subset of X if $S \subseteq X$ and $S \neq X$; in particular, there exists an $x \in X$ such that $x \notin S$. We denote this as $S \subset X$ or $S \subsetneq X$.

  In other words, S is a subset that is "strictly smaller than X".

- The power set of a set X is the set of *all* subsets of X, including X and $\varnothing$. We denote it 2^X,

$$2^X = \{S \ : \ S \subseteq X\}$$

- The cardinality of a finite set is the number of elements in the set. For a set X, its cardinality is denoted $\#X$ or $|X|$, and we have $|2^X| = 2^{|X|}$.

  For a finite set X, we often write $|X| < \infty$ to indicate that its cardinality is finite.

Discussion 1.2.2 (Set Operations)**.** We assume that $A,\ B \subseteq U$, for some set U (soon, for us, $U = \mathbb{R}$), labelled the "universal set".

- The union of A and B is the set

$$A \cup B \coloneqq \{x \in U \ : \ x \in A \text{ or } x \in B\}$$

  Here our use of "or" allows the possibility of x being in both A and B.

- The intersection of A and B is the set

$$A \cap B \coloneqq \{x \in U \ : \ x \in A \text{ and } x \in B\}$$

  It is clear definitionally that we always have $A \cap B \subseteq A \cup B$.

- A and B are said to be disjoint if $A \cap B = \varnothing$.

- The complement of A in U is the set

$$A^{\mathsf{c}} \coloneqq U \setminus A = \{x \in U \ : \ x \notin A\}$$

  That is, it is the set of exactly those elements in U that don't belong to A.

- The difference of A and B is the set

$$A \setminus B \coloneqq A \cap B^{\mathsf{c}} = \{x \in U \ : \ x \in A \text{ and } x \notin B\}$$

  That is, it is the set that's left when we remove the part of B that belonged to A.

- If A and B are finite, then

$$|A \cup B| = |A| + |B| - |A \cap B| \leqslant |A| + |B|.$$

  Therefore we have $|A \cup B| = |A| + |B|$ if and only if A and B are disjoint.

Discussion 1.2.3 (Cartesian Product)**.** Let A and B sets.

- The cartesian product $A \times B$ is the set of ordered pairs

$$A \times B \coloneqq \{(a, b) \ : \ a \in A,\ b \in B\}$$

  Note that $A \times B \neq B \times A$, unless $A = B$.

- If A and B are finite, then $|A \times B| = |A| \cdot |B|$.

Discussion 1.2.4 (Functions)**.** Let $X,\ Y$ be sets

- A function
$$f : X \to Y$$
is a "rule" that assigns to each $x \in X$ a *unique* element $f(x) \in Y$.

- X is called the domain of f and denoted $X = \mathrm{dom}(f)$, while Y is called the codomain of f and denoted $Y = \mathrm{codom}(f)$.

- The subset $\{f(x) \ : \ x \in X\} \subseteq Y$ is called the range of f and denoted $f(X)$.

- If $g : X \to Y$ is another function, then we say f is equal to g, denoted $f = g$, if
$$f(x) = g(x)$$
for every $x \in X$.

- The set of all functions from X to Y is denoted as Y^X.

- If X and Y are finite, then Y^X is also finite and $|Y^X| = |Y|^{|X|}$.

Lecture 3

- Let $S \subseteq X$, then the image of S under f is
$$f_*(S) \coloneqq \{f(x) \ : \ x \in S\} \subseteq Y$$
Note that $f(X) = f_*(X)$.

- Let $T \subseteq Y$, then the preimage of T under f is
$$f^*(T) \coloneqq \{x \in X \ : \ f(x) \in T\} \subseteq X$$
Note that $f^*(Y) = X$.

- f is said to be injective (or one-to-one) if "f maps distinct elements to distinct elements"; that is, if for every $a, b \in X$, $a \neq b$ implies $f(a) \neq f(b)$.

  Equivalently, taking the contrapositive, f is injective if for every $a, b \in X$, $f(a) = f(b)$ implies $a = b$. This definition is more convenient for proofs.

- f is said to be surjective (or onto) if $f(X) = Y$, i.e., the codomain equals the range.

  Equivalently, "if every element in Y is mapped to by f", that is, for every $y \in Y$, there exists an $x \in X$ such that $y = f(x)$ (this is just saying $Y \subseteq f(x)$).

- f is bijective if it is both injective and surjective.

- For any set X, one can define the (bijective) identity function $\mathrm{id}_X : X \to X$, $x \mapsto x$.

- Let Z be another set and $g : Y \to Z$ a function, then the composition of f and g is the function
$$g \circ f : X \to Z$$
$$x \mapsto g(f(x))$$

- f is bijective if and only if there exists a function $g : Y \to X$ such that

$$f \circ g = \mathrm{id}_Y \quad \text{and} \quad g \circ f = \mathrm{id}_X .$$

  g is called the inverse function of f and denoted f^{-1}.

We now our attention to cardinality of sets and how we think about this notion for infinite sets. We first prove the following proposition to motivate the idea.

Proposition 1.2.5. *Let A and B be two finite sets, then $|A| = |B|$ if and only if there exists a bijective function $f : A \to B$.*

Proof. Since A and B are finite, let $n = |A|$ and $m = |B|$, and we may write $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_m\}$.

($\Rightarrow$) Suppose $n = |A| = |B| = m$, then consider the function

$$\begin{aligned} f : A &\to B \\ a_i &\mapsto b_i \end{aligned}$$

This is necessarily a bijection.

($\Leftarrow$) Suppose there exists a bijective function $f : A \to B$. Then, since in particular f is injective, we necessarily have

$$|f(A)| = |\{f(a_1), \ldots, f(a_n)\}| = n.$$

Furthermore, since f is surjective, we have $B = f(A)$. Therefore $m = |B| = |f(A)| = n$, and hence $|B| = |A|$. □

Definition 1.2.6 (Countability)**.** Let $\mathbb{N} := \mathbb{N}_1 = \{1, 2, 3, \ldots\}$, and consider a set A.

- A is said to be denumerable if there exists a bijection $f : \mathbb{N} \to A$ (or equivalently a bijective function $g : A \to \mathbb{N}$).

  This means we can enumerate, label with positive integers, the elements of A. So, we can write $A = \{a_1, a_2, \ldots\}$ where $a_i = f(i)$.

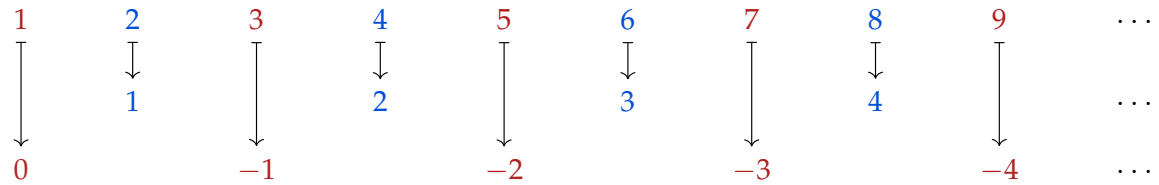
- A is said to be countable if $|A| < \infty$ or $|A| = |\mathbb{N}|$, that is, if A is finite or denumerable.

- A is said to be uncountable if A is not countable.

Example 1.2.7. The following sets are countable.

- The set of positive even integers $2\mathbb{N} = \{2n \ : \ n \in \mathbb{N}\}$ is countable since $f : \mathbb{N} \to 2\mathbb{N}$, $n \mapsto 2n$ is a bijection.

- The set of integers $\mathbb{Z}$ is countable, one deduces this by sending even numbers to positive integers, and odd numbers to non-positive integers.



  (try and write this function down explicitly)

- The set of rational numbers $\mathbb{Q}$ is countable.
- A countable union of countable sets is countable.
- A finite cartesian product of countable sets is countable.

Theorem 1.2.8. *The power set $2^{\mathbb{N}}$ is uncountable.*

Before we prove this we need a useful characterisation of power sets.

Proposition 1.2.9. *Let X be any set, then 2^X is in bijection with $\{0,1\}^X$.*

Proof. The idea of the proof is to send a subset of X to its "indicator function".
Consider the following function

$$\Phi : 2^X \to \{0,1\}^X,\ S \mapsto \chi_S$$

where $\chi_S : X \to \{0,1\}$ is the function

$$\chi_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases}$$

We note that Φ is surjective, as given any function $f : X \to \{0,1\}$, then $f^*(1) \subseteq X$ and we claim

$$f = \Phi(f^*(1)) = \chi_{f^*(1)}.$$

Note that for $x \in X$, we have $\chi_{f^*(1)}(x) = 1$ if and only if $x \in f^*(1)$ if and only if $f(x) = 1$. This tells us that $f = \chi_{f^*(1)}$.
Φ is also injective; to see this, let $S, T \subseteq X$ such that $\chi_S = \Phi(S) = \Phi(T) = \chi_T$. Let $t \in T$, then $1 = \chi_T(t) = \chi_S(t)$ which tells us $t \in S$ and thus $T \subseteq S$. We can similarly observe that $S \subseteq T$ and thus $S = T$.
Therefore Φ is our required bijection. □

Lecture 4

Proof of Theorem 1.2.8. For the sake of contradiction assume that $2^{\mathbb{N}}$ is countable, then so is $\{0,1\}^{\mathbb{N}}$ by Proposition 1.2.9. Hence we can enumerate its elements. So, let $\{0,1\}^{\mathbb{N}} = \{\chi_1, \chi_2, \ldots\}$.
We will produce an element of $\{0,1\}^{\mathbb{N}}$ distinct from each χ_i. Consider $\chi \in \{0,1\}^{\mathbb{N}}$ such that

$$\chi(i) = 1 - \chi_i(i) = \begin{cases} 1 & \text{if } \chi_i(i) = 0 \\ 0 & \text{if } \chi_i(i) = 1 \end{cases}$$

Then, $\chi \neq \chi_i$ for every i, since for any i we have $\chi(i) \neq \chi_i(i)$. Thus, we have produced an element in the set distinct from the enumerated element which should have been an exhaustive list; and so we have arrived at a contradiction.
Hence $\{0,1\}^{\mathbb{N}}$ is uncountable and therefore so is $2^{\mathbb{N}}$. □

2. The Reals

2.1. Rational Numbers

Discussion 2.1.1 (Notation)**.** Throughout this course, we will denote by $\mathbb{Z}$ the set of integers

$$\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

Our convention in this course will be that $\mathbb{N}$ is the set of (strictly) positive integers

$$\{1, 2, 3, \ldots\}$$

while $\mathbb{N}_0$ denotes the set of non-negative integers $\{0, 1, 2, 3, \ldots\}$. Finally, $\mathbb{Q}$ denotes the set of rational numbers

$$\left\{\frac{a}{b} \ : \ a, b \in \mathbb{Z},\ b \neq 0\right\}$$

Remark 2.1.2 (Constructions)**.**

- One can start with $\mathbb{N}$ and we quickly encounter a defect: there exists no natural number z such that $n + z = n$ for every $n \in \mathbb{N}$ and so we *construct* the number 0.
- Now, with $\mathbb{N}_0$ we encounter another defect: there exist no natural number n that can be a solution to the equation $a + n = b$ where $a, b \in \mathbb{N}$ such that $a > b$. We *construct* $\mathbb{Z}$ to fix this issue.
- Unfortunately, we still have some defects with $\mathbb{Z}$: there exist no integers x such that $ax = b$ where $b \neq a \neq 0$, and we *construct* $\mathbb{Q}$ to fix this issue.

We will soon see that $\mathbb{Q}$ itself has a defect and $\mathbb{R}$ is the set where this defect is resolved. We take this opportunity to remark on the mathematical input that goes into these constructions, although we will not discuss the constructions themselves. We will take for granted that $\mathbb{Q}$ exists.

Definition 2.1.3 (Equivalence Relations)**.** Given a set X, an equivalence relation $\sim$ on X is a relation satisfying three properties (modelling equality).

(i) (reflexivity) for every $x \in X$, we have $x \sim x$.

(ii) (symmetry) for every $x, y \in X$, if $x \sim y$ then $y \sim x$.

(iii) (transitivity) for every $x, y, z \in X$, if $x \sim y$ and $y \sim z$ then $x \sim z$.

Given $x \in X$, the equivalence class of x is defined to be the subset

$$[x] := \{y \in X \ : \ y \sim x\} \subseteq X.$$

By properties (i) – (iii), two equivalence classes are either disjoint or equal, and we may define the quotient $X/\sim$ to be the set of equivalence classes, that is,

$$X/\sim \ := \{[x] \ : \ x \in X\}.$$

Such a set always has at most as many elements as X itself, as the canonical map

$$\pi : X \to X/\sim, \quad x \mapsto [x]$$

is surjective.

Example 2.1.4. We define the following equivalence relation $\sim$ on the set $Q = \mathbb{Z} \times \mathbb{Z}^\circ$, where $\mathbb{Z}^\circ = \mathbb{Z} \setminus \{0\}$. We say,

$$(a, m) \sim (b, n) \quad \text{if and only if} \quad an = bm$$

Let us show that this is indeed an equivalence relation.

(i) (reflexivity) since $an = an$, we have $(a, n) \sim (a, n)$, for any $(a, n) \in Q$.

(ii) (symmetry) if $(a, n) \sim (b, m)$, then $am = bn$; so $bn = am$, and therefore $(b, m) \sim (a, n)$, for any $(a, n), (b, m) \in Q$.

(iii) (transitivity) consider $(a, n), (b, m), (c, k) \in Q$ such that $(a, n) \sim (b, m)$ and $(b, m) \sim (c, k)$. Therefore,

$$am = bn \quad \text{and} \quad bk = cm$$

We then obtain $amk = bnk$ and $bkn = cmn$; since $bnk = bkn$, we obtain $amk = cmn$. Since $m \neq 0$, we can cancel it and therefore $ak = cn$. Hence $(a, n) \sim (c, k)$.

What is the equivalence class of $(1, 2)$? We note

$$\begin{aligned}[(1,2)] &= \{(a, n) \in \mathbb{Z} \times \mathbb{Z}^\circ \ : \ (a, n) \sim (1, 2)\} \\ &= \{(a, n) \in Q \ : \ n = 2a\} \\ &= \{(a, 2a) \ : \ a \in \mathbb{Z}\}\end{aligned}$$

Remark 2.1.5 (Continuing Remark 2.1.2)**.** Equivalence relations are extremely useful to construct new sets with additional properties; in particular, the new sets are the resulting quotients.

This is how one may construct the integers $\mathbb{Z}$ out of $\mathbb{N}_0 \times \mathbb{N}_0$, the rationals $\mathbb{Q}$ out of $\mathbb{Z} \times \mathbb{N}$ (see Appendix **??**), and also the set of real numbers $\mathbb{R}$ out of a specifically designed set.

Discussion 2.1.6 (Properties & Issues with $\mathbb{Q}$)**.** The set of rational numbers has properties that make it amenable to arithmetic.

- They are *nested*: between any two rational numbers is another one. And hence, between any two there are infinitely many others.
- They are closed under addition, subtraction, multiplication and division:
  - ▷ If $a, b \in \mathbb{Q}$, then $a + b \in \mathbb{Q}$.
  - ▷ If $a, b \in \mathbb{Q}$, then $a - b \in \mathbb{Q}$.
  - ▷ If $a, b \in \mathbb{Q}$, then $ab \in \mathbb{Q}$.
  - ▷ If $a, b \in \mathbb{Q}$, then $\dfrac{a}{b} \in \mathbb{Q}$.

But they have the following algebraic and analytic defects.

- It's possible that $r \in \mathbb{Q}$ but, if it exists, $\sqrt{r} \notin \mathbb{Q}$.
- It has a lot of "gaps": it's possible that $r_i \in \mathbb{Q}$, for each $i \in \mathbb{N}$, "cluster" very close to each other but do not "converge" to a rational number. If we will the "fill in the gaps", we in fact see that "most" numbers are not rational!

Let us produce a number that is not rational.

Proposition 2.1.7. *The number $\sqrt{2}$, if it exists, is not rational.*

Proof. For the sake of contradiction, assume that $\sqrt{2}$ is rational. Therefore, there exist integers $a, b \in \mathbb{Z}$ with $b \neq 0$ such that

$$\sqrt{2} = \frac{a}{b}$$

We may further assume that this expression is reduced, that is, a and b have no common factors. Then,

$$\begin{aligned} b\sqrt{2} &= a \\ 2b^2 &= a^2 \end{aligned} \tag{$*$}$$

By definition of divisibility, we have $2 \mid a^2$. Since 2 is a prime, we have that $2 \mid a$; and so $a = 2k$ for some integer k. Plugging this in $(*)$, we obtain

$$\begin{aligned} 2b^2 &= (2k)^2 \\ 2b^2 &= 4k^2 \\ b^2 &= 2k^2 \end{aligned}$$

By definition of divisibility, we have $2 \mid b^2$; and since 2 is a prime, we have that $2 \mid b$. Thus, we have produced a common factor of both a and b, contradicting our assumption that a and b had no common factors. Hence, necessarily, $\sqrt{2}$ cannot be rational. □

So, the rationals are not enough. But they do have the necessary algebraic properties that we would want. They are what we call an *ordered field*.

2.2. Ordered Fields

Lecture 5

Definition 2.2.1 (Fields)**.** A field is a non-empty set $\mathbb{F}$, along with two binary operations *addition* $(+)$ and *multiplication* $(\cdot)$ satisfying the following axioms.

(F1) (Associative Law) If $a, b, c \in \mathbb{F}$, then $a + (b + c) = (a + b) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

(F2) (Distributive Law) If $a, b, c \in \mathbb{F}$, then $a \cdot (b + c) = a \cdot b + a \cdot c$.

(F3) (Commutative Law) If $a, b \in \mathbb{F}$, then $a + b = b + a$ and $a \cdot b = b \cdot a$.

(F4) (Identity Law) There exist elements $0, 1 \in \mathbb{F}$ such that $a + 0 = a$ and $a \cdot 1 = a$ for all $a \in \mathbb{F}$.

(F5) (Inverse Law) For each $a \in \mathbb{F}$, there exists an element $-a \in \mathbb{F}$ such that $a + (-a) = 0$. If $a \neq 0$, then there also exists a element $a^{-1} \in \mathbb{F}$ (also written as $1/a$) such that $a \cdot a^{-1} = 1$.

Example 2.2.2. Below are some examples and some non-examples of fields.

- The natural numbers $\mathbb{N}$ and $\mathbb{Z}$ do not form a field.
- The integers $\mathbb{Z}$ almost form a field.
- The rationals $\mathbb{Q}$ form a field.

Discussion 2.2.3. We want to motivate the need to consider the set of real numbers $\mathbb{R}$, and for that we wish to identify the properties $\mathbb{R}$ has, some which may be an extension to those of $\mathbb{Q}$, and some that may that fix a defect of $\mathbb{Q}$. Moreover, we would like these properties to uniquely characterise $\mathbb{R}$, i.e., they need to be minimal in some sense.

Let's focus on $\mathbb{Q}$; apart from its field properties, there are three properties we haven't axiomatically developed yet.

(1) There are infinitely many rationals.

(2) There is an ordering on $\mathbb{Q}$, i.e., not only elements the elements distinct, for a pair of rationals, one is greater than the other. In particular, this is a consequence of having a notion of *positive elements*, as by definition we have $r > s$ if and only if $r - s > 0$.

(3) Finally, we would like a metric or a distance function; a way to determine the *distance* of a number from 0, or more generally, the distance between any two rational numbers.

All of these are achieved by asking for one additional axiom be added to the field axioms.

Definition 2.2.4 (Ordered Fields)**.** An ordered field $\mathbb{F}$ is a field, along with the additional axiom

(OF1) (Order Axiom) There is a non-empty subset $P \subseteq \mathbb{F}$, the set of *positive elements*, such that

(a) If $a, b \in P$, then $a + b \in P$ and $a \cdot b \in P$.

(b) (Trichotomy Law) If $a \in \mathbb{F}$ and $a \neq 0$, then either $a \in P$ or $-a \in P$, but not both.

Definition 2.2.5 (Total Order)**.** If $\mathbb{F}$ is an ordered field, P its set of positive element, and $a, b \in \mathbb{F}$, then we say that b *is greater than* a, denoting it as $b > a$, if $b - a \in P$. In particular, we write $a > 0$ if $a \in P$. Then, the Trichotomy Law states that, for any $a \in \mathbb{F}$, either $a = 0$ or $a > 0$ or $a < 0$.

Furthermore, we write $b \geqslant a$, if $b - a \in P$ or $b - a = 0$.

Remark 2.2.6. One may use the axioms of an ordered field to obtain the following properties that we know to be standard for inequalities.

(a) If $b > a$, then $b + c > a + c$

(b) If $b > a$ and $c > b$, then $c > a$

(c) If $b > a$, then $bc > ac$ if $c > 0$, and $bc < ac$ if $c < 0$.

(d) If $a \neq 0$, then $a^2 > 0$

Remark 2.2.7. We have been able to axiomatise (2) in Discussion 2.2.3 using Definition 2.2.5. The axioms of an ordered field also allow us to conclude (1) in Discussion 2.2.3, i.e., an ordered field is necessarily infinite.

Note that necessarily $1 \in P$ (why?). Thus, so are $1 + 1$, $1 + 1 + 1$, ... and moreover $1 < 1 + 1 < 1 + 1 + 1 < \cdots$, and hence we have produced an infinite list of elements in $\mathbb{F}$.

We have yet to produce a distance function, to axiomatically realise (3) in Discussion 2.2.3. This is achieved by defining an *absolute value* function.

Definition 2.2.8 (Max and Min)**.** For an ordered field $\mathbb{F}$, we define the maximum function

$$\max : \mathbb{F} \times \mathbb{F} \to \mathbb{F}, \qquad \max(a,b) = \begin{cases} b & \text{if } b \geqslant a \\ a & \text{if } b < a \end{cases}$$

We note that $\max(a,b) = \max(b,a)$.

The minimum function is defined as

$$\min : \mathbb{F} \times \mathbb{F} \to \mathbb{F}, \qquad \min(a,b) = -\max(-a,-b) = \begin{cases} a & \text{if } b \geqslant a \\ b & \text{if } b < a \end{cases}$$

Similarly to the maximum function, we note that $\min(a,b) = \min(b,a)$.

Definition 2.2.9 (Absolute Value)**.** For an ordered field $\mathbb{F}$, we define the absolute value function

$$|\cdot| : \mathbb{F} \to \mathbb{F}, \qquad |x| = \max\{x,-x\} = \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0 \end{cases}$$

Remark 2.2.10. One may use the axioms of an ordered field to obtain the following properties that we know to be standard for the absolute value. Let $a, b \in \mathbb{F}$, an ordered field.

(a) $|a| \geqslant 0$, and moreover $|a| = 0$ if and only if $a = 0$

(b) $|-a| = |a|$

(c) $-|a| \leqslant a \leqslant |a|$

(d) $|a \cdot b| = |a| \cdot |b|$

(e) $|1/a| = 1/|a|$, if $a \neq 0$

(f) $|a/b| = |a| \,/\, |b|$, if $b \neq 0$

(g) $|a| \leqslant b$ if and only if $-b \leqslant a \leqslant b$

We now prove one of the most indispensable properties of the absolute value.

Theorem 2.2.11 (The Triangle Inequalities)**.** *Let $x, y \in \mathbb{F}$, an ordered field. Then*

(T1) $|x+y| \leqslant |x| + |y|$ (the triangle inequality)

(T2) $||x| - |y|| \leqslant |x-y|$ (the reverse triangle inequality)

Proof. We first prove (T1). If we wish to prove

$$|x+y| \leqslant |x| + |y|\,,$$

then by Remark 2.2.10 (g), it suffices for us to that

$$-(|x| + |y|) \leqslant x + y \leqslant (|x| + |y|)$$

By Remark 2.2.10 (c), one has

$$-|x| \leqslant x \leqslant |x| \quad \text{and} \quad -|y| \leqslant y \leqslant |y|\,.$$

Adding these two together, we obtain

$$-(|x| + |y|) \leqslant x + y \leqslant (|x| + |y|)$$

which is what we needed to conclude $|x + y| \leqslant |x| + |y|$.

We now prove (T2); if we wish to prove

$$||x| - |y|| \leqslant |x - y|\,,$$

then by Remark 2.2.10 (g), it suffices for us to that

$$-|x - y| \leqslant |x| - |y| \leqslant |x - y|$$

We will prove both inequalities separately, by using a clever application of (T1).

Let us first focus on the inequality of the right: $|x| - |y| \leqslant |x - y|$.

$$\begin{aligned} |x| = |x + (y - y)| &= |(x - y) + y| \\ &\leqslant |x - y| + |y| && \text{by (T1)} \\ |x| - |y| &\leqslant |x - y| && \text{subtracting } |y| \end{aligned}$$

Therefore we have proved the inequality we wished to prove. Let us now focus on the inequality of the left: $|y| - |x| = -(|x| - |y|) \leqslant |x - y|$. We achieve this by repeating the above computation, replacing x with y

$$\begin{aligned} |y| = |y + (x - x)| &= |(y - x) + x| \\ &\leqslant |y - x| + |x| && \text{by (T1)} \\ &\leqslant |x - y| + |x| && \text{by Remark 2.2.10 (b)} \\ |y| - |x| &\leqslant |x - y| && \text{subtracting } |x| \\ -|x - y| &\leqslant |x| - |y| && \text{multiplying } -1 \end{aligned}$$

which is what we needed to conclude $||x| - |y|| \leqslant |x - y|$. □

We conclude the section by defining our distance function just how one would expect.

Definition 2.2.12 (Distance)**.** Let $\mathbb{F}$ be an ordered field, then we define the distance function

$$d : \mathbb{F} \times \mathbb{F} \to \mathbb{F}, \qquad d(x, y) = |x - y|$$

Note that $d(x, y) = d(y, x)$; moreover $|x| = d(x, 0)$.

Asking for one more axiom, the *completeness axiom*, gives us $\mathbb{R}$.

2.3. The Completeness Axiom and the Reals

Lecture 6

Discussion 2.3.1. The *completeness axiom* is formally an axiom about *bounded sets*. One way to account for all real numbers, apart from the rational ones, is by asking for each bounded set to admit a *least upper bound*. For example, consider the set $\{r \in \mathbb{Q} : r^2 < 2\}$. This set is *bounded above*, e.g., numbers such as 1.5, 2 and 6 are all numbers that are great than any number in this set; they are all *upper bounds*. Amongst the set of all upper bounds, a special one is the smallest amongst all of them – the *least upper bound*, which will in fact have to be $\sqrt{2}$. In this way, we will have admitted $\sqrt{2}$ as a real number.

Definition 2.3.2 (Bounded Sets)**.** Let $\mathbb{B}$ be an ordered field, and $A \subseteq \mathbb{B}$ be non-empty.

(i) The set A is said to be bounded above, if there exists a $b \in \mathbb{B}$ such that $x \leqslant b$ for every $x \in A$. In this case, b is called an upper bound

(ii) The least upper bound of A, is some $b_0 \in \mathbb{B}$ such that

(L1) b_0 is an upper bound of A; and

(L2) if b is any other upper bound of A, then $b_0 \leqslant b$.

Such a b_0 is also called the supremum of A and written as $\sup(A)$.

(iii) The set A is said to be bounded below, if there exists a $b \in \mathbb{B}$ such that $b \leqslant x$ for every $x \in A$. In this case, b is called a lower bound

(iv) The greated lower bound of A, is some $b_0 \in \mathbb{B}$ such that

(G1) b_0 is a lower bound of A; and

(G2) if b is any other lower bound of A, then $b \leqslant b_0$.

Such a b_0 is also called the infimum of A and written as $\inf(A)$.

(v) If a set is both bounded above and bounded below, then it is called bounded.

Example 2.3.3. Treat any set in these examples as a subset of $\mathbb{Q}$.

(a) The set $\mathbb{N}$ has no upper bounds, while $\inf(\mathbb{N}) = 1$.

(b) $\mathbb{Q}$ has no upper or lower bounds, and therefore $\sup(\mathbb{Q})$ and $\inf(\mathbb{Q})$ don't exist.

(c) Let $T = \{1/n : n \in \mathbb{N}\}$, then $\sup(T) = 1$ and $\inf(T) = 0$. Note that while $\sup(T) \in T$ but $\inf(T) \notin T$.

(d) Let $S = \{n/(n+1) : n \in \mathbb{N}\}$, then $\sup(S) = 1$ and $\inf(S) = 1/2$. Note that while $\sup(S) \notin S$ but $\inf(S) \in S$.

(e) As discussed, $\sup(\{r \in \mathbb{Q} : r^2 < 2\})$ doesn't exist in $\mathbb{Q}$. But it will exist in $\mathbb{R}$, and moreover it will be $\sqrt{2}$.

Definition 2.3.4 (Complete Fields)**.** Let $\mathbb{B}$ be an ordered field. Then $\mathbb{B}$ is said to be complete if given any non-empty bounded above set $A \subseteq \mathbb{B}$, then A has a supremum in $\mathbb{B}$; i.e., $\sup(A) \in \mathbb{B}$.

Lemma 2.3.5. *Let $\mathbb{B}$ be an ordered field that is complete and assume $A \subseteq \mathbb{B}$ is a non-empty bounded below set. Then* $\inf(A)$ *exists.*

Proof. Consider the set $-A \coloneqq \{-x \,:\, x \in A\}$; we will prove that $-A$ is bounded above and that $\inf(A) = -\sup(-A)$. Since A is bounded below, then there exists an $L \in \mathbb{B}$ such that

$$L \leqslant x, \quad \text{for all } x \in A$$

Hence $-x \leqslant -L$, for all $x \in A$, and thus $-A$ is bounded above. Since $\mathbb{B}$ is complete, $\sup(-A)$ exists. Let us now prove that $-\sup(-A)$ is the infimum of A. We wish to prove that

(a) $-\sup(-A)$ is a lower bound of A; and

(b) if L is any other lower bound of A, then $L \leqslant -\sup(-A)$.

Since $\sup(-A)$ is the supremum of A, and so an upper bound of $-A$, and therefore $-x \leqslant \sup(-A)$ for all $x \in A$. Hence, $-\sup(-A) \leqslant x$, for all $x \in A$, and thus $-\sup(-A)$ is a lower bound of A, and we have proven (a).

For (b), let us consider any lower bound L of A, as observed in the beginning, then $-L$ is an upper bound of $-A$. Since $\sup(-A)$ is the supremum of A, we have $\sup(-A) \leqslant -L$, and therefore $L \leqslant -\sup(-A)$.

Hence, we have proven that $\inf(A) = -\sup(-A)$. □

Example 2.3.6.

(a) We have already remarked that $\mathbb{Q}$ is not complete.

(b) While $\mathbb{Z}$ is not a field, it is true that for any non-empty bounded above $A \subseteq \mathbb{Z}$, one has that $\sup(A) \in \mathbb{Z}$.

Theorem 2.3.7 (Uniqueness and Existence of $\mathbb{R}$)**.** *There exists a unique complete ordered field. We call this field the field of real numbers and denote it $\mathbb{R}$.*

From here on out, our universal set is $\mathbb{R}$, unless specified. It possesses an order and is complete.

Lemma 2.3.8 (Uniqueness of Suprema)**.** *If the supremum of $A \subseteq \mathbb{R}$ exists, then it is unique.*

Proof. Suppose α and β are suprema of some set $A \subseteq \mathbb{R}$, to conclude the uniquess of suprema, we wish to show that $\alpha = \beta$.

Since β is a supremum, it is an upper bound (L1); then by (L2) applied to the supremum α, one obtains $\alpha \leqslant \beta$. Similarly applying (L1) to α and (L2) to β, one has $\beta \leqslant \alpha$. Therefore $\alpha = \beta$. Hence suprema are unique. □

Remark 2.3.9. Since we can realise infima in terms of suprema (see Lemma 2.3.5), we obtain, as a corollary, that infima, when they exist, are unique.

Proposition 2.3.10 (Suprema Analytically)**.** *Let $A \subseteq \mathbb{R}$, then $\alpha = \sup(A)$ if and only if*

(i) α *is an upper bound of* A*; and*

(ii) for every $\varepsilon > 0$, $\alpha - \varepsilon$ *is* not *an upper bound of* A*, i.e., there exists an* $x \in A$ *such that* $x > \alpha - \varepsilon$.

Similarly, $\beta = \inf(A)$ *if and only if*

(i) β *is a lower bound of* A*; and*

(ii) for every $\varepsilon > 0$, $\beta + \varepsilon$ *is* not *a lower bound of* A*, i.e., there exists an* $x \in A$ *such that* $\beta + \varepsilon > x$.

Proof. We prove the case with the supremum, a proof for the case with infimum follows a similar argument.

Assuming $\alpha = \sup(A)$, let us prove (i) and (ii). Since $\alpha = \sup(A)$, then in particular it is an upper bound of A, and therefore we have (i). Since α is the least upper bound, i.e., smaller than any upper bound, and since $\alpha - \varepsilon < \alpha$, we conclude that $\alpha - \varepsilon$ is not an upper bound of A. Therefore, there exists an $x \in A$ such that $\alpha - \varepsilon < x$.

Let us now assume that (i) and (ii) are true, and prove that $\alpha = \sup(A)$. By (i) we know that α is an upper bound, therefore we must prove that if M is any upper bound of A, then $\alpha \leqslant M$. We will prove the contrapositive instead, we will show that if $M < \alpha$, then M is not an upper bound of A. Define $\varepsilon = \alpha - M > 0$, then applying (ii) to this ε, we have that $\alpha - \varepsilon = \alpha - (\alpha - M) = M$ is not an upper bound of A, which is what we wanted to prove. □

2.4. The Archimedean Principle

Lecture 7 This principle is fundamental to $\mathbb{R}$ and aids us in our proofs of determining superma of many sets.

Lemma 2.4.1 (Archimedean Principle)**.** *For any* $x \in \mathbb{R}$*, there exists an* $N \in \mathbb{N}$ *such that* $N > x$.

Proof. Note that for any real number $x \leqslant 0$, taking $N = 1$, gives us our result. Therefore, we may focus on the case when $x > 0$.

Let $x > 0$ be a real number, and assume for the sake of contradiction that no such natural number in the statement exists, i.e., for every $n \in \mathbb{N}$, we have $n \leqslant x$. Therefore x is an upper bound of $\mathbb{N}$, and so $\mathbb{N}$ is bounded above.

By the completeness of $\mathbb{R}$, $\alpha = \sup(\mathbb{N})$ exists. Since α is the supremum, by 2.3.10, $\alpha - 1$ is not an upper bound of $\mathbb{N}$. Therefore, there exists an $m \in \mathbb{N}$, such that $\alpha - 1 < m$. Hence $\alpha < m + 1$; but note that $m + 1 \in \mathbb{N}$. We have therefore arrived at a contradiction, since α is meant to be a supremum of $\mathbb{N}$ and thus no natural number can be greater than it.

Thus, there must exist a natural number N such that $N > x$. □

Corollary 2.4.2. *For any* $\varepsilon > 0$*, there exists an* $N \in \mathbb{N}$ *such that* $1/N < \varepsilon$.

Proof. Since $\varepsilon > 0$, so is $1/\varepsilon$. Applying Lemma 2.4.1, we know that there exists an $N \in \mathbb{N}$ such that $N > 1/\varepsilon$, and hence $1/N < \varepsilon$. □

Example 2.4.3. $\inf\left(\left\{\frac{1}{n} \;:\; n \in \mathbb{N}\right\}\right) = 0$

Proof. Let

$$A = \left\{ \frac{1}{n} \ : \ n \in \mathbb{N} \right\},$$

and we note that 0 is a lower bound of A since $0 < 1/n$ for all $n \in \mathbb{N}$. To show that 0 is the infimum, we will show that anything larger cannot be a lower bound of A (Proposition 2.3.10).

Consider any $0 + \varepsilon$ for $\varepsilon > 0$, then by the Archimedean principle, there exists an $N \in \mathbb{N}$ such that $1/N < \varepsilon$. Since $1/N \in A$, therefore ε is not a lower bound of A, and hence 0 is the infimum. □

Theorem 2.4.4 (Existence of Square Roots)**.** *If $a \in \mathbb{R}$ and $a > 0$, then $\sqrt{a} \in \mathbb{R}$.*

Proof. Consider the set $A = \left\{ x \in \mathbb{R} \ : \ x^2 < a \right\}$, we first prove that A is bounded above; note that it non-empty since $0 \in A$. We consider two cases, $a \geqslant 1$ or $a < 1$.

Suppose $a \geqslant 1$, then we claim that a is an upper bound of A. Since $a \geqslant 1$, then $a^2 \geqslant a$, and assume for the sake of contradiction that a is not an upper bound of A. Then there exists an element $x \in A$ such that $x > a$. Therefore, $a > x^2 > a^2 \geqslant a$, giving us $a > a$ and thereby handing a contradiction. Hence, a is an upper bound of A.

Suppose $a < 1$, then we claim that 1 is an upper bound of A. Assume for the sake of contradiction that 1 is not an uupper bound of A. Then there exists an element $x \in A$ such that $x > 1$. Therefore, $a > x^2 > 1^2 = 1$, giving $a > 1$ and thereby handing us a contradiction. Hence, 1 is an upper bound of A.

Hence, A is a bounded above set and its supremum exists in $\mathbb{R}$. Let $s = \sup(A)$, we wish to show that $s^2 = a$; since $0 \in A$, we have that $0 < s$. Assume for the sake of contradiction that $s^2 \neq a$

CASE I. Assume $s^2 < a$. We will show that s cannot be an upper bound by producing an $N \in \mathbb{N}$ such that $s + 1/N \in A$, as $s + 1/N > s$.

> *Scratch Work.* To produce such an N, let's investigate what it means for $s + 1/N$ to be an element of A.
>
> $$\begin{aligned} a &> \left(s + \frac{1}{N} \right)^2 \\ &= s^2 + \frac{2s}{N} + \frac{1}{N^2} \\ &\leqslant s^2 + \frac{2s}{N} + \frac{1}{N} \\ &= s^2 + \frac{2s+1}{N} \end{aligned}$$
>
> $$a - s^2 > \frac{2s+1}{N}$$
>
> $$\frac{1}{N} < \frac{a - s^2}{2s+1}$$
>
> Since, by assumption, $s^2 < a$, we get $a - s^2 > 0$; therefore $(a - s^2)/(2s+1) > 0$. Thus, the existence of this N is guaranteed by the Archiemedean Principle.

Since, by assumption, $s^2 < a$, we get $a - s^2 > 0$; therefore $(a - s^2)/(2s + 1) > 0$. Thus, by the Archimedean Principle, there exists an $N \in \mathbb{N}$ such that

$$\frac{1}{N} < \frac{a - s^2}{2s + 1}$$

which tells us that

$$\left(s + \frac{1}{N}\right)^2 = s^2 + \frac{2s}{N} + \frac{1}{N^2} \leqslant s^2 + \frac{2s}{N} + \frac{1}{N} = s^2 + \frac{2s + 1}{N} < s^2 + (2s + 1) \cdot \frac{a - s^2}{2s + 1} = a$$

and hence $s + 1/N \in A$, but $s < s + 1/N$, contradicting the fact that s is an upper bound.

CASE II. Assume $s^2 > a$. We will show that s cannot be the least upper bound by producing an $N \in \mathbb{N}$ such that $s - 1/N$ is an upper bound of A, as $s - 1/N < s$. We will produce an $N \in \mathbb{N}$ such that $(s - 1/N) \notin A$, which is equivalent to saying $(s - 1/N)^2 > a$; which will allow us to conclude that $s - 1/N$ is an upper bound of A. Since, if it was not, then there would exist an $x \in A$ such that $s - 1/N < x$ and then

$$a^2 < \left(s - \frac{1}{N}\right)^2 < x^2 < a^2,$$

which is a contradiction since we cannot have $a^2 < a^2$.

> *Scratch Work.* To produce such an N, let's investigate what it means for $s - 1/N$ to not be an element of A.
>
> $$\begin{aligned} a &< \left(s - \frac{1}{N}\right)^2 \\ &= s^2 - \frac{2s}{N} + \frac{1}{N^2} \\ &\geqslant s^2 - \frac{2s}{N} \end{aligned}$$
>
> $$\frac{2s}{N} < s^2 - a$$
>
> $$\frac{1}{N} < \frac{s^2 - a}{2s}$$
>
> Since, by assumption, $s^2 > a$, we get $s^2 - a > 0$; therefore $(s^2 - a)/(2s) > 0$. Thus, the existence of this N is guaranteed by the Archiemedean Principle.

Since, by assumption, $s^2 > a$ and $s^2 - a > 0$; therefore therefore $(s^2 - a)/(2s) > 0$. Thus, by the Archimedean Principle, there exists an $N \in \mathbb{N}$ such that

$$\frac{1}{N} < \frac{s^2 - a}{2s}$$

which tells us that

$$\left(s - \frac{1}{N}\right)^2 = s^2 - \frac{2s}{N} + \frac{1}{N^2} \geqslant s^2 - \frac{2s}{N} > s^2 - 2s \cdot \frac{s^2 - a}{2s} = a$$

and hence $s - 1/N \notin A$, and by our reasoning earlier, it is an upper bound of A. But $s - 1/N < s$, contradicting the fact that s is the *least* upper bound.

Hence $s^2 = a$, i.e. the square root of a exists as a real number. □

Definition 2.4.5 (Dense Sets)**.** Suppose $\mathbb{B}$ is an ordered field and $S \subseteq \mathbb{B}$ is non-empty, then S is dense in $\mathbb{B}$, if for every $x, y \in \mathbb{B}$, there exists an $r \in S$ such that $x < r < y$.

We wish to show that $\mathbb{Q}$ is dense in $\mathbb{R}$. To prove that, we first prove a preliminary lemma: that if two real numbers are far away from each other enough, there's an integer lying between them.

Lemma 2.4.6. *Let x, y be real numbers such that $y - x > 1$, then there exist an $M \in \mathbb{Z}$ such that $x < M < y$.*

Proof. Let us first assume that $x, y \geqslant 0$, and consider the set

$$A = \{n \in \mathbb{N}_0 \,:\, n \leqslant x\}$$

If $x = 0$, then $A = \{0\}$, and is therefore finite. If $x > 0$, then by the Archimedean Principle, there exists an $N \in \mathbb{N}$ such that $x < N$. Then,

$$A \subseteq \{n \in \mathbb{N} \,:\, n \leqslant N\},$$

since the latter set is clearly finite, so is A. Hence, in either case, we have that A is finite and thus contains a maximum, say $m = \max(A)$. We claim that $M \coloneqq m + 1$ is the integer we are looking for. Since $m \in \mathbb{N}_0$, we obtain $M \in \mathbb{N}$.

Moreover, since $M > m$, the largest element in A, we have $M \notin A$, and therefore $x < M$. Since $m \in A$, we obtain $m \leqslant x$ and hence

$$M = m + 1 \leqslant x + 1 < y$$

This gives us $x < M < y$.

We are left with two cases: suppose both $x, y < 0$, then we can apply the above result to $-x$ and $-y$ to produce an integer M such that $-y < M < -x$, and then $x < -M < y$. And if one of them is non-negative and the other negative, then 0 works as the integer we want to lie between x and y. □

Theorem 2.4.7 (Density of $\mathbb{Q}$)**.** *$\mathbb{Q}$ is dense in $\mathbb{R}$.*

Proof. Consider any $x, y \in \mathbb{R}$, then $x < y$. Since $y - x > 0$, then by the Archimedean Principle there exists an $n \in \mathbb{N}$ such that

$$\frac{1}{y - x} < n$$

Therefore $ny - nx = n(y - x) > 1$, and hence by Lemma 2.4.6 there exists an $m \in \mathbb{Z}$ such that

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

which concludes the proof. □

We record some final definitions. The first one is a consequence of the proof of Lemma 2.4.6.

Definition 2.4.8 (Ceiling & Floor)**.** Let $x \in \mathbb{R}$.

- The ceiling of x, denoted $\lceil x \rceil$, is the unique integer n such that $x \leqslant n < x + 1$. That is, this is the smallest integer we can round up our real number to.
- The floor of x, denoted $\lfloor x \rfloor$, is the unique integer n such that $x - 1 < n \leqslant x$. That is, this is the largest integer we can round down our real number to.

Definition 2.4.9 (Interval Notations)**.** We define the open interval as

$$(a, b) = \{x \in \mathbb{R} \ : \ a < x < b\}$$

and closed interval as

$$[a, b] = \{x \in \mathbb{R} \ : \ a \leqslant x \leqslant b\}$$

While the right/left open rays and left/right closed rays intervals are given as

$$(a, \infty) = \{x \in \mathbb{R} \ : \ a < x\} \qquad [a, \infty) = \{x \in \mathbb{R} \ : \ a \leqslant x\}$$
$$(-\infty, b) = \{x \in \mathbb{R} \ : \ x < b\} \qquad (-\infty, b] = \{x \in \mathbb{R} \ : \ x \leqslant b\}$$

Proposition 2.4.10. *We have the following for the intervals and rays introduced above.*

$$\inf((a, b)) = \inf([a, b]) = a \qquad \sup((a, b)) = \sup([a, b]) = b$$
$$\inf((a, \infty)) = \inf([a, \infty)) = a \qquad \sup((-\infty, b)) = \sup((-\infty, b]) = b$$

While $\sup((a, \infty))$, $\sup([a, \infty))$ *and* $\inf((-\infty, b))$, $\inf((-\infty, b])$ *does not exist.*

3. Sequences

3.1. Sequences & Convergent Sequences

Lecture 8

Definition 3.1.1 (Sequences)**.** A sequence of real numbers is a function $a : \mathbb{N} \to \mathbb{R}$.

Remark 3.1.2 (Sequence Notation)**.** Given a sequence $a : \mathbb{N} \to \mathbb{R}$.

- We usually write a_n for $a(n)$, and refer to it as a term of the sequence.
- We usually write the sequence as an ordered list $(a_1, a_2, a_3, \ldots)$, which we shorten as

$$(a_n)_{n\in\mathbb{N}} \quad \text{or} \quad (a_n)_{n=1}^{\infty} \quad \text{or} \quad (a_n)_{n\geqslant 1} \quad \text{or} \quad (a_n)_n \quad \text{or} \quad (a_n).$$

Definition 3.1.3 (Bounded Sequences)**.** A sequence (a_n) is bounded if the range $\{a_n : n \in \mathbb{N}\}$ is bounded. That is, there exists a lower bound $L \in \mathbb{R}$ and an upper bound $U \in \mathbb{R}$ such that

$$L \leqslant a_n \leqslant U$$

for every $n \in \mathbb{N}$.

Proposition 3.1.4. *A sequence (a_n) is bounded if and only if there exists an $M \in \mathbb{R}$ such that $|a_n| \leqslant M$ for every $n \in \mathbb{N}$.*

Proof. ($\Leftarrow$) Suppose there existed an $M \in \mathbb{R}$ such that $|a_n| \leqslant M$ for every $n \in \mathbb{N}$, then

$$-M \leqslant a_n \leqslant M, \quad \text{for all } n \in \mathbb{N}$$

Therefore $-M$ is a lower bound, and M is an upper bound, and hence (a_n) is bounded.

($\Rightarrow$) Suppose (a_n) is bounded, therefore there exists a lower bound $L \in \mathbb{R}$ and an upper bound $U \in \mathbb{R}$ such that $L \leqslant a_n \leqslant U$ for every $n \in \mathbb{N}$. Define

$$M = \max(|L|, U),$$

then $U \leqslant M$ and $|L| \leqslant M$, so $-M \leqslant -|L|$. Hence, we have

$$-M \leqslant -|L| \leqslant L \leqslant a_n \leqslant U \leqslant M$$

Thus $|a_n| \leqslant M$, for all $n \in \mathbb{N}$. □

Example 3.1.5. Some examples of (bounded) sequences.

(a) Consider the sequence (a_n) given by

$$a_n = (-1)^n \frac{1}{n}, \quad \text{i.e., } (a_n) = \left(-1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, -\frac{1}{5}, \ldots\right).$$

Then,

$$|a_n| = \left|(-1)^n \frac{1}{n}\right| = \frac{1}{n} \leqslant 1$$

Therefore (a_n) is bounded.

(b) Consider the sequence (a_n) given by $a_n = 2 + \sin(n)$,

$$(2 + \sin(1), 2 + \sin(2), 2 + \sin(3), 2 + \sin(4), \ldots)$$

This sequence is bounded, to see that we note

$$-1 \leqslant \sin(n) \leqslant 1, \quad \text{and so } \; 1 \leqslant 2 + \sin(n) \leqslant 3$$

for every $n \in \mathbb{N}$. Therefore (a_n) is bounded.

Definition 3.1.6 (Convergent Sequences)**.** A sequence (a_n) converges to $a \in \mathbb{R}$ if *for all* $\varepsilon > 0$, *there exists an* $N \in \mathbb{N}$ *such that*

$$|a_n - a| < \varepsilon,$$

for all $n \geqslant N$.

Discussion 3.1.7. The final condition in Definition 3.1.6 that we are aiming to achieve is $|a_n - a| < \varepsilon$ for an appropriate n. Note that this is the same as $-\varepsilon < a_n - a < \varepsilon$ which is equivalent to

$$a - \varepsilon < a_n < a + \varepsilon$$

So, our sequence has to live inside of this interval after some point N.

That is, eventually (i.e. after N), the points of the sequence (a_n) get *arbitrary close* to a, as this needs to hold true for every $\varepsilon > 0$, and therefore for smaller and smaller ε, in particular.

Remark 3.1.8 (Convergent Sequence Notation)**.** The following are notationally equivalent.

- The sequence (a_n) *converges* to a.
- $a_n \to a$ as $n \to \infty$.
- $a_n \to a$.
- $\lim\limits_{n\to\infty} a_n = a$.

Example 3.1.9. Show that the sequence (a_n) with $a_n = \dfrac{1}{n}$ converges to 0. That is,

$$\lim_{n\to\infty} \frac{1}{n} = 0$$

Solution. We begin with some scratch work.

> *Scratch Work.* Given an arbitrary $\varepsilon > 0$, we need to find a specific N such that whenever $n \geqslant N$ we have $|a_n - 0| < \varepsilon$. We first observe that
>
> $$|a_n - 0| = \left|\frac{1}{n} - 0\right| = \frac{1}{n}$$
>
> and since $n \geqslant N$, we have
>
> $$\frac{1}{n} \leqslant \frac{1}{N}$$

Therefore, if we can find an N such that $\frac{1}{N} < \varepsilon$, we will obtain

$$|a_n - 0| = \frac{1}{n} \leqslant \frac{1}{N} < \varepsilon$$

The existence of this N is guarantee by the Archiemedean property.

Fix any $\varepsilon > 0$, then by the Archimedean property there exists an $N \in \mathbb{N}$ such that $1/N < \varepsilon$. Then, for any $n \geqslant N$, we get

$$|a_n - 0| = \frac{1}{n} \leqslant \frac{1}{N} < \varepsilon.$$

Therefore, we have proved $\frac{1}{n} \to 0$. □

Example 3.1.10. Show that the sequence (a_n) with $a_n = 5 - \frac{1}{n^2}$ converges to 5. That is,

$$\lim_{n\to\infty} \left(5 - \frac{1}{n^2}\right) = 5$$

Solution. We begin with some scratch work.

Scratch Work. Given an arbitrary $\varepsilon > 0$, we need to find a specific N such that whenever $n \geqslant N$ we have $|a_n - 5| < \varepsilon$. We first observe that

$$|a_n - 5| = \left|5 - \frac{1}{n^2} - 5\right| = \left|-\frac{1}{n^2}\right| = \frac{1}{n^2}$$

and since $n \geqslant N$, we have

$$\frac{1}{n^2} \leqslant \frac{1}{N^2}$$

Therefore, if we can find an N such that $\frac{1}{N^2} < \varepsilon$, or equivalently $\frac{1}{N} < \sqrt{\varepsilon}$, we will obtain

$$|a_n - 5| = \frac{1}{n^2} \leqslant \frac{1}{N^2} < \varepsilon$$

Using the Archiemedean property.

Fix any $\varepsilon > 0$, then by the Archimedean property there exists an $N \in \mathbb{N}$ such that $1/N < \sqrt{\varepsilon}$. Then, for any $n \geqslant N$, we get

$$|a_n - 5| = \frac{1}{n^2} \leqslant \frac{1}{N^2} < \varepsilon.$$

Therefore, we have proved $5 - \frac{1}{n^2} \to 5$. □

Example 3.1.11. Show that the sequence (a_n) with $a_n = \frac{3n+1}{n+2}$ converges to 3. That is,

$$\lim_{n\to\infty} \frac{3n+1}{n+2} = 3$$

Solution. We begin with some scratch work.

> *Scratch Work.* Given an arbitrary $\varepsilon > 0$, we need to find a specific N such that whenever $n \geqslant N$ we have $|a_n - 3| < \varepsilon$. We first observe that
>
> $$\begin{aligned} |a_n - 3| = \left|\frac{3n+1}{n+2} - 3\right| &= \left|\frac{3n+1}{n+2} - 3\left(\frac{n+2}{n+2}\right)\right| \\ &= \left|\frac{3n+1-3(n+2)}{n+2}\right| \\ &= \left|\frac{-5}{n+2}\right| \\ &= \frac{5}{n+2} \end{aligned}$$
>
> and since $n \geqslant N$, we have
>
> $$\frac{5}{n+2} \leqslant \frac{5}{N+2}$$
>
> Therefore, if we can find an N such that $\dfrac{5}{N+2} < \varepsilon$, or equivalently (isolating N)
>
> $$\begin{aligned} \frac{5}{\varepsilon} &< N+2 \\ \frac{5}{\varepsilon} - 2 &< N \\ \frac{5-2\varepsilon}{\varepsilon} &< N \\ \frac{1}{N} &< \frac{\varepsilon}{5-2\varepsilon}, \end{aligned}$$
>
> we will obtain
>
> $$|a_n - 3| = \frac{5}{n+2} \leqslant \frac{5}{N+2} < \varepsilon$$
>
> The existence of this N is guarantee by the Archiemedean property.

Fix any $\varepsilon > 0$. If $\varepsilon \neq 5/2$, then by the Archimedean property there exists an $N \in \mathbb{N}$ such that $1/N < \varepsilon/(5-2\varepsilon)$. Then, for any $n \geqslant N$, we get

$$|a_n - 3| = \frac{5}{n+2} \leqslant \frac{5}{N+2} < \varepsilon.$$

Now, assume that $\varepsilon = 5/2$, then for any $n \geqslant 1$, we have $|a_n - 3| = \dfrac{5}{n+2} < \dfrac{5}{2}$.

Therefore, we have proved $\dfrac{3n+1}{n+2} \to 3$. □

Lemma 3.1.12 (Constant Sequences). *Let $c \in \mathbb{R}$, and consider the* constant sequence (a_n) *where $a_n = c$ for every $n \in \mathbb{N}$. Then, $a_n \to c$.*

Solution. We begin with some scratch work.

> *Scratch Work.* Given an arbitrary $\varepsilon > 0$, we need to find a specific N such that whenever $n \geqslant N$ we have $|a_n - c| < \varepsilon$. We observe that $|a_n - c| = |c - c| = 0$.

Fix any $\varepsilon > 0$, then for any $n \geqslant 1$, we get

$$|a_n - c| = |c - c| = 0 < \varepsilon.$$

Therefore, we have proved $a_n \to c$. □

Lecture 9

Definition 3.1.13 (Divergent Sequences)**.** If a sequence (a_n) does not converge, then it diverges.

Divergence appears in three flavours.

- (a_n) diverges to ∞ if, for all $M > 0$, there exists an $N \in \mathbb{N}$ such that $a_n > M$ for all $n \geqslant N$; denoted $\lim_{n\to\infty} a_n = \infty$.
- (a_n) diverges to $-\infty$ if, for all $M < 0$, there exists an $N \in \mathbb{N}$ such that $a_n < M$ for all $n \geqslant N$; denoted $\lim_{n\to\infty} a_n = -\infty$.
- Otherwise, the limit of (a_n) does not exist.

Example 3.1.14. Show that the sequence (a_n) with $a_n = n^2$ diverges to ∞.

Solution. We begin with some scratch work.

> *Scratch Work.* Given an arbitrary $M > 0$, we need to find a specific N such that whenever $n \geqslant N$ we have $n^2 = a_n > M$.
>
> Since $n \geqslant N$, we have $n^2 \geqslant N^2$; therefore, if we can find an N such that $N^2 > M$, or $N > \sqrt{M}$, we will obtain
>
> $$n^2 \geqslant N^2 > M$$
>
> The existence of this N is guarantee by the Archiemedean property.

Fix any $M > 0$, then by the Archimedean property there exists an $N \in \mathbb{N}$ such that $N > \sqrt{M}$. Then, for any $n \geqslant N$, we get

$$a_n = n^2 \geqslant N^2 > M.$$

Therefore, we have proved $n^2 \to \infty$. □

Example 3.1.15. Show that the limit of the sequence (a_n), with $a_n = (-1)^n$, does not exist.

Solution. Assume, for the sake of contradiction, there exists an $a \in \mathbb{R}$ such that $a_n = (-1)^n \to a$. Therefore, for any $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|(-1)^n - a| < \varepsilon$, for all $n \geqslant N$.

Since this statement is true for every $\varepsilon > 0$, in particular, it is true for $\varepsilon = 1/2$. For which, there exists an $N \in \mathbb{N}$ such that

$$|(-1)^n - a| < \frac{1}{2}, \quad \text{for all } n \geqslant N$$

Case I. Suppose n is even and $n \geqslant N$, then

$$|1 - a| \leqslant \frac{1}{2}$$

Case II. Suppose n is odd and $n \geqslant N$, then

$$|1 + a| = |-1 - a| \leqslant \frac{1}{2}$$

Using these two expressions and the triangle inequality, we note that

$$\begin{aligned} 2 = |1 - (-1)| &= |1 - (-1) + (a - a)| \\ &= |(1 - a) + (1 + a)| \leqslant |1 - a| + |1 + a| < \frac{1}{2} + \frac{1}{2} < 1 \end{aligned}$$

a contradiction. Thus, the limit of (a_n) does not exist. □

Proposition 3.1.16 (Uniqueness of Limits)**.** *A sequence, if it converges, has a unique limit.*

Proof. Let (a_n) be a convergent sequence, and suppose it converges to real numbers α and β; we wish to show $\alpha = \beta$. Given an $\varepsilon > 0$, we also have $\varepsilon/2 > 0$.

Since $a_n \to \alpha$, there exists an $N_1 \in \mathbb{N}$ such that $|a_n - \alpha| < \dfrac{\varepsilon}{2}$, for all $n \geqslant N_1$.

Since $a_n \to \beta$, there exists an $N_2 \in \mathbb{N}$ such that $|a_n - \beta| < \dfrac{\varepsilon}{2}$, for all $n \geqslant N_2$.

Define $N = \max(N_1, N_2)$, then for any $n \geqslant N$, we note

$$\begin{aligned} |\alpha - \beta| &= |\alpha - \beta + (a_n - a_n)| \\ &= |(a_n - \beta) + (\alpha - a_n)| \\ &\leqslant |a_n - \beta| + |\alpha - a_n| && \text{triangle inequality} \\ &= |a_n - \beta| + |a_n - \alpha| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} && \text{since } n \geqslant N \text{ means } n \geqslant N_1 \text{ and } n \geqslant N_2 \\ &= \varepsilon \end{aligned}$$

That is, $|\alpha = \beta| < \varepsilon$, for every $\varepsilon > 0$. By Problem 17, we conclude that $\alpha = \beta$. □

Proposition 3.1.17 (Convergent Sequences are Bounded)**.** *If (a_n) is a convergent sequence, then (a_n) is bounded.*

Proof. Suppose (a_n) converges to a. Then, in particular, for $\varepsilon = 1$, there exists an $N \in \mathbb{N}$ such that, for all $n \geqslant N$

$$|a_n - a| < 1, \quad \text{and so } \; a - 1 < a_n < a + 1.$$

Set

$$U := \max(a_1, \ldots, a_{N-1}, a + 1)$$

$$L := \min(a_1, \ldots, a_{N-1}, a - 1)$$

Note that for $n < N$, we have $L \leqslant a_n \leqslant U$, since these a_n are included in the list we're taking minimum and maximum over. While, for $n \geqslant N$, we have

$$L \leqslant a - 1 < a_n < a + 1 \leqslant U$$

Thus, for every n, we have obtained $L \leqslant a_n \leqslant U$. Hence (a_n) is bounded. □

Remark 3.1.18. Note that the converse of Proposition 3.1.17 is not true: bounded sequences need not converge. For example, take the sequence in Example 3.1.15. Note that $|a_n| = |(-1)^n| = 1 \leqslant 1$, i.e., (a_n) is bounded. But as we saw in the example, it does not converge.

The contrapositive of Proposition 3.1.17 is useful: *an unbounded sequence diverges*.

3.2. Limit Laws

Lecture 10

Theorem 3.2.1 (Limit Laws for Sequences)**.** *Assume that (a_n) and (b_n) converge to real numbers a and b respectively. Then*

(i) $(a_n + b_n)$ *converges to* $a + b$.

(ii) $(a_n b_n)$ *converges to* ab.

(iii) if $b_n \neq 0$*, for every* n*, and* $b \neq 0$, $\left(\dfrac{1}{b_n}\right)$ *converges to* $\dfrac{1}{b}$.

Proof. We leave (i) as Problem ??; we prove (ii) and (iii).

(ii) We first remark that since (b_n) converges, it is bounded by Proposition 3.1.17, therefore $|b_n| < M$, for some $M > 0$, for every $n \in \mathbb{N}$.

Consider any $\varepsilon > 0$, then $\dfrac{\varepsilon}{2M} > 0$ and $\dfrac{\varepsilon}{2(1+|a|)} > 0$.

Since $a_n \to a$, there exists an $N_1 \in \mathbb{N}$ such that $|a_n - a| < \dfrac{\varepsilon}{2M}$, for all $n \geqslant N_1$.

Since $b_n \to b$, there exists an $N_2 \in \mathbb{N}$ such that $|b_n - b| < \dfrac{\varepsilon}{2(1+|a|)}$, for all $n \geqslant N_2$.

Define $N = \max(N_1, N_2)$, then for any $n \geqslant N$, we note

$$\begin{aligned}
|a_n b_n - ab| &= |(a_n b_n - ab_n) + (ab_n - ab)| \\
&\leqslant |a_n b_n - ab_n| + |ab_n - ab| && \text{triangle inequality} \\
&= |b_n(a_n - a)| + |a(b_n - b)| \\
&= |b_n|\,|a_n - a| + |a|\,|b_n - b| \\
&< M \cdot \frac{\varepsilon}{2M} + |a| \cdot \frac{\varepsilon}{2(1+|a|)} && \text{since } n \geqslant N \geqslant N_1 \text{ and } n \geqslant N \geqslant N_2 \\
&< \frac{\varepsilon}{2} + \frac{|a|}{1+|a|} \cdot \frac{\varepsilon}{2} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon
\end{aligned}$$

Therefore, $a_n b_n \to ab$.

(iii) We first remark that since $b \neq 0$, we have that $|b| > 0$. Since (b_n) converges, then for $\varepsilon = |b|/2 > 0$, there exists an $N_1 \in \mathbb{N}$ such that, for every $n \geqslant N_1$, we have

$$|b_n - b| < \frac{|b|}{2}$$

Using the reverse triangle inequality (Theorem 2.2.11 (T2)), we have $||b_n| - |b|| < |b_n - b|$, and so

$$||b_n| - |b|| < \frac{|b|}{2}$$

$$-\frac{|b|}{2} < |b_n| - |b| < \frac{|b|}{2}$$

$$\frac{|b|}{2} = -\frac{|b|}{2} + |b| < |b_n| < \frac{|b|}{2} + |b| = \frac{3\,|b|}{2}$$

Therefore, we have

$$|b_n| > \frac{|b|}{2}, \quad \text{for every } n \geqslant N_1 \tag{$*$}$$

Consider any $\varepsilon > 0$, then $\dfrac{|b|^2}{2} \cdot \varepsilon > 0$. Since $b_n \to b$, there exists an $N_2 \in \mathbb{N}$ such that

$$|b_n - b| < \frac{|b|^2}{2} \cdot \varepsilon, \quad \text{for all } n \geqslant N_2 \tag{$\star$}$$

Define $N = \max(N_1, N_2)$, then for any $n \geqslant N$, we note

$$\begin{aligned}
\left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b - b_n}{b \cdot b_n}\right| = \frac{|b - b_n|}{|b| \cdot |b_n|} &= \frac{|b_n - b|}{|b| \cdot |b_n|} \\
&< \frac{|b_n - b|}{|b|} \cdot \frac{2}{|b|} && \text{since } n \geqslant N \geqslant N_1 \text{ and } (*) \\
&= |b_n - b| \cdot \frac{2}{|b|^2} \\
&< \frac{|b|^2}{2} \cdot \varepsilon \cdot \frac{2}{|b|^2} && \text{since } n \geqslant N \geqslant N_2 \text{ and } (\star) \\
&= \varepsilon
\end{aligned}$$

Therefore, $\dfrac{1}{b_n} \to \dfrac{1}{b}$.

□

Corollary 3.2.2. *Assume that (a_n) and (b_n) converge to real numbers a and b respectively. Then*

(iv) for any $c \in \mathbb{R}$, the sequence $(c \cdot b_n)$ converges to $c \cdot b$.

(v) $(a_n - b_n)$ converges to $a - b$.

(vi) if $b_n \neq 0$, for every n, and $b \neq 0$, $\left(\frac{a_n}{b_n}\right)$ converges to $\frac{a}{b}$.

Proof. Letting (a_n) be a constant sequence, given as $a_n = c$, for $c \in \mathbb{R}$, we obtain (iv) from Theorem 3.2.1 (ii). In particular, letting $c = -1$, we obtain that $-b_n \to -b$, then using Theorem 3.2.1 (i) we conclude (v) that $(a_n - b_n) = (a_n + (-b_n))$ converges to $a + (-b) = a - b$. Finally, (vi) follows immediately from Theorem 3.2.1 (ii) and (iii). □

Theorem 3.2.3 (Squeeze Theorem for Sequences)**.** *Assume $a_n \leqslant x_n \leqslant b_n$ for every $n \in \mathbb{N}$. If (a_n) and (b_n) are such that*

$$a_n \to L \quad \textit{and} \quad b_n \to L,$$

then $x_n \to L$.

Proof. Let $\varepsilon > 0$, then

since $a_n \to L$, there exists some $N_1 \in \mathbb{N}$ such that for any $n \geqslant N_1$ we have

$$|a_n - L| < \varepsilon, \quad \text{and so } \; L - \varepsilon < a_n < L + \varepsilon$$

since $b_n \to L$, there exists some $N_2 \in \mathbb{N}$ such that for any $n \geqslant N_2$ we have

$$|b_n - L| < \varepsilon, \quad \text{and so } \; L - \varepsilon < b_n < L + \varepsilon$$

Let $N = \max(N_1, N_2)$, and then for $n \geqslant N$, we have

$$L - \varepsilon < a_n \leqslant x_n \leqslant b_n < L + \varepsilon$$

as $n \geqslant N_1$ and $n \geqslant N_2$. Giving us,

$$|x_n - L| < \varepsilon$$

Therefore $L = \lim_{n\to\infty} x_n = L$. □

Example 3.2.4. For any $p \in \mathbb{N}$, we have $\lim\limits_{n\to\infty} \dfrac{1}{n^p} = 0$.

Proof. We note that for any $n \in \mathbb{N}$, we have

$$0 \leqslant \frac{1}{n^p} \leqslant \frac{1}{n}$$

We note that $\lim\limits_{n\to\infty} 0 = \lim\limits_{n\to\infty} \dfrac{1}{n} = 0$. Therefore, Theorem 3.2.3 tells us that $\lim\limits_{n\to\infty} \dfrac{1}{n^p} = 0$. □

Definition 3.2.5 (Monotone Sequences)**.** Given a sequence (a_n), we say

- (a_n) is monotonically increasing if $a_n \leqslant a_{n+1}$, for every $n \in \mathbb{N}$.
- (a_n) is monotonically decreasing if $a_n \geqslant a_{n+1}$, for every $n \in \mathbb{N}$.

A sequence is said to be monotone (or monotonic) if it is either monotonically increasing *or* monotonically decreasing.

Theorem 3.2.6 (Monotone Convergence Theorem)**.** *Suppose a sequence is monotone (a_n), then (a_n) converges if and only if it bounded. Moreover,*

- *if (a_n) is monotonically increasing, then either $a_n \to \infty$ or*

$$\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n.$$

- *if (b_n) is monotonically decreasing, then either $b_n \to -\infty$ or*

$$\lim_{n\to\infty} b_n = \inf_{n\in\mathbb{N}} b_n.$$

Proof. Assume that (a_n) is monotonically increasing. Let's suppose first that (a_n) is not bounded. We show that $a_n \to \infty$. Since, (a_n) is not bounded, then for every $M > 0$, there exists some $N \in \mathbb{N}$ such that $a_N > M$. Since (a_n) is monotonically increasing, for any $n \geqslant N$, we have $a_n \geqslant a_N > M$. Therefore, by definition, $a_n \to \infty$.

Now, suppose that (a_n) is bounded, i.e., the set $\{a_n \,:\, n \in \mathbb{N}\}$ is bounded. Then, by completeness of $\mathbb{R}$, the supremum of this set exists, let's call it α. We now wish to prove that $\alpha = \lim_{n\to\infty} a_n$, i.e., we wish to show that every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|a_n - \alpha| < \varepsilon$, for any $n \geqslant N$.

Consider any $\varepsilon > 0$, since α is the supremum, $\alpha - \varepsilon$ is not an upper bound of the set $\{a_n \,:\, n \in \mathbb{N}\}$. Therefore, there exists an $N \in \mathbb{N}$ such that $\alpha - \varepsilon < a_N$. Since (a_n) is monotonically increasing, we have for $n \geqslant N$, we have $a_n \geqslant a_N$; this gives us that $\alpha - \varepsilon < a_N \leqslant a_n$ for any $n \geqslant N$. Now, since α is the supremum of $\{a_n \,:\, n \in \mathbb{N}\}$, hence $a_n \leqslant \alpha < \alpha + \varepsilon$, for any n.

Together, for any $n \geqslant N$, we obtain that

$$\alpha - \varepsilon < a_n < \alpha + \varepsilon, \quad \text{and so} \quad -\varepsilon < a_n - \alpha < \varepsilon$$

Thus, $|a_n - \alpha| < \varepsilon$, for any $n \geqslant N$. Therefore, $\lim_{n\to\infty} a_n = \alpha$.

The case of monotonically decreasing sequence (b_n) follows by applying the above result to the sequence $(-b_n)$, which will be monotonically increasing. □

3.3. Subsequences

Definition 3.3.1 (Subsequences)**.** Let (a_n) be a sequences and let

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

be an increasing sequence of natural numbers. Then,

$$(a_{n_1}, a_{n_2}, a_{n_3}, \ldots, a_{n_k}, \ldots)$$

is called a subsequence of (a_n), and is denoted $(a_{n_k})_k$ or (a_{n_k}).

We remark that, necessarily, we always have $n_k \geqslant k$.

Example 3.3.2. Some examples.

- Consider the sequence (a_n) given by $a_n = \dfrac{1}{n}$. Then,

  (a) $\left(1, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \dots\right)$ is a subsequence where

  $$n_1 = 1, \quad n_2 = 3, \quad n_3 = 5, \quad n_4 = 7, \quad \dots, \quad n_k = 2k-1, \quad \dots$$

  That is, we may write this subsequence as (a_{2k-1}).

  (b) $\left(\frac{1}{10}, \frac{1}{100}, \frac{1}{1000}, \frac{1}{10000}, \dots\right)$ is a subsequence where

  $$n_1 = 10, \quad n_2 = 100, \quad n_3 = 1000, \quad n_4 = 10000, \quad \dots, \quad n_k = 10^k, \quad \dots$$

  That is, we may write this subsequence as (a_{10^k}).

  (c) $\left(\frac{1}{2}, \frac{1}{25}, \frac{1}{16129}, \frac{1}{100007}, \dots\right)$ is a subsequence where

  $$n_1 = 2, \quad n_2 = 25, \quad n_3 = 16129, \quad n_4 = 100007, \quad \dots$$

  (d) $\left(1, \frac{1}{5}, \frac{1}{3}, \frac{1}{7}, \frac{1}{2}, \dots\right)$ is *not* a subsequence since we do not have $n_1 < n_2 < n_3 < \cdots$.

  (e) $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right)$ is *not* a subsequence since we do not have $n_1 < n_2 < n_3 < \cdots$.

- Consider the sequence (a_n) given by $a_n = n^2$. Then,

  (a) $(4, 16, 36, 64, \dots)$ is a subsequence where

  $$n_1 = 1, \quad n_2 = 4, \quad n_3 = 6, \quad n_4 = 8, \quad \dots, \quad n_k = 2k, \quad \dots$$

  That is, we may write this subsequence as (a_{2k}).

  (b) $(1, 25, 81, 169, \dots)$ is a subsequence where

  $$n_1 = 1, \quad n_2 = 5, \quad n_3 = 9, \quad n_4 = 13, \quad \dots, \quad n_k = 4k-3, \quad \dots$$

  That is, we may write this subsequence as (a_{4k-3}).

  (c) $(1, 121, 144, 484, 10201, \dots)$ is a subsequence where

  $$n_1 = 1, \quad n_2 = 11, \quad n_3 = 12, \quad n_4 = 22, \quad n_5 = 101, \quad \dots$$

Lecture 11

Proposition 3.3.3 (Subsequences and Convergent Sequences)**.** *A sequence (a_n) to $a \in \mathbb{R}$ if and only if every subsequence of (a_n) also converges to a.*

Proof. ($\Rightarrow$) Suppose (a_n) converges to a, and consider a subsequence (a_{n_k}) of a_n. Consider $\varepsilon > 0$, then there exists an $N \in \mathbb{N}$ such that

$$|a_n - a| < \varepsilon, \quad \text{for every } n \geqslant N \tag{$*$}$$

We wish to prove that $a_{n_k} \to a$, i.e., produce an $N' \in \mathbb{N}$ such that $|a_{n_k} - a| < \varepsilon$ for every $k > N'$.

Since $n_k \geqslant k$; therefore, letting $N' = N$, we see that for every $k > N$ we obtain $n_k \geqslant k > N$ and so, by $(*)$,

$$|a_{n_k} - a| < \varepsilon$$

Thus, $a_{n_k} \to a$.

$(\Leftarrow)$ Suppose every subsequence of (a_n) also converges to a. Since (a_n), in particular, is a subsequence of itself, we conclude that $a_n \to a$. □

Corollary 3.3.4. *Suppose (a_n) has subsequences (a_{n_k}) and (a_{n_ℓ}) such that $a_{n_k} \to \alpha$ and $a_{n_\ell} \to \beta$, where $\alpha \neq \beta$; that is, this pair of subsequences converge to different limits. Then (a_n) diverges.*

Proof. Suppose, for the sake of contradiction, assume (a_n) converges to $L \in \mathbb{R}$. Then, by Proposition 3.3.3, every subsequence of (a_n) also converges to L. In particular, $a_{n_k} \to L$ and $a_{n_\ell} \to L$. By uniqueness of limit (Proposition 3.1.16), we conclude $\alpha = L = \beta$. Contradicting our assumption that $\alpha \neq \beta$, and hence (a_n) diverges. □

Example 3.3.5. Using Corollary 3.3.4, we may give an alternate proof of the divergence of the sequence (a_n), where $a_n = (-1)^n$. We consider two subsequences of (a_n), one is (a_{2k}) and the other is (a_{2k-1}). Note that (a_{2k}) is a constant sequence, since $a_{2k} = 1$, and therefore converges to 1; the other subsequence (a_{2k-1}) is also a constant sequence, since $a_{2k} = -1$, and therefore converges to -1. Since $1 \neq -1$, necessarily (a_n) diverges, using Corollary 3.3.4.

The question we ask ourselves is: given a sequence, how nice of a subsequence is that sequence guaranteed to have?

Lemma 3.3.6. *Every sequence has a monotone subsequence.*

Proof. Consider a sequence (a_n); we call a term a_n a *peak* if $a_n \geqslant a_m$ for all $m > n$. Either there are infinitely many peaks or finitely many.

Assume first that our sequence has infinitely many peaks. Define a_{n_k} to be the k^{th} peak; then, by definition of peaks, we have $a_{n_k} \geqslant a_{n_{k+1}}$. Therefore, (a_{n_k}) is a montonically decreasing sequence.

Now, assume that our sequence has finitely many peaks, denoted $a_{p_1}, a_{p_2}, \ldots, a_{p_m}$, where $1 \leqslant p_1 < p_2 < \cdots < p_m$. Let $n_1 = p_m + 1$, then a_{n_1} is not a peak seeing as how a_{p_m} was the last peak. Therefore, there exists an $n_2 > n_1$ such that $a_{n_1} < a_{n_2}$. Now, since $n_2 > n_1 > p_m$, hence a_{n_2} is not a peak; thus, there exists an $n_3 > n_2$ such that $a_{n_2} < a_{n_3}$. Once again, since $n_3 > n_2 > n_1 > p_m$, hence a_{n_3} is not a peak; thus, there exists an $n_4 > n_3$ such that $a_{n_3} < a_{n_4}$.

Continuing inductively, we have thus constructed a subsequence (a_{n_k}) which is monotonically increasing.

In either case, we have produced a monotone subsequence (a_{n_k}) of (a_n). □

Theorem 3.3.7 (Bolzano-Weierstrass Theorem)**.** *Every bounded sequence has a convergent subsequence.*

Proof. Assume (a_n) is a bounded sequence. Using Lemma 3.3.6, we know that it has a monotone subsequence (a_{n_k}); since (a_n) is bounded, so is (a_{n_k}). Therefore, (a_{n_k}) is a monotone bounded sequence, and hence using Theorem 3.2.6 we conclude that (a_{n_k}) converges. □

3.4. Cauchy Sequences

Definition 3.4.1 (Cauchy Sequences)**.** A sequence (a_n) is Cauchy if *for all* $\varepsilon > 0$*, there exists an* $N \in \mathbb{N}$ *such that*

$$|a_m - a_n| < \varepsilon,$$

for all $m, n \geqslant N$*.*

Example 3.4.2. Show that the sequence (a_n) with $a_n = 7 + \dfrac{1}{n}$ is Cauchy.

Solution. We begin with some scratch work.

> *Scratch Work.* Given an arbitrary $\varepsilon > 0$, we need to find a specific N such that whenever $m, n \geqslant N$ we have $|a_m - a_n| < \varepsilon$. We first observe that
>
> $$|a_m - a_n| = \left|\left(7 + \frac{1}{m}\right) - \left(7 + \frac{1}{n}\right)\right| = \left|\frac{1}{m} + \frac{-1}{n}\right| \leqslant \left|\frac{1}{m}\right| + \left|\frac{-1}{n}\right| = \frac{1}{m} + \frac{1}{n}$$
>
> and since $m, n \geqslant N$, we have
>
> $$\frac{1}{m} < \frac{1}{n} \leqslant \frac{1}{N}$$
>
> Therefore, if we can find an N such that $\dfrac{1}{N} < \dfrac{\varepsilon}{2}$, we will obtain
>
> $$|a_m - a_n| \leqslant \frac{1}{m} + \frac{1}{n} \leqslant \frac{1}{N} + \frac{1}{N} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$
>
> The existence of this N is guarantee by the Archiemedean property.

Fix any $\varepsilon > 0$, then by the Archimedean property there exists an $N \in \mathbb{N}$ such that $1/N < \varepsilon/2$. Then, for any $m, n \geqslant N$, we get

$$\begin{aligned} |a_m - a_n| = \left|\left(7 + \frac{1}{m}\right) - \left(7 + \frac{1}{n}\right)\right| &= \left|\frac{1}{m} + \frac{-1}{n}\right| \\ &\leqslant \left|\frac{1}{m}\right| + \left|\frac{1}{n}\right| = \frac{1}{m} + \frac{1}{n} \leqslant \frac{1}{N} + \frac{1}{N} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Therefore, we have proved (a_n) is a Cauchy sequence. □

Lecture 12

Lemma 3.4.3 (Cauchy Sequences are Bounded)**.** *If* (a_n) *is Cauchy, then* (a_n) *is bounded.*

Proof. Suppose (a_n) is Cauchy. Then, in particular, for $\varepsilon = 1$, there exists an $N \in \mathbb{N}$ such that

$$|a_m - a_n| < 1,$$

for all $m, n \geqslant N$. Therefore, letting $m = N + 1$, we have that for every $n \geqslant N$

$$|a_n - a_{N+1}| < \varepsilon$$

Hence, we obtain

$$a_{N+1} - 1 < a_n < a_{N+1} + 1$$

Set

$$U \coloneqq \max(a_1, \ldots, a_{N-1}, a_{N+1} + 1) \qquad L \coloneqq \min(a_1, \ldots, a_{N-1}, a_{N+1} - 1)$$

Note that for $n < N$, we have $L \leqslant a_n \leqslant U$, since these a_n are included in the list we're taking minimum and maximum over. While, for $n \geqslant N$, we have

$$L \leqslant a_{N+1} - 1 < a_n < a_{N+1} + 1 \leqslant U$$

Thus, for every n, we have obtained $L \leqslant a_n \leqslant U$. Hence (a_n) is bounded. □

Theorem 3.4.4 (Cauchy Criterion for Convergence)**.** *A sequence converges if and only if it is Cauchy.*

Proof. ($\Rightarrow$) Suppose (a_n) converges to some $a \in \mathbb{R}$. Let $\varepsilon > 0$; since $\varepsilon/2 > 0$, there exists an $N \in \mathbb{N}$ such that for any $n \geqslant N$, we have

$$|a_n - a| < \frac{\varepsilon}{2}$$

Therefore, for all $m, n \geqslant N$,

$$\begin{aligned} |a_m - a_n| = |a_m - a_n + (a - a)| &= |(a_m - a) + (a - a_n)| \\ &\leqslant |a_m - a| + |a - a_n| \\ &= |a_m - a| + |a_n - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Hence, (a_n) is Cauchy.

($\Leftarrow$) Suppose (a_n) is Cauchy, then by Lemma 3.4.3 we have that (a_n) is bounded. Therefore, by the Bolzano-Weierstrass Theorem (Theorem 3.3.7), (a_n) has a convergent subsequence. Let (a_{n_k}) be this convergent subsequence, converging to some $a \in \mathbb{R}$. We will show $a_n \to a$.

Let $\varepsilon > 0$; since $\varepsilon/2 > 0$ and (a_n) is Cauchy, there exists some N_1 such that

$$|a_m - a_n| < \frac{\varepsilon}{2}, \quad \text{for all } m, n \geqslant N_1$$

Moreover, since $\varepsilon/2 > 0$ and $a_{n_k} \to a$, there exists some N_2 such that

$$|a_{n_k} - a| < \frac{\varepsilon}{2}, \quad \text{for all } k \geqslant N_2$$

Let $M = \max(N_1, N_2)$, then for any $n \geqslant M$, we note that

$$\begin{aligned} |a_n - a| &= |a_n - a + (a_{n_M} - a_{n_M})| \\ &= |(a_n - a_{n_M}) + (a_{n_M} - a)| \\ &\leqslant |a_n - a_{n_M}| + |a_{n_M} - a| \\ &= |a_n - a_{n_M}| + |a_{n_M} - a| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \qquad \text{as } n, n_M \geqslant M \geqslant N_1 \text{ and } n_M \geqslant M \geqslant N_2 \\ &= \varepsilon \end{aligned}$$

Thus, $a_n \to a$. □

4. Series

4.1. Series & Convergence Tests

Definition 4.1.1 (Series)**.** A sum of infinitely many numbers is called an infinite series, which we write as

$$\sum_{k=1}^{\infty} a_k$$

- The numbers a_k are the terms of the series.
- The sequence of partial sums is the sequence (s_n) where

$$s_n = \sum_{k=1}^{n} a_k$$

- The series converges to $L \in \mathbb{R}$, written $\sum_{k=1}^{\infty} a_k = L$, if $s_n \to L$, and diverges if (s_n) does.
- We say the series is bounded or monotone if (s_n) is.

Corollary 4.1.2 (Limit Laws for Series)**.** *Assume that* $\sum_{k=1}^{\infty} a_k = \alpha$ *and* $\sum_{k=1}^{\infty} b_k = \beta$*. Then,*

(i) $\sum_{k=1}^{\infty} (a_k + b_k) = \alpha + \beta.$

(ii) $\sum_{k=1}^{\infty} (a_k - b_k) = \alpha - \beta.$

(iii) for any $c \in \mathbb{R}$, $\sum_{k=1}^{\infty} (c \cdot a_k) = c \cdot \alpha.$

Proof. This is a consequence of applying Theorem 3.2.1 and Corollary 3.2.2 to the relevant sequence of partial sums. □

Remark 4.1.3. We note that no rule like the one stated in Theorem 3.2.1 for product of sequences holds for product of series. You should try and convince yourself why this is the case.

Proposition 4.1.4 (Term Divergence Test)**.** *If* $a_k \not\to 0$*, then* $\sum_{k=1}^{\infty} a_k$ *diverges.*

Proof. We will prove the contrapositive: if $\sum_{k=1}^{\infty}$ converges, then $a_k \to 0$.

Let (s_n) be the sequence of partial sums, that is,

$$s_n = \sum_{k=1}^{n} a_k.$$

Assume that the series converges, that is, (s_n) converges. Say, $s_n \to L$, and therefore any subsequence of (s_n) also converges to L, in particular, $s_{n+1} \to L$. Observe that,

$$\begin{aligned}
\lim_{n\to\infty} a_{n+1} &= \lim_{n\to\infty} \left(\sum_{k=1}^{n+1} a_k - \sum_{k=1}^{n} a_k \right) \\
&= \lim_{n\to\infty} (s_{n+1} - s_n) \\
&= L - L && \text{limit laws} \\
&= 0
\end{aligned}$$

Therefore, the subsequence (a_{n+1}) of (a_k) converges to 0. We prove that, in fact, $a_k \to 0$. For an $\varepsilon > 0$, since $a_{n+1} \to 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_{n+1} - 0| < \varepsilon, \quad \text{for all } n \geqslant N$$

Then, for any $k \geqslant N+1$, we observe that $k-1 \geqslant N$, and hence

$$|a_k - 0| = \left| a_{(k-1)+1} - 0 \right| < \varepsilon$$

for all $k \geqslant N+1$. Thus, $a_k \to 0$. □

Definition 4.1.5 (Geometric Series)**.** A geometric series is a series of the form

$$\sum_{k=0}^{\infty} ar^k = a + ar + ar^2 + ar^3 + \cdots$$

where $a, r \in \mathbb{R}$.

Lemma 4.1.6. *For any $r \in \mathbb{R}$, the sequence $(r^n)_n$ has the following behaviour.*

$$\begin{cases} (r^n) \textit{ converges when } r \in (-1,1]\textit{; if } r \in (-1,1) \textit{ then } r^n \to 0\textit{, and if } r = 1\textit{, then } r^n \to 1. \\ (r^n) \textit{ diverges when } r \notin (-1,1] \end{cases}$$

Proof. Let $r \in \mathbb{R}$, we will divide our problem in multiple cases.

CASE I. $r = 1$

The sequence (r^n) is the constant sequence $(1, 1, 1, \ldots)$, and therefore converges to 1.

CASE II. $r = -1$

The sequence (r^n) is the sequence $((-1)^n)$, which we know diverges.

CASE III. $0 \leqslant r < 1$

Since $0 \leqslant r < 1$, we have $r^{n+1} \leqslant r^n$ and therefore the sequence (r^n) is monotonically decreasing. Moreover, we have that $0 \leqslant r^n < 1$ for every $n \in \mathbb{N}$, and so (r^n) is bounded. Hence, (r^n) is a bounded monotone sequence, and thus converges by the Monotone Convergence Theorem

(Theorem 3.2.6); assume $r^n \to L$. Thus, any of its subsequences also converge to L, in particular, the subsequence $(r^{n+1}) = (r^2, r^3, \ldots)$.

Note that, we have $r^{n+1} = r \cdot r^n$. Then, using limit laws, we may observe that

$$L = \lim_{n\to\infty} r^{n+1} = \lim_{n\to\infty} r \cdot r^n = r \cdot \lim_{n\to\infty} r^n = r \cdot L$$

This gives us $(r-1) \cdot L = 0$. Now, as $r < 1$, we necessarily have $L = 0$; and thus $r^n \to 0$.

CASE IV. $-1 < r < 0$

Since $-1 < r < 0$, we have $0 < |r| < 1$. Therefore, by Case III, we have that $|r|^n \to 0$. Using limit laws, we may further conclude that $-|r|^n \to 0$.

Consider any $n \in \mathbb{N}$. Suppose n is odd, then $n = 2k-1$ for some $k \in \mathbb{N}$ and

$$-|r|^{2k-1} = -\left|r^{2k-1}\right| < r^{2k-1} < 0$$

Since $-|r|^n \to 0$, then for its subsequence $(-|r|^{2k-1})$ we also have that $-|r|^{2k-1} \to 0$. Using the Squeeze Theorem (Theorem 3.2.3), we may conclude $r^{2k-1} \to 0$.

Suppose n is even, then $n = 2k$ for some $k \in \mathbb{N}$ and

$$0 < r^{2k} < \left|r^{2k}\right| = |r|^{2k}$$

Since $|r|^n \to 0$, then for its subsequence $(|r|^{2k})$ we also have that $|r|^{2k} \to 0$. Using the Squeeze Theorem (Theorem 3.2.3), we may conclude $r^{2k} \to 0$.

Hence, the subsequences (r^{2k-1}) and (r^{2k}) both converge to 0, then by Problem ?? we conclude that $r^n \to 0$.

CASE V. $r > 1$

Since $r > 1$, we observe that $0 < 1/r < 1$ and therefore by Case III we have that $1/r^n \to 0$. We wish to prove that $r^n \to \infty$.

Consider any $M > 0$, then since $1/M > 0$, there exists an $N \in \mathbb{N}$ such that

$$\frac{1}{r^n} = \left|\frac{1}{r^n} - 0\right| < \frac{1}{M}, \quad \text{for all } n \geqslant N$$

which gives us

$$r^n > M, \quad \text{for all } n \geqslant N$$

Hence, we have proven that $r^n \to \infty$.

CASE VI. $r < -1$

Suppose $r^n \to L$, for some $L \in \mathbb{R}$. Then, so do the subsequences (r^{2k}) and (r^{2k-1}).

Let us focus on the subsequence (r^{2k}), we note that the sequence is monotonically increasing. Then, for $r^2 > 0$, there exists an $N_1 \in \mathbb{N}$ such that

$$\left|r^{2k} - L\right| < r^2, \quad \text{for all } k \geqslant N_1, \qquad \text{and so } -r^2 < r^{2k} - L < r^2$$

Therefore, $r^{2k} - L < r^2$ and hence $L > r^{2k} - r^2 > 0$, since (r^{2k}) is monotonically increasing.

Let us focus on the subsequence (r^{2k-1}), we note that the sequence is monotonically decreasing. Then, for $-r > 0$, there exists an $N_2 \in \mathbb{N}$ such that

$$\left|r^{2k-1} - L\right| < -r, \quad \text{for all } k \geqslant N_2, \qquad \text{and so } r = -(-r) < r^{2k-1} - L < -r$$

Therefore, $r < r^{2k-1} - L$ and hence $L < r^{2k-1} - r < 0$, since (r^{2k-1}) is monotonically decreasing.

We have thus arrived at a contradiction as we cannot have that $L > 0$ and $L < 0$.

Therefore, the limit of (r^n) does not exist.

□

Lecture 13

Proposition 4.1.7 (Geometric Series). *Assume $r \neq 0$, then*

$$\sum_{k=0}^{\infty} r^k = \begin{cases} \dfrac{1}{1-r} & \text{if } |r| < 1 \\ \text{diverges} & \text{if } |r| \geqslant 1 \end{cases}$$

Proof. By Lemma 4.1.6, we observe that $r^n \not\to 0$ whenever $|r| \geqslant 1$; therefore, by the Term Divergence Test (Proposition 4.1.4), the series diverges.

Now, let us assume that $|r| < 1$. Let (s_n) be the sequence of partial sum of $\sum_{k=0}^{\infty} r^k$. Note,

$$\begin{aligned}(1-r)\cdot s_n = (1-r)\cdot \sum_{k=0}^{n} r^k &= (1-r)\cdot(1 + r + r^2 + \cdots + r^n) \\ &= (1 + r + r^2 + \cdots + r^n) - r\cdot(1 + r + \cdots + r^{n-1} + r^n) \\ &= (1 + r + r^2 + \cdots + r^n) - (r + r^2 + \cdots + r^n + r^{n+1}) \\ &= 1 - r^{n+1}\end{aligned}$$

$$s_n = \frac{1 - r^{n+1}}{1-r},$$

By Lemma 4.1.6, we have $r^n \to 0$ and so $r^{n+1} \to 0$, as (r^{n+1}) is a subsequence of (r^n). Hence,

$$\sum_{k=0}^{\infty} r^k = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1 - r^{n+1}}{1-r} = \frac{1-0}{1-r} = \frac{1}{1-r}$$

as we wanted to show. □

Corollary 4.1.8 (General Geometric Series). *Assume $a, r \neq 0$, then*

$$\sum_{k=0}^{\infty} ar^k = \begin{cases} \dfrac{a}{1-r} & \text{if } |r| < 1 \\ \text{diverges} & \text{if } |r| \geqslant 1 \end{cases}$$

Proof. This follows from Proposition 4.1.7 and Corollary 4.1.2 (iii). □

Proposition 4.1.9. *If $a_k \geqslant 0$, for every $k \in \mathbb{N}$, then either $\sum_{k=0}^{\infty} a_k$ either converges or diverges to ∞.*

Proof. Let $s_n = \sum_{k=0}^{n} a_k$ be the n^{th} partial sum, then since $a_k \geqslant 0$, we note that $s_{n+1} \geqslant s_n$. Therefore (s_n) is a monotonically increasing sequence. The result then follows from Monotone Convergence Theorem (Theorem 3.2.6). □

Proposition 4.1.10 (Comparison Test)**.** *Assume $0 \leqslant a_k \leqslant b_k$ for all k.*

(i) If $\sum_{k=1}^{\infty} b_k$ converges, then $\sum_{k=1}^{\infty} a_k$ converges. *(ii) If $\sum_{k=1}^{\infty} a_k$ diverges, then $\sum_{k=1}^{\infty} b_k$ diverges.*

Proof. Let (s_n) and (t_n) be the sequence of partial sums of $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ respectively. Since $0 \leqslant a_k \leqslant b_k$ for every k, we note that

$$0 \leqslant s_n = a_1 + \cdots + a_n \leqslant b_1 + \cdots + b_n = t_n, \quad \text{for all } n \tag{$*$}$$

We shall prove (i); note that by Proposition 4.1.9 and the Monotone Convergence Theorem, it suffices to prove that if (t_n) converges then (s_n) is bounded. Since (t_n) converges, we can conclude that (t_n) is bounded (Proposition 3.1.17), say, by M. That is, for every n, we have $t_n \leqslant M$. Hence, by $(*)$, we obtain

$$0 \leqslant s_n \leqslant M, \quad \text{for every } n.$$

Thus, (s_n) is bounded, and since (s_n) is monotonically increasing, we conclude (s_n) converges.

(ii) is the contrapositive of (i). □

Corollary 4.1.11 (General Comparison Test)**.** *Assume $0 \leqslant a_k \leqslant b_k$ for all $k \geqslant N$, for some N.*

(i) If $\sum_{k=1}^{\infty} b_k$ converges, then $\sum_{k=1}^{\infty} a_k$ converges. *(ii) If $\sum_{k=1}^{\infty} a_k$ diverges, then $\sum_{k=1}^{\infty} b_k$ diverges.*

Proposition 4.1.12 (Harmonic Series Diverges)**.** *The* Harmonic series $\sum_{k=1}^{\infty} \frac{1}{k}$ *diverges.*

Proof. For any n, we will define

$$x_n := \frac{1}{2^{n-1}+1} + \cdots + \frac{1}{2^{n-1}+2^{n-1}} = \sum_{i=1}^{2^{n-1}} \frac{1}{2^{n-1}+i}$$

That is,

$$x_1 = \frac{1}{2}, \qquad x_2 = \frac{1}{3} + \frac{1}{4}, \qquad x_3 = \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}, \qquad \cdots$$

Therefore, we note that

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + x_1 + x_2 + x_3 + \cdots$$

We now note that

$$x_n = \frac{1}{2^{n-1}+1} + \cdots + \frac{1}{2^{n-1}+2^{n-1}} = \sum_{i=1}^{2^{n-1}} \frac{1}{2^{n-1}+i}$$
$$> \sum_{i=1}^{2^{n-1}} \frac{1}{2^{n-1}+2^{n-1}} = \frac{2^{n-1}}{2^{n-1}+2^{n-1}} = \frac{1}{2}$$

Hence,

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + x_1 + x_2 + x_3 + \cdots > 1 + \sum_{k=1}^{\infty} \frac{1}{2}$$

Since the latter sequence diverges (using the Term Divergence test, for example), and thus the Harmonic series diverges by the Comparison test. □

Proposition 4.1.13 (Series p-Test)**.** *The series* $\sum_{k=1}^{\infty} \frac{1}{k^p}$ *converges if and only if* $p > 1$.

Proof. Note that when $p \leqslant 1$, then $1/k \leqslant 1/k^p$.

Since, by Proposition 4.1.12, the series $\sum_{k=1}^{\infty} 1/k$ converges, therefore the series $\sum_{k=1}^{\infty} 1/k^p$ diverges by the Comparison Test.

Now, assume that $p > 1$. For any n, we will define

$$x_n \coloneqq \frac{1}{(2^n)^p} + \frac{1}{(2^n+1)^p} + \cdots + \frac{1}{(2^n+(2^n-1))^p} = \sum_{i=0}^{2^n-1} \frac{1}{(2^n+i)^p}$$

That is,

$$x_1 = \frac{1}{2^p} + \frac{1}{3^p},$$
$$x_2 = \frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p},$$
$$x_3 = \frac{1}{8^p} + \frac{1}{9^p} + \frac{1}{10^p} + \frac{1}{11^p} + \frac{1}{12^p} + \frac{1}{13^p} + \frac{1}{14^p} + \frac{1}{15^p},$$
$$\vdots$$

Therefore, we note that

$$\sum_{k=1}^{\infty} \frac{1}{k^p} = 1 + x_1 + x_2 + x_3 + \cdots$$

We now note that

$$x_n = \frac{1}{(2^n)^p} + \frac{1}{(2^n+1)^p} + \cdots + \frac{1}{(2^n+(2^n-1))^p} = \sum_{i=0}^{2^n-1} \frac{1}{(2^n+i)^p}$$
$$< \sum_{i=0}^{2^n-1} \frac{1}{(2^n)^p} = \frac{2^n}{(2^n)^p} = \frac{1}{(2^n)^{p-1}}$$

Hence,

$$\sum_{k=1}^{\infty} \frac{1}{k^p} = 1 + x_1 + x_2 + x_3 + \cdots < \sum_{k=0}^{\infty} \left(\frac{1}{2^{p-1}} \right)^k$$

Since $p > 1$, we observe that $1/2^{p-1} < 1$, and therefore the latter sequence converges as a geometric series. Therefore the series

$$\sum_{k=1}^{\infty} \frac{1}{k^p}$$

converges by the Comparison test. □

Remark 4.1.14. The case of immediate interest is the case when $p = 2$, that is, to what real number does the series

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots$$

converge to. This was historically called the *Basel Problem*. Euler, at the age of 24, proved that

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$$

4.2. Absolute Convergence

Lecture 14

Proposition 4.2.1 (Alternating Series Test)**.** *Assume that (a_k) is a monotonically decreasing sequence and $a_k \to 0$. Then the series*

$$\sum_{k=1}^{\infty} (-1)^{k+1} a_k = a_1 - a_2 + a_3 - a_4 + \cdots$$

converges.

Proof. Note that since (a_k) is monotonically decreasing and converging to 0, by the Monotone Convergence Theorem (Theorem 3.2.6), necessarily $0 = \inf\{a_k : k \in \mathbb{N}\}$. Thus, $0 \leqslant a_k$ for every k.

Let (s_n) be the sequence of partial sum of our series, and let us consider the subsequences (s_{2n}) and (s_{2n-1}). We will show that these subsequences are bounded monotone sequences, at which point we may use the Monotone Convergence Theorem.

Consider the sequence (s_{2n}), we have

$$s_{2n} = (a_1 - a_2) + \cdots + (a_{2n-1} - a_{2n})$$

As our sequence (a_k) is monotonically decreasing, we have $a_i \geqslant a_{i+1}$ for every i, i.e. $a_i - a_{i+1} \geqslant 0$ for every $i \in \mathbb{N}$. So,

$$s_{2n+2} = (a_1 - a_2) + \cdots + (a_{2n-1} - a_{2n}) + (a_{2n+1} - a_{2n+2}) \geqslant (a_1 - a_2) + \cdots + (a_{2n-1} - a_{2n}) = s_{2n}$$

and thus (s_{2n}) is monotonically increasing. Moreover, note that

$$\begin{aligned} s_{2n} &= (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2n-1} - a_{2n}) \\ &= a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n} \leqslant a_1 \end{aligned}$$

since $a_i - a_{i+1} \geqslant 0$ and $a_i \geqslant 0$ for every i; and so (s_n) is a bounded above sequence. Therefore, (s_{2n}) converges, by the Monotone Convergence Theorem; suppose $s_{2n} \to U$.

Consider the sequence (s_{2n-1}). A similar argument shows that (s_{2n-1}) is a monotonically decreasing sequence and bounded below by $a_1 - a_2$. Therefore, (s_{2n-1}) converges, by the Monotone Convergence Theorem; suppose $s_{2n-1} \to V$.

Now, note that

$$\begin{aligned} U - V &= \lim_{n\to\infty} (s_{2n} - s_{2n-1}) \\ &= \lim_{n\to\infty} \left(\sum_{k=1}^{2n} (-1)^{k+1} a_k - \sum_{k=1}^{2n-1} (-1)^{k+1} a_n \right) \\ &= \lim_{n\to\infty} (-1)^{2n+1} a_{2n} \\ &= \lim_{n\to\infty} -a_{2n} \\ &= -\lim_{n\to\infty} a_{2n} \\ &= 0 \end{aligned}$$

the latter is true since $a_k \to 0$ by assumption and (a_{2n}) is a subsequence of (a_k). Therefore, $U = V$, and hence both subsequences (s_{2n}) and (s_{2n-1}) converge to U. Thus, by Problem ?? (from HW 4), we conclude that (s_n) converges. □

Proposition 4.2.2. *If* $\sum_{k=1}^{\infty} |a_k|$ *converges, then* $\sum_{k=1}^{\infty} a_k$ *converges.*

Proof. Let (s_n) and (t_n) be the sequence of partial sum of $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} |a_k|$ respectively.

We first observe that for $n, m \in \mathbb{N}$ for $n > m$ we observe that

$$\begin{aligned} &|s_n - s_m| \\ &\quad = \left| \sum_{k=1}^{n} a_k - \sum_{k=1}^{m} a_k \right| \\ &\quad = \left| \sum_{k=m+1}^{n} a_k \right| \\ &\quad \leqslant \sum_{k=m+1}^{n} |a_k| \\ &\quad = \sum_{k=1}^{n} |a_k| - \sum_{k=1}^{m} |a_k| \\ &\quad = t_n - t_m \\ &\quad \leqslant |t_n - t_m| \qquad (*) \end{aligned}$$

Now, since by assumption (t_n) converges, by the Cauchy Criterion (Theorem 3.4.4), this sequence is Cauchy. We will show that (s_n) is also Cauchy. For any $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for any $n, m \geqslant N$ we have

$$|t_n - t_m| < \varepsilon$$

By $(*)$, we conclude that for any $n, m \geqslant N$, we have

$$|s_n - s_m| < |t_n - t_m| < \varepsilon$$

Hence, (s_n) is Cauchy. By the Cauchy Criterion (Theorem 3.4.4), (s_n) converges. □

Remark 4.2.3. We note that the converse of Proposition 4.2.2 does not hold. Consider the *alternating Harmonic series*

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \tag{†}$$

This series converges by the Alternating Series test (Proposition 4.2.1), and converges to $\ln 2$. But

$$\sum_{k=1}^{\infty} \left| (-1)^{k+1} \frac{1}{k} \right| = \sum_{k=1}^{\infty} \frac{1}{k}$$

is the Harmonic series, which we know does not converge.

Moreover it is not true that when the hypothesis of Proposition 4.2.2 satisfied, that the series

$$\sum_{k=1}^{\infty} |a_k| \quad \text{and} \quad \sum_{k=1}^{\infty} a_k$$

converge to the same real number. For example, consider the series

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)} \tag{⋆}$$

The corresponding series with absolute terms is the series

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} \tag{$*$}$$

Let us first show that the latter series converges. Let (s_n) be the sequence of partial sum of $(*)$, the note that

$$\begin{aligned} s_n &= \sum_{k=1}^{n} \frac{1}{k(k+1)} \\ &= \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1} \right) \\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= 1 - \frac{1}{n+1} \end{aligned}$$

Therefore,

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \left(1 - \frac{1}{n+1}\right) = 1$$

Therefore $(\star)$ converges by Proposition 4.2.2. We will compute the limit of this series. Let (t_n) be the sequence of partial sum of $(\star)$, then

$$\begin{aligned}
t_n &= \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k(k+1)} \\
&= \sum_{k=1}^{n} (-1)^{k+1} \left(\frac{1}{k} - \frac{1}{k+1}\right) \\
&= \sum_{k=1}^{n} \left(\frac{(-1)^{k+1}}{k} + \frac{(-1)^{k+2}}{k+1}\right) \\
&= \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} + \sum_{k=1}^{n} \frac{(-1)^{k+2}}{k+1} \\
&= \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} + \sum_{k=2}^{n+1} \frac{(-1)^{k+1}}{k} \\
&= \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} + \left(\sum_{k=1}^{n+1} \frac{(-1)^{k+1}}{k} - 1\right) \\
&= u_n + u_{n+1} - 1
\end{aligned}$$

where (u_n) is the sequence of partial sums of the alternating Harmonic series (†). Therefore,

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)} = \lim_{n\to\infty} t_n = \lim_{n\to\infty} (u_n + u_{n+1} - 1) = \ln 2 + \ln 2 - 1 = 2\ln 2 - 1$$

Definition 4.2.4 (Absolute & Conditional Convergence)**.** Consider the series $\sum_{k=1}^{\infty} a_k$.

- If $\sum_{k=1}^{\infty} |a_k|$ converges, then we say $\sum_{k=1}^{\infty} a_k$ converges absolutely.
- If $\sum_{k=1}^{\infty} a_k$ converges but $\sum_{k=1}^{\infty} |a_k|$ diverges, then $\sum_{k=1}^{\infty} a_k$ converges conditionally.

Remark 4.2.5. Proposition 4.2.2 tells us that if a series converges absolutely, the series converges.

Example 4.2.6. The alternating Harmonic series converges conditionally.

While the series

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)}$$

converges absolutely.

Definition 4.2.7 (Rearrangement)**.** A rearrangement of a series $\sum_{k=1}^{\infty} a_k$ is a series

$$\sum_{k=1}^{\infty} b_k$$

for which there is a bijection $f : \mathbb{N} \to \mathbb{N}$ such that $b_k = a_{f(k)}$.

Example 4.2.8. The alternating harmonic series is

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots$$

A rearrangement of this series is

$$\frac{1}{5} - \frac{1}{2} - \frac{1}{152} + \frac{1}{57} - \frac{1}{1847604} - \frac{1}{8} + \cdots$$

A different rearrangement is

$$1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} + \frac{1}{7} + \frac{1}{9} + \frac{1}{11} - \frac{1}{4} + \cdots$$

Proposition 4.2.9. *If a series $\sum_{k=1}^{\infty} a_k$ converges absolutely, then any rearrangement $\sum_{k=1}^{\infty} b_k$ converges, and*

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} b_k$$

Theorem 4.2.10 (Rearrangement Theorem)**.** *If a series*

$$\sum_{k=1}^{\infty} a_k$$

converges conditionally, then for any $L \in \mathbb{R}$ there exists some rearrangement of the series which converges to L, or a rearrangement of the series which diverges to $\pm\infty$.

5. The Topology of Reals

5.1. Open Sets

Lecture 15

Definition 5.1.1 (δ-neighbourhood)**.** For $x \in \mathbb{R}$ and some $\delta > 0$, the open interval

$$(x-\delta, x+\delta) = \{y \in \mathbb{R} \,:\, x-\delta < y < x+\delta\} = \{y \in \mathbb{R} \,:\, |y-x| < \delta\}$$

is called the δ-neighbourhood of x. We denote this set as $B_\delta(x) = (x-\delta, x+\delta)$.

Definition 5.1.2 (Open Sets)**.** A set $U \subseteq \mathbb{R}$ is called open if *for every $x \in U$, there exists a $\delta > 0$ such that $(x-\delta, x+\delta) \subseteq U$.*

That is, a set U is open if for every $x \in U$, there exists a δ-neighbourhood of x, such that $B_\delta(x) \subseteq U$.

Example 5.1.3. Let us look at some examples of open sets.

(a) $\mathbb{R}$ is an open set. For any $x \in \mathbb{R}$, we may take, for example, $\delta = \pi$, and then

$$B_\pi(x) = (x-\pi, x+\pi) \subseteq \mathbb{R}$$

Hence, $\mathbb{R}$ is an open set.

(b) The empty set $\varnothing$ is open. Let us prove this by contradiction; assume $\varnothing$ is not an open set. Therefore, there exists an $x \in \varnothing$ such that for every $\delta > 0$, we have $B_\delta(x) \not\subseteq \varnothing$. But no such x can exist, as the empty set has no elements. Therefore, we have arrived at a contradiction. Hence, $\varnothing$ is an open set.

(c) The *open interval* $(a,b) = \{x \in \mathbb{R} \,:\, a < x < b\}$ is an open set. Consider any $x \in (a,b)$, we wish to produce a $\delta > 0$ such that $(x-\delta, x+\delta) \subseteq (a,b)$.

Take $\delta = \min(x-a, b-x)$, and we claim that $(x-\delta, x+\delta) \subseteq (a,b)$. Consider any $y \in (x-\delta, x+\delta)$, that is, assume $x-\delta < y < x+\delta$;

CASE I. Assume $\delta = x-a$, that is, assume $x-a \leqslant b-x$. Then,

$$x-\delta = x-(x-a) = a$$

$$x+\delta = x+(x-a) \leqslant x+(b-x) = b$$

Therefore, we observe

$$a = x-\delta < y < x+\delta \leqslant b$$

and hence $y \in (a,b)$.

CASE II. Assume $\delta = b-x$, that is, assume $b-x \leqslant x-a$. Then,

$$x-\delta = x-(b-x) \geqslant x-(x-a) = a$$

$$x+\delta = x+(b-x) = b$$

Therefore, we observe

$$a \leqslant x-\delta < y < x+\delta = b$$

and hence $y \in (a,b)$.

Thus, $(x-\delta, x+\delta) \subseteq (a,b)$, and therefore (a,b) is an open set.

(d) The *open rays*

$$(a,\infty) = \{x \in \mathbb{R} \,:\, x > a\} \quad \text{and} \quad (-\infty, a) = \{x \in \mathbb{R} \,:\, x < a\}$$

are open sets.

(e) We now give a non-example: the set $[3,7) = \{x \in \mathbb{R} \,:\, 3 \leqslant x < 7\}$ is not an open set. For $3 \in [3,7)$, and any $\delta > 0$, we have that

$$(3-\delta, 3+\delta) \not\subseteq [3,7)$$

Since, any $y \in (3-\delta, 3+\delta)$ such that $3-\delta < y < 3$ is such that $y \notin [3,7)$. For instance, one may take

$$y = \frac{(3-\delta)+3}{2}$$

Proposition 5.1.4.

(i) If $\{U_\alpha\}_{\alpha \in T}$ is an arbitrary collection of open sets, then $\bigcup_{\alpha \in T} U_\alpha$ is an open set.

(i) If $\{U_1, U_2, \ldots, U_n\}$ is a finite collection of open sets, then $\bigcap_{i=1}^{n} U_i$ is an open set.

Proof. We will prove (i) first. To show $\cup_{\alpha \in T} U_\alpha$ is an open set, we will consider an $x \in \cup_{\alpha \in T} U_\alpha$ and produce a $\delta > 0$ such that $B_\delta(x) \subseteq \cup_{\alpha \in T} U_\alpha$.

Since $x \in \cup_{\alpha \in T} U_\alpha$, therefore there exists an $\alpha_0 \in T$ such that $x \in U_{\alpha_0}$. Since U_{α_0} is an open set, hence there exists a $\delta > 0$ such that $B_\delta(x) \subseteq U_{\alpha_0}$. Thus, $B_\delta(x) \subseteq U_{\alpha_0} \subseteq \cup_{\alpha \in T} U_\alpha$. Therefore, $\cup_{\alpha \in T} U_\alpha$ is open.

Let us now prove (ii). To show $\cap_{i=1}^{n} U_i$ is an open set, we will consider an $x \in \cap_{i=1}^{n} U_i$ and produce a $\delta > 0$ such that $B_\delta(x) \subseteq \cap_{i=1}^{n} U_i$.

Since $x \in \cap_{i=1}^{n} U_i$, therefore $x \in U_i$, for every $1 \leqslant i \leqslant n$. Since each U_i is an open set, hence there exists a $\delta_i > 0$ such that $B_{\delta_i}(x) \subseteq U_i$. Define $\delta = \min(\delta_1, \delta_2, \ldots, \delta_n)$, then $\delta \leqslant \delta_i$ for every i. Then, for any $y \in B_\delta(x)$, we have

$$x - \delta_i \leqslant x - \delta < y < x + \delta \leqslant x + \delta_i$$

for every i. Thus, $B_\delta(x) \subseteq B_{\delta_i}(x) \subseteq U_i$, and therefore $B_\delta(x) \subseteq \cap_{i=1}^{n} U_i$. Hence, $\cap_{i=1}^{n} U_i$ is open. □

Remark 5.1.5. We remark on the notion of a *union of an arbitrary collection of sets*. Given an arbitrary collection of sets $\{X_\alpha\}_{\alpha \in T}$, means that there are as many sets in this collection as there are elements in T. For each element of $\alpha \in T$, there is a corresponding set X_α. The union

$$\bigcup_{\alpha \in T} X_\alpha$$

is where we are taking the union of every single set in the given collection. Then, by definition, if we consider an $x \in \cup_{\alpha \in T} U_\alpha$, then x needs to be contained in at least set in the given collection, and since each set corresponds to an element of the set T, this means that there exists an element $\alpha_0 \in T$ such that $x \in X_{\alpha_0}$, the corresponding test.

For example, if T is finite, then one can write $T = \{1, 2, \ldots, n\}$, and for each $i \in T$, we have a corresponding set X_i, then this collection may be written as $\{X_1, X_2, \ldots, X_n\}$. Their union is

$$\bigcup_{i \in T} X_\alpha = X_1 \cup X_2 \cup \cdots \cup X_n = \bigcup_{i=1}^{n} X_\alpha$$

If T is countably finite, then one may assume $T = \mathbb{N}$, then for each natural number k, we have a corresponding set X_k, then this collection may be written as $\{X_1, X_2, \ldots\} = \{X_k\}_{k=1}^{\infty}$. Their union is

$$\bigcup_{k \in T} X_\alpha = X_1 \cup X_2 \cup \cdots = \bigcup_{k=1}^{\infty} X_\alpha$$

If T is uncountably finite, assume $T = \mathbb{R}$, then for each natural number y, we have a corresponding set X_y, then this collection may be written as $\{X_y\}_{y \in \mathbb{R}}$. Their union is

$$\bigcup_{y \in T} X_y = \bigcup_{y \in \mathbb{R}} X_y$$

So, if $x \in \cup_{y \in \mathbb{R}} X_y$, then there exists a real number y_0 such that $x \in X_{y_0}$.

A similar discussion holds for intersections.

Example 5.1.6. Let us look at some more examples of open sets using Proposition 5.1.4 and Example 5.1.3.

(a) Since we have seen that open intervals (a, b) are open sets, therefore so are their unions. For example, $(a, b) \cup (c, d)$ is open; and so is, the following union of an infinite collection of sets

$$\bigcup_{n \in \mathbb{N}} (-n, n),$$

using the Archimedean Principle, once can show that the union above equals $\mathbb{R}$, which we know to be an open set.

(b) Likewise, $(a, b) \cap (c, d)$ is open. For instance,

$$(0, 3) \cap (1, 4) = (1, 3)$$
$$(-2, -1) \cap (-1, 3) = \varnothing$$

are open sets. One can show that the intersection of open intervals is another open interval or the empty set.

(c) Similarly, unions and intersections of open rays are open. That is,

$$(-\infty, a) \cup (b, \infty)$$

is an open set. So is,

$$((-\infty, 2) \cup (3, \infty)) \cap ((-\infty, -1) \cap (2, \infty))$$

which may show is $(-\infty, -1) \cup (3, \infty)$, which we have observed is open.

(d) Note that arbitrary intersections of open sets may not be open. TO see this, consider for each $n \in \mathbb{N}$, then open interval $I_n = (-1/n, 1/n)$; we claim that

$$\bigcap_{n \in \mathbb{N}} I_n = \{0\}$$

This can be proven by using the fact that $0 = \inf\{1/n \ : \ n \in \mathbb{N}\}$.

But observe that $\{0\}$ is *not open*, since for any $\delta > 0$, we note that

$$(0 - \delta, 0 + \delta) = (-\delta, \delta) \not\subseteq \{0\},$$

as $\delta/2 \notin \{0\}$, for example.

Proposition 5.1.7 (Representation Theorem for Open Sets)**.** *Every proper open subset of $\mathbb{R}$ is a countable union of disjoint open intervals or rays.*

Proof. Let U be an open set, then for each $x \in U$ we may find a $\delta > 0$ such that x is contained in the open interval $(x - \delta, x + \delta) \subseteq U$. This may not be largest open interval in U containing x, we may find a larger open interval. Let $I_x \subseteq U$ be the largest open interval (or ray) in U containing x. Formally, $I_x = (\alpha_x, \beta_x)$ where

$$\alpha_x = \inf\{a \ : \ (a, x) \subseteq U\} \quad \text{and} \quad \beta_x = \sup\{b \ : \ (x, b) \subseteq U\}$$

We let $\alpha_x = -\infty$ if the relevant set is not bounded below, and $\beta_x = \infty$ if the relevant set is not bounded above.

We claim that the collection $\{I_x\}_{x \in U}$ is a countable collection and its union is U. We first observe that for any pair $x, y \in U$, we either have $I_x \cap I_y = \varnothing$ or $I_x = I_y$. Suppose $I_x \neq I_y$ and $I_x \cap I_y \neq \varnothing$, then $I_x \cup I_y$ is an open interval containing both x and y, contradicting the fact that I_x and I_y were the largest such intervals. Thus, $I_x \cap I_y = \varnothing$. And since $I_x \subseteq U$ for each $x \in U$, and moreover since each $x \in U$ is contained in its corresponding interval I_x, we have shown that $\{I_x\}_{x \in U}$ is a collection of disjoint open intervals whose union is U.

The collection $\{I_x\}$ $x \in U$ has repetitions as a pair of these intervals I_x and I_y are either equal or disjoint. Let us collect the distinct disjoint intervals together into a set $\mathcal{C}$. That is, $I_x, I_y \in \mathcal{C}$ if and only if $I_x \cap I_y = \varnothing$. We continue to have that

$$U = \bigcup_{I_x \in \mathcal{C}} I_x$$

For each $I_x \in \mathcal{C}$, by density of rationals, there exists a $r_x \in \mathbb{Q}$ such that $r_x \in I_x$. Then, we claim that

$$\Phi : \mathcal{C} \to \mathbb{Q}, \ I_x \mapsto r_x$$

is an injective function. Suppose $r_x = r_y$ for some pair $x, y \in U$, then for the corresponding open intervals I_x and I_y, we have $r_x = r_y \in I_x \cap I_y$. That is, $I_x \cap I_y \neq \varnothing$, and necessarily $I_x = I_y$. Thus, Φ is injective. Therefore, $\mathcal{C}$ injects into $\mathbb{Q}$, a countable set, and hence $\mathcal{C}$ is countable. □

5.2. Closed Sets

Definition 5.2.1 (Closed Sets)**.** A set $A \subseteq \mathbb{R}$ is called closed if its complement A^c is open.

Recall that $A^c = \mathbb{R} \setminus A = \{x \in \mathbb{R} \ : \ x \notin A\}$.

Remark 5.2.2. Consider any open set U, and then consider its complement U^c. Since $(U^c)^c = U$, an open set, this tells us that U^c is closed. Therefore, not only is the complement of a closed set an open set (by definition), the complement of an open set is closed.

Example 5.2.3. Let us look at some example of closed sets.

(a) Note that $\mathbb{R}$ and $\varnothing$ are closed sets, since their complements $\varnothing$ and $\mathbb{R}$ respectively are open (Example 5.1.3).

(b) The *closed interval* $[a, b] = \{x \in \mathbb{R} \ : \ a \leqslant x \leqslant b\}$ is a closed set. Because it complement is the set $(-\infty, a) \cup (b, \infty)$, which is open (Example 5.1.6).

(c) The *closed rays* are closed sets,

$$(-\infty, b] = \{x \in \mathbb{R} \ : \ x \leqslant b\} \quad \text{and} \quad [b, \infty) = \{x \in \mathbb{R} \ : \ x \geqslant b\},$$

since their complements

$$(-\infty, b]^c = (b, \infty) \quad \text{and} \quad [b, \infty)^c = (-\infty, b)$$

are open sets (Example 5.1.6).

(c) The *singleton* $\{x\}$ for any $x \in \mathbb{R}$ is closed, since its complement $(-\infty, x) \cup (x, \infty)$ is open (Example 5.1.6).

Remark 5.2.4. Given a set $X \subseteq \mathbb{R}$, there are four possibilities for this set.

- X may be open, e.g., if $X = (1, 2)$.
- X may be closed, e.g., if $X = [4, 5]$.
- X may be open and closed, e.g., if $X = \mathbb{R}$ or $X = \varnothing$.
  In fact, these are the only two subsets of $\mathbb{R}$ that are open and closed.
- X may be neither open nor closed, e.g., if $X = [3, 7)$.
  We saw in Example 5.1.3 that $[3, 7)$ is not open. To show it is not closed, we wish to show that its complement $(-\infty, 3) \cup [7, \infty)$ is not open. This can be shown using a similar argument using $7 \in (-\infty, 3) \cup [7, \infty)$, as the one for showing $[3, 7)$ is not open using $3 \in [3, 7)$.

In particular, the negation of being an open set is not being a closed set.

Lecture 16

Definition 5.2.5. For a set A, any $x \in \mathbb{R}$ is a limit point if there exists a sequence (a_n) in A, i.e. $a_n \in A$ for every n, such that $a_n \to x$.

Example 5.2.6. For the interval $(0,1]$, both 0 and 1 are limit points.

0 is a limit point since $\frac{1}{n} \in (0,1]$ for every $n \in \mathbb{N}$, and $\frac{1}{n} \to 0$.

1 is a limit point since $\frac{n}{n+1} \in (0,1]$ for every $n \in \mathbb{N}$, and $\frac{n}{n+1} \to 1$.

Theorem 5.2.7 (Characterisation of Closed Sets)**.** *A set A is closed if and only if A contains all its limit points.*

Proof. ($\Rightarrow$) We prove this by contradiction. Suppose A is closed and there exists a limit point $x \in \mathbb{R}$ of A such that $x \notin A$. Since x is a limit point of A, there exists a sequence (a_n) in A such that $a_n \to x$.

By assumption $x \notin A$, therefore $x \in A^c$. Since A is closed, A^c is open. Hence, there exists a $\delta > 0$ such that $(x-\delta, x+\delta) \subseteq A^c$.

Since $a_n \to x$, for $\varepsilon = \delta > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - x| < \delta, \quad \text{for every } n \geqslant N.$$

In particular, we have

$$|a_N - x| < \delta$$

Therefore, $a_N \in (x-\delta, x+\delta) \subseteq A^c$, i.e. $a_N \in A^c$ or equivalently $a_N \notin A$. This is a contradiction, as we know that $a_N \in A$. Thus, A contains all its limit points.

($\Leftarrow$) We prove the contrapositive, that is, we will prove that if A is not closed, then A does not contain all its limit points. Since A is not closed, we have that A^c is not open. Therefore there exists an $x \in A^c$ such that for every $\delta > 0$ we have $(x-\delta, x+\delta) \not\subseteq A^c$.

For this $x \in A^c$, equivalently $x \notin A$, and $1/n > 0$, we have

$$\left(x - \frac{1}{n}, x + \frac{1}{n}\right) \not\subseteq A^c$$

That is, there exists an $a_n \in (x - 1/n, x + 1/n)$ such that $a_n \notin A^c$. Since $a_n \notin A^c$, we have that $a_n \in A$ for every n. Moreover, since

$$x - \frac{1}{n} < a_n < x + \frac{1}{n}$$

for every n, and hence $a_n \to x$ by the Squeeze Theorem (Theorem 3.2.3). Hence, $x \notin A$ is a limit point of A and thus A does not contain all its limit points. □

Proposition 5.2.8.

(i) If $\{A_\alpha\}_{\alpha \in T}$ is an arbitrary collection of closed sets, then $\bigcap_{\alpha \in T} A_\alpha$ is a closed set.

(i) If $\{A_1, A_2, \ldots, A_n\}$ is a finite collection of closed sets, then $\bigcup_{i=1}^{n} A_i$ is a closed set.

Proof. This follows from Proposition 5.1.4 and De Morgan's Laws of how complements interacts with unions and intersections. □

5.3. Compact Sets

Definition 5.3.1 (Open Covers)**.** Let $X \subseteq \mathbb{R}$ be a set.

- A set $\{U_\alpha\}_{\alpha \in T}$ of subsets of $\mathbb{R}$ is a cover of X if
$$X \subseteq \bigcup_{\alpha \in T} U_\alpha$$
- If each U_α is open, then $\{U_\alpha\}_{\alpha \in T}$ is an open cover of X.
- Suppose there exists a finite subset of $\{U_\alpha\}_{\alpha \in T}$, written $\{U_\alpha\}_{\alpha \in F}$, where $F \subseteq T$ is finite, which is still a cover of X. Then $\{U_\alpha\}_{\alpha \in F}$ is called a finite subcover of X.

Example 5.3.2.

(a) An open cover of $(0,5)$ is
$$\left\{\left(\frac{1}{k}, 5-\frac{1}{k}\right)\right\}_{k=1}^{\infty} = \left\{(1,3), \left(\frac{1}{2}, \frac{9}{2}\right), \left(\frac{1}{3}, \frac{14}{3}\right), \dots\right\}$$
In fact,
$$(0,5) = \bigcup_{k=1}^{\infty} \left(\frac{1}{k}, 5-\frac{1}{k}\right)$$
We see the latter is a subset of the former, now consider any $x \in (0,5)$, that is, $0 < x < 5$, which also gives us $5 - x > 0$. By the Archimedean Principle, there exist $k_1, k_2 \in \mathbb{N}$ such that
$$\frac{1}{k_1} < x \quad \text{and} \quad \frac{1}{k_2} < 5 - x.$$
This gives us, $\frac{1}{k_1} < x < 5 - \frac{1}{k_2}$. Define $k = \max(k_1, k_2)$, then
$$\frac{1}{k} < \frac{1}{k_2} < x < 5 - \frac{1}{k_2} < 5 - \frac{1}{k}$$
and so $x \in \left(\frac{1}{k}, 5-\frac{1}{k}\right)$, giving us the equality of sets we needed.

(b) Another open cover of $(0,5)$ is
$$\{(-\infty, 2), (1,3), (1,2), (0.5, 3.5), (3,7)\}$$
Note that not all sets in this cover are subsets of $(0,5)$, this is fine, as the definition of covering doesn't require them to be.

Furthermore, note that not all sets in this cover are necessary, removing $(1,2)$ still yields a cover of $(0,5)$. That is,
$$\{(-\infty, 2), (1,3), (0.5, 3.5), (3,7)\}$$
is a finite subcover.

(c) An open cover of $[0,5]$ is

$$\left\{\left(\frac{1}{k}, 5-\frac{1}{k}\right)\right\}_{k=1}^{\infty} \cup \{(-0.2, 0.2), (4.9, 5.1)\}$$

Definition 5.3.3 (Compact Sets)**.** A set $K \subseteq \mathbb{R}$ is compact if *every open cover of K contains a finite subcover of K.*

Example 5.3.4. Let us revisit Example 5.3.2

(a) The open cover

$$\left\{\left(\frac{1}{k}, 5-\frac{1}{k}\right)\right\}_{k=1}^{\infty}$$

of $(0,5)$ has *no finite subcover*. To see this, assume there was one, then we may write it as

$$\left\{\left(\frac{1}{k_1}, 5-\frac{1}{k_1}\right), \left(\frac{1}{k_2}, 5-\frac{1}{k_2}\right), \ldots, \left(\frac{1}{k_n}, 5-\frac{1}{k_n}\right)\right\}$$

Then, we may consider $N = \max(k_1, k_2, \ldots, k_n)$, then we note that

$$\left(\frac{1}{k_i}, 5-\frac{1}{k_i}\right) \subseteq \left(\frac{1}{N}, 5-\frac{1}{N}\right)$$

for every $1 \leqslant i \leqslant n$, which tells us that

$$\left(\frac{1}{k_1}, 5-\frac{1}{k_1}\right) \cup \left(\frac{1}{k_2}, 5-\frac{1}{k_2}\right) \cup \cdots \cup \left(\frac{1}{k_n}, 5-\frac{1}{k_n}\right) = \left(\frac{1}{N}, 5-\frac{1}{N}\right)$$

Now, we note that

$$(0,5) \neq \left(\frac{1}{N}, 5-\frac{1}{N}\right) = \left(\frac{1}{k_1}, 5-\frac{1}{k_1}\right) \cup \left(\frac{1}{k_2}, 5-\frac{1}{k_2}\right) \cup \cdots \cup \left(\frac{1}{k_n}, 5-\frac{1}{k_n}\right)$$

since the former set contains, for example $1/2N$, while the latter set does not.

Therefore, $(0,5)$ is *not compact*. A similar argument shows that open intervals are not compact.

(b) But note that the following open cover

$$\left\{\left(\frac{1}{k}, 5-\frac{1}{k}\right)\right\}_{k=1}^{\infty} \cup \{(-0.2, 0.2), (4.9, 5.1)\}$$

of $[0,5]$ does have a finite subcover. You may take

$$\left\{\left(\frac{1}{12}, 5-\frac{1}{12}\right), (-0.2, 0.2), (4.9, 5.1)\right\}$$

Hence, $[0,5]$ may just be compact.

Lecture 17

Theorem 5.3.5 (Heine-Borel Theorem). *A set $K \subseteq \mathbb{R}$ is compact if and only if K is closed and bounded.*

Proof. ($\Rightarrow$) Assume that K is a compact set, we wish to show K is closed and bounded.

I. We will first show K is bounded. Define $I_n = (-n, n)$, and we observe that

$$K \subseteq \mathbb{R} = \bigcup_{n \in \mathbb{N}} (-n, n)$$

Therefore $\{I_n\}_{n=1}^{\infty}$ is an open cover of K, and hence it contains a finite subcover of K as K is compact. We may write this finite subcover as

$$\{I_{n_1}, I_{n_2}, \ldots, I_{n_m}\}$$

We may assume that $n_1 < n_2 < \cdots < n_m$, which tells us that

$$I_{n_1} \subseteq I_{n_2} \subseteq \cdots \subseteq I_{n_m},$$

giving us

$$K \subseteq \bigcup_{i=1}^{m} I_{n_i} = I_{n_m} = (-n_m, n_m)$$

Hence, K is bounded, as $K \subseteq I_{n_m}$ tells us that $|x| < n_m$ for every $x \in K$..

II. We will now show K is closed. Suppose, for the sake of contradiction, that K is not closed. Therefore, by Theorem 5.2.7, there exists some $x \notin K$ and a sequence (a_n) in K such that $a_n \to x$. We will use this to construct an open cover of K and the existence of a finite subcover will provide us with our contradiction.

Consider, for every $n \in \mathbb{N}$, the open sets

$$U_n = \left(-\infty, x - \frac{1}{n}\right) \cup \left(x + \frac{1}{n}, \infty\right),$$

and using the Archimedean Principle one may see that

$$\bigcup_{n=1}^{\infty} U_n = \mathbb{R} \setminus \{x\}$$

Since $x \notin K$, we observe that $K \subseteq \mathbb{R} \setminus \{x\}$, and thus $\{U_n\}_{n \geqslant 1}$ is an open cover of K. Since K is compact, this open cover contains a finite subcover of K, which me write as

$$\{U_{n_1}, U_{n_2}, \ldots, U_{n_m}\}$$

We may assume that $n_1 < n_2 < \cdots < n_m$, which tells us that

$$U_{n_1} \subseteq U_{n_2} \subseteq \cdots \subseteq U_{n_m},$$

giving us

$$K \subseteq \bigcup_{i=1}^{m} U_{n_i} = U_{n_m}$$

That is,

$$K \subseteq \left(-\infty, x - \frac{1}{n_m}\right) \cup \left(x + \frac{1}{n_m}, \infty\right)$$

This, in particular, tells us that

$$K \cap \left(x - \frac{1}{n_m}, x + \frac{1}{n_m} \right) = \varnothing \tag{$*$}$$

This provides us with a contradiction as follows: since $a_n \to x$, for $\varepsilon = 1/n_m > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - x| < \frac{1}{n_m} \quad \text{for every } n \geqslant N$$

Expanding this given inequality tells us

$$x - \frac{1}{n_m} < a_n < x + \frac{1}{n_m}$$

for every $n \geqslant N$. Since $a_n \in K$ for every n, this is telling us that

$$K \cap \left(x - \frac{1}{n_m}, x + \frac{1}{n_m} \right) \neq \varnothing$$

contradicting $(*)$. Thus, our assumption was false, and therefore K is closed.

($\Leftarrow$) skipped. □

Theorem 5.3.6 (Characterisation of Compact Sets)**.** *For a set $K \subseteq \mathbb{R}$, the following are equivalent.*

(i) K is compact.

(ii) K is closed and bounded.

(iii) Every sequence in K has a convergent subsequence whose limit lies in K.

Proof. The equivalence of (i) and (ii) is precisely the Heine-Borel theorem (Theorem 5.3.5). We prove the equivalence of (ii) and (iii).

(ii) $\Rightarrow$ (iii) Assume K is closed and bounded, and consider any sequence (a_n) is a sequence, with $a_n \in K$ for every n. Since K is bounded, we conclude that so is (a_n). Thus, by the Bolzano-Weierstrass Theorem (Theorem 3.3.7), there exists a convergent subsequence of (a_n), say (a_{n_k}). Let $a_{n_k} \to a$, we wish to show that $a \in K$. Since $a_{n_k} \in K$ for every k, this tells us that a is a limit point of K, and since K is closed, we conclude $a \in K$ by Theorem 5.2.7.

(iii) $\Rightarrow$ (ii) Assume that every sequence in K has a convergent subsequence whose limit lies in K.

We will first prove that K is closed; we will show that it contains all its limit points (Theorem 5.2.7). Let $x \in \mathbb{R}$ be a limit point of K, therefore there exists a sequence (a_n) in K such that $a_n \to x$. By assumption, (a_n) has a subsequence (a_{n_k}) whose limit is in K. Necessarily, $a_{n_k} \to x$ and we then conclude that $x \in K$. Hence, K contains all its limit points and therefore K is closed.

Let us now show that K is bounded; assume for the sake of contradiction that K is not bounded. We have two cases, either K is no bounded below or K is not bounded above. We deal with the case K is not bounded above, the case of K not being bounded below can be dealt with similarly. Since K is not bounded above, for every $n \in \mathbb{N}$, there exists an $a_n \in K$ such that $a_n > n$. Since the sequence (n), and any of its subsequences, are such that they diverge to ∞, we conclude that $a_n \to \infty$ and so do any of its subsequences. This contradicts the assumption that every sequence in K has a convergent subsequence whose limit lies in K. Hence K is bounded. □

6. Continuity

6.1. Functional Limits and Properties

Definition 6.1.1 (Functional Limit)**.** Let $f : X \to \mathbb{R}$ be a function, and let a be a limit point of X. Then we say *as x approaches a, the limit of f is L*, denoted

$$\lim_{x \to a} f(x) = L,$$

if *for all $\varepsilon > 0$, there exists a $\delta > 0$ such that, for every $x \in X$, if $0 < |x - a| < \delta$, then $|f(x) - L| < \varepsilon$.*

Remark 6.1.2. We make some pertinent remarks about this definition.

1. We ask for $x \in X$ in the definition just so that $f(x)$ makes sense. Note that if $x \notin X$, the domain of f, it does not make sense to write $f(x)$. Since, we do not require to write $f(a)$, it is okay if $a \notin X$; in fact, this is the case of interest.
2. We require a to be a limit point of X because we wish to make sure that elements in X actually do approach a.
3. The reason we ask for $0 < |x - a|$, equivalently $x \neq a$, is because limits are of interest when we approach a, not what is occurring at a. We may define $f(a)$ to be anything we like, but it does not dictate the limiting behaviour of the function while approaching a.
4. When the context is clear, we do not usually write $x \in X$ and $0 < |x - a|$, but they remain part of the definition.
5. To show that limit of $f(x)$, as x approaches a, is not L, we must prove the negation of the definition of the functional limit: *there exists $\varepsilon > 0$ such that for every $\delta > 0$ there exists an $x \in X$ for which $0 < |x - a| < \delta$ but $|f(x) - L| \geqslant \varepsilon$.*

Lecture 18

Example 6.1.3. Consider $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = 7x + 3$. Let us prove $\lim_{x \to 2} f(x) = 17$.

Solution. We begin with some scratch work.

> *Scratch Work.* Given an arbitrary $\varepsilon > 0$, we need to find a specific $\delta > 0$ such that whenever $0 < |x - 2| < \delta$ we have $|f(x) - 17| < \varepsilon$. We first observe that
>
> $$|f(x) - 17| = |7x + 3 - 17| = |7x - 14| = |7(x - 2)|$$
>
> Note that if we had our conclusion $|f(x) - 17| < \varepsilon$, then
>
> $$|7(x - 2)| < \varepsilon, \quad \text{hence} \quad |x - 2| < \frac{\varepsilon}{2}$$
>
> So, this tells us what we can define our δ to be; it should be $\delta = \varepsilon/7$.

Fix any $\varepsilon > 0$, then for $\delta = \varepsilon/7 > 0$, whenever $0 < |x - 2| < \delta = \varepsilon/7$, we have

$$|f(x) - 17| = |7(x - 2)| = 7 \cdot |x - 2| < 7\delta = 7 \cdot \frac{\varepsilon}{7} = \varepsilon$$

Therefore $\lim_{x \to 2} f(x) = 17$. □

Example 6.1.4. Consider $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$. Let us prove $\lim_{x\to 5} f(x) = 25$.

Solution. We begin with some scratch work.

> *Scratch Work.* Given an arbitrary $\varepsilon > 0$, we need to find a specific $\delta > 0$ such that whenever $0 < |x-5| < \delta$ we have $|f(x) - 25| < \varepsilon$. We first observe that
>
> $$|f(x) - 25| = \left|x^2 - 25\right| = |(x-5)(x+5)| = |x-5| \cdot |x+5|$$
>
> Like the previous example, if we had to start with our conclusion $|f(x) - 25| < \varepsilon$, that is,
>
> $$|x-5| \cdot |x+5| < \varepsilon$$
>
> It doesn't seem like we can isolate the $|x-5|$ in a way such that the right hand side is independent of x, and engage in our "reverse engineering" strategy. Therefore, we need a new strategy.
>
> Let us forget about ε for a bit, and consider those $x \in \mathbb{R}$ such that $0 < |x-5| < 1$. The choice of 1 is arbitrary, you may choose any other number; the aim is to start our approach to 5 after a concrete point. If $|x-5| < 1$, then to say something about $\left|x^2 - 25\right|$ we wish to understand what we can do with $|x+5|$. We have,
>
> $$|x+5| = |x-5+10| \leqslant |x-5| + 10 < 11$$
>
> Let use this inequality as our starting point, if $|x+5| < 11$, then
>
> $$|f(x) - 25| = |x-5| \cdot |x+5| < 11 \cdot |x-5|$$
>
> Therefore, if we had $|x-5| < \varepsilon/11$, this would give us
>
> $$|f(x) - 25| = |x-5| \cdot |x+5| < 11 \cdot |x-5| < 11 \cdot \frac{\varepsilon}{11} = \varepsilon$$
>
> which is what we wished to show. Therefore, we wish to find a δ such that whenever $|x-5| < \delta$, we have $|x-5| < \varepsilon/11$ and $|x+5| < 11$. Recall that the latter condition is a consequence of the case when $|x-5| < 1$. Hence, we wish to find a δ such that whenever $|x-5| < \delta$, we have $|x-5| < \varepsilon/11$ and $|x-5| < 1$.
>
> So, this tells us what we can define our δ to be. It should be $\delta = \min(\varepsilon/11, 1)$.

Fix any $\varepsilon > 0$, and consider $\delta = \min(\varepsilon/11, 1)$. Whenever $0 < |x-5| < \delta$, we have

$$\begin{aligned}
|f(x) - 25| = \left|x^2 - 25\right| &= |x-5| \cdot |x+5| \\
&< 11 \cdot |x-5| && \text{since } |x-5| < \delta \leqslant 1 \\
&< 11 \cdot \frac{\varepsilon}{11} && \text{since } |x-5| < \delta \leqslant \frac{\varepsilon}{11} \\
&= \varepsilon
\end{aligned}$$

Therefore $\lim_{x\to 5} f(x) = 25$. □

Proposition 6.1.5. *The limit of a function, if it exists, is unique.*

Proof. Consider a function $f : X \to \mathbb{R}$, and let $a \in \mathbb{R}$ be a limit point of X. Suppose

$$\lim_{x\to a} f(x) = \alpha \quad \text{and} \quad \lim_{x\to a} f(x) = \beta;$$

we wish to show $\alpha = \beta$. Given an $\varepsilon > 0$, we also have $\varepsilon/2 > 0$.

Since $\lim_{x\to a} f(x) = \alpha$, there exists a $\delta_1 > 0$ such that if $0 < |x - a| < \delta_1$ then $|f(x) - \alpha| < \varepsilon/2$.

Since $\lim_{x\to a} f(x) = \beta$, there exists a $\delta_2 > 0$ such that if $0 < |x - a| < \delta_2$ then $|f(x) - \beta| < \varepsilon/2$.

Define $\delta = \min(\delta_1, \delta_2)$, then if $0 < |x - a| < \delta$, we have

$$\begin{aligned}
|\alpha - \beta| &= |\alpha - \beta + (f(x) - f(x))| \\
&= |(f(x) - \beta) + (\alpha - f(x))| \\
&\leqslant |f(x) - \beta| + |\alpha - f(x)| && \text{triangle inequality} \\
&= |f(x) - \beta| + |f(x) - \alpha| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} && \text{since } 0 < |x - a| < \delta \leqslant \delta_1,\ \delta_2 \\
&= \varepsilon
\end{aligned}$$

That is, $|\alpha = \beta| < \varepsilon$, for every $\varepsilon > 0$. By Problem 17, we conclude that $\alpha = \beta$. □

Theorem 6.1.6 (Sequential Characterisation of Functional Limits)**.** *Let $X \subseteq \mathbb{R}$, and consider a function $f : X \to \mathbb{R}$ and a limit point a of X. Then*

$$\lim_{x\to a} f(x) = L \quad \textit{if and only if}$$

for every sequence (x_n) in X, for which $x_n \neq a$ for every n, such that $x_n \to a$, we have $f(x_n) \to L$.

Proof. Throughout this proof assume that $x \neq a$, and $x_n \neq a$ for every n.

($\Rightarrow$) Assume that $\lim_{x\to a} f(x) = L$, and consider a sequence (x_n) in X such that $x_n \to a$, we wish to show that $f(x_n) \to L$.

Consider an $\varepsilon > 0$, since $\lim_{x\to a} f(x) = L$, there exists a $\delta > 0$ such that

$$\text{if } 0 < |x - a| < \delta, \text{ then } |f(x) - L| \tag{$*$}$$

Since $x_n \to a$ and $\delta > 0$, there exists an $N \in \mathbb{N}$ such that

$$|x_n - a| < \delta, \quad \text{for every } n \geqslant N$$

For every $n \geqslant N$, and since $x_n \neq a$ for all $n \in \mathbb{N}$, we have

$$0 < |x_n - a| < \delta$$

Therefore by ($*$) we conclude that

$$|f(x_n) - L| < \varepsilon$$

That is, we have proved that for every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|f(x_n) - L| < \varepsilon$. Thus, $f(x_n) \to L$.

($\Leftarrow$) Suppose every sequence (x_n) in X, for which $x_n \neq a$ for every n, is such that whenever $x_n \to a$, we have $f(x_n) \to L$. Assume for the sake of contradiction that $\lim_{x\to a} f(x) \neq L$. That is, there exists an $\varepsilon > 0$ such that for every $\delta > 0$ there exists an $x \in X$ such that $0 < |x - a| < \delta$ and $|f(x) - L| \geqslant \varepsilon$.

In particular, for this $\varepsilon > 0$, take $\delta = 1/n$, then there exists an $y_n \in X$ such that

$$0 < |y_n - a| < \frac{1}{n} \quad \text{and} \quad |f(y_n) - L| \geqslant \varepsilon$$

Thus, we have found a sequence (y_n) in X, for which $y_n \neq a$ for all n, such that $y_n \to a$ (using the Squeeze Theorem), for which $|f(y_n) - L| \geqslant \varepsilon$ for every n; therefore $f(y_n) \not\to L$. Hence, we have arrived at a contradiction, and thus $\lim_{x\to a} f(x) = L$. □

Remark 6.1.7. Theorem 6.1.6 gives a criterion to determine that a functional limit does not exist: if you can find two sequences (x_n) and (y_n) in X that both converge to a but for which

$$\lim_{n\to\infty} f(x_n) \neq \lim_{n\to\infty} f(y_n)$$

then we can conclude that $\lim_{x\to a} f(x)$ does not exist.

Example 6.1.8. Consider the characteristic function of the rationals $\chi_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ where

$$\chi_{\mathbb{Q}}(t) = \begin{cases} 1 & \text{if } t \in \mathbb{Q} \\ 0 & \text{if } t \notin \mathbb{Q} \end{cases}$$

We will use Remark 6.1.7 to show that $\lim_{t\to a} \chi_{\mathbb{Q}}(t)$ does not exist for any $a \in \mathbb{R}$. That is, we produce two sequences (r_n) and (x_n) such that $r_n \to a$ and $x_n \to a$ but

$$\lim_{n\to\infty} \chi_{\mathbb{Q}}(r_n) \neq \lim_{n\to\infty} \chi_{\mathbb{Q}}(x_n).$$

Using the density of $\mathbb{Q}$, for $a \in \mathbb{R}$, we may find a sequence (r_n), where $r_n \in \mathbb{Q}$ for ever n and $r_n \to a$. Applying this to $a/\sqrt{2} \in \mathbb{R}$, we may again find a sequence (q_n), where $q_n \in \mathbb{Q}$ for ever n and $q_n \to a/\sqrt{2}$. Using limit laws of sequences, we deduce $\sqrt{2} \cdot q_n \to a$; define $x_n = \sqrt{2} \cdot q_n$, and since $q_n \in \mathbb{Q}$ and $\sqrt{2} \notin \mathbb{Q}$, we conclude that $x_n \notin \mathbb{Q}$ for every n.

Therefore, for any $a \in \mathbb{R}$, we have produced two sequences (r_n) and (x_n) such that $r_n \to a$ and $x_n \to a$, where $r_n \in \mathbb{Q}$ and $x_n \notin \mathbb{Q}$ for every n.

Now note that $\chi_{\mathbb{Q}}(r_n) = 1$ for every n, and $\chi_{\mathbb{Q}}(x_n) = 0$ for every n, and thus

$$\lim_{n\to\infty} \chi_{\mathbb{Q}}(r_n) = 1 \neq 0 = \lim_{n\to\infty} \chi_{\mathbb{Q}}(x_n).$$

Therefore $\lim_{t\to a} \chi_{\mathbb{Q}}(t)$ does not exist for any $a \in \mathbb{R}$.

Lecture 19

Corollary 6.1.9 (Functional Limit Laws)**.** *Let f and g be function with domain X, and let a be a limit point of X. Assume that*

$$\lim_{x\to a} f(x) = \alpha \quad \textit{and} \quad \lim_{x\to a} g(x) = \beta$$

Then,

(i) $\lim_{x\to a}(k\cdot f(x)) = k\cdot\alpha$, *for any* $k\in\mathbb{R}$.

(ii) $\lim_{x\to a}(f(x)+g(x)) = \alpha+\beta$.

(iii) $\lim_{x\to a}(f(x)\cdot g(x)) = \alpha\cdot\beta$.

(iv) $\lim_{x\to a}\dfrac{f(x)}{g(x)} = \dfrac{\alpha}{\beta}$, *if* $\beta\neq 0$ *and* $g(x)\neq 0$ *for any* $x\in X$.

Proof. The result follows from the sequential limit laws (Theorem 3.2.1 and Corollary 3.2.2), and the sequential characterisation of limits (Theorem 6.1.6). □

Theorem 6.1.10 (Functional Squeeze Theorem)**.** *Let* f, g *and* h *be function with domain* X, *and let* a *be a limit point of* X. *Assume that*

$$f(x)\leqslant g(x)\leqslant h(x),\quad \text{for every } x\in X,$$

and that $\lim_{x\to a} f(x) = L$ *and* $\lim_{x\to a} h(x) = L$. *Then,*

$$\lim_{x\to a} g(x) = L$$

Proof. The result follows from the sequential squeeze theorem (Theorem 3.2.3), and the sequential characterisation of limits (Theorem 6.1.6). □

6.2. Continuity

Definition 6.2.1 (Continuous Functions)**.** Let $f: X\to\mathbb{R}$ be a function. f is continuous at $a\in X$ if *for all* $\varepsilon>0$, *there exists a* $\delta>0$ *such that, for every* $x\in X$, *if* $|x-a|<\delta$, *then* $|f(x)-f(a)|<\varepsilon$.

If f is *continuous at every point in its domain*, then the f is said to be continuous.

Theorem 6.2.2 (Characterisation of Continuity)**.** *Let* $f: X\to\mathbb{R}$ *and* $a\in X$. *Then the following are equivalent.*

(i) f *is continuous at* a.

(ii) *for all* $\varepsilon>0$, *there exists a* $\delta>0$ *such that, for every* $x\in X$, *if* $|x-a|<\delta$, *then* $|f(x)-f(a)|<\varepsilon$.

(iii) *for any* ε*-neighbourhood* $B_\varepsilon(f(a))$ *of* $f(a)$, *there exists a* δ*-neighbourhood* $B_\delta(a)$ *of* a *such that, for every* $x\in X$, *if* $x\in B_\delta(a)$, *then* $f(x)\in B_\varepsilon(f(a))$.

(iv) $\lim_{x\to a} f(x) = f(a)$.

(v) *for all sequences* (x_n) *in* X *such that* $x_n\to a$, *we have* $f(x_n)\to f(a)$.

Proof. (i) ⇔ (ii) is a consequence of the definition of being continuous at a point.

(ii) ⇔ (iii) is also a consequence of the definition of neighbourhoods.

(iv) ⇔ (v) is a consequence of Theorem 6.1.6.

(ii) ⇔ (iv) also follows from unpacking the definition. □

Remark 6.2.3. Theorem 6.2.2 gives a criterion to determine that a function is continuous at a point: if you can find a sequence (x_n) in X such that $x_n \to a$ but $f(x_n) \not\to f(a)$.

Example 6.2.4. The function $\chi_{\mathbb{Q}}$ from Example 6.1.8 is not continuous at any real number.

Corollary 6.2.5 (Continuity Limit Laws)**.** *Let f and g be function with domain X, and let $a \in X$. Then,*

(i) $k \cdot f(x)$ *is continuous at a, for any $k \in \mathbb{R}$.*

(ii) $f(x) + g(x)$ *is continuous at a.*

(iii) $f(x) \cdot g(x)$ *is continuous at a.*

(iv) $\dfrac{f(x)}{g(x)}$ *is continuous at a, if $g(x) \neq 0$ for any $x \in X$.*

Proof. Using Theorem 6.2.2 (iv), we can translate our statements into functional limits and then the result follows from Functional Limit Laws. □

Example 6.2.6. For any $c \in \mathbb{R}$, consider the constant function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto c$; this function is continuous. To show this, we will use the sequential characterisation of continuity: consider any $a \in \mathbb{R}$ and a sequence (x_n) where $x_n \to a$. Then, $f(x_n) = c$ for every n, and so $(f(x_n))$ is the constant sequence c. Thus, $f(x_n) \to c = f(a)$ and therefore we conclude that f is continuous. That is,

$$\lim_{x \to a} c = c.$$

The identity function $\mathrm{id} : \mathbb{R} \to \mathbb{R}$, $x \mapsto x$ is continuous. To show this, we will use the sequential characterisation of continuity: consider any $a \in \mathbb{R}$ and a sequence (x_n) where $x_n \to a$. Then, $\mathrm{id}(x_n) = x_n \to a = \mathrm{id}(a)$, and hence we conclude that id is continuous. In particular, we have

$$\lim_{x \to a} x = a$$

Hence, using limit laws for continuity we may then conclude that any polynomial is continuous. Consider any polynomial $p : \mathbb{R} \to \mathbb{R}$ where

$$p(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0$$

here $c_i \in \mathbb{R}$ and $c_n \neq 0$. Thus,

$$\begin{aligned}
\lim_{x \to a} p(x) &= \lim_{x \to a} (c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0) \\
&= \lim_{x \to a} c_n x^n + \lim_{x \to a} c_{n-1} x^{n-1} + \cdots + \lim_{x \to a} c_1 x + \lim_{x \to a} c_0 \\
&= c_n \cdot \lim_{x \to a} x^n + c_{n-1} \cdot \lim_{x \to a} x^{n-1} + \cdots + c_1 \cdot \lim_{x \to a} x + \lim_{x \to a} c_0 \\
&= c_n \left(\lim_{x \to a} x \right)^n + c_{n-1} \left(\lim_{x \to a} x \right)^{n-1} + \cdots + c_1 \cdot \lim_{x \to a} x + \lim_{x \to a} c_0 \\
&= c_n a^n + c_{n-1} a^{n-1} + \cdots + c_1 a + c_0 \\
&= p(a)
\end{aligned}$$

Hence p is a continuous function.

Moreover, this tells us that *rational functions* are also continuous. Consider any rational function $r : X \to \mathbb{R}$ where

$$r(x) = \frac{p(x)}{q(x)}$$

here $p(x)$ and $q(x)$ are polynomials, and $X = \{x \in \mathbb{R} \ : \ q(x) \neq 0\}$. Thus,

$$\lim_{x\to a} r(x) = \lim_{x\to a} \frac{p(x)}{q(x)} = \frac{p(a)}{q(a)} = r(a)$$

Hence r is a continuous function.

Remark 6.2.7 (Fact)**.** You may take for granted that the following standard functions are continuous. You may also take for granted their algebraic rules.

(i) The *exponential function* $\exp : \mathbb{R} \to \mathbb{R},\ x \mapsto \exp(x) = e^x$

(ii) The *logarithmic function* $\ln : (0, \infty) \to \mathbb{R},\ x \mapsto \ln(x)$

(i) The *sine function* $\sin : \mathbb{R} \to \mathbb{R},\ x \mapsto \sin(x)$

(i) The *cosine function* $\cos : \mathbb{R} \to \mathbb{R},\ x \mapsto \cos(x)$

Proposition 6.2.8. *Consider functions $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ such that $f(X) \subseteq Y$. If f is continuous at $a \in X$, and g is continuous at $f(a)$, then $g \circ f : X \to \mathbb{R}$ is continuous at $a \in X$.*

Proof. We shall use the sequential characterisation of continuity. Consider a sequence (x_n) in X such that $x_n \to a$; we wish to show

$$(g \circ f)(x_n) \to (g \circ f)(a)$$

Since f is continuous at a, therefore $f(x_n) \to f(a)$. Now, observe that since $f(X) \subseteq Y$, we have that $(f(x_n))$ is a sequence in Y. Hence, as g is continuous at $f(a)$, we get

$$(g \circ f)(x_n) = g(f(x_n)) \to g(f(a)) = (g \circ f)(a).$$

Thus, $g \circ f$ is continuous at $a \in X$. □

Example 6.2.9. Consider the functions $f, g : \mathbb{R} \to \mathbb{R}$ where $f(x) = \sin(x)$ and $g(x) = x^2$. Then, by Proposition 6.2.8, the functions $g \circ f$ and $f \circ g$, where

$$(f \circ g)(x) = f(g(x)) = f(x^2) = \sin x^2$$

are continuous.

Lecture 20

Remark 6.2.10. Recall that for a function $f : X \to Y$, and a subset $V \subseteq Y$, we can consider the *inverse image of V under f*

$$f^*(V) = \{x \in X \ : \ f(x) \in V\}$$

- Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ and $V = (4, 25)$, then

$$\begin{aligned} f^*((4,25)) &= \{x \in \mathbb{R} \ : \ x^2 \in (4,25)\} \\ &= \{x \in \mathbb{R} \ : \ x \in (2,5) \text{ or } x \in (-5,-2)\} \\ &= (-5,-2) \cup (2,5) \end{aligned}$$

- Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 7x$ and $V = (-1, 14) \cup (17, \infty)$, then

$$\begin{aligned} f^*((-1,35)) &= \{x \in \mathbb{R} \ : \ 7x \in (-1,14) \cup (17,\infty)\} \\ &= \{x \in \mathbb{R} \ : \ x \in (-1/7,2) \cup (17/7,\infty)\} \\ &= (-1/7,2) \cup (17/7,\infty) \end{aligned}$$

Therefore, it seems like the inverse image of an open set is another open set under a continuous function. This is not an accident and holds *almost* generally; there is just one thing preventing it from holding, which we illustrate in the following example.

- Consider the function $f : [0, 10] \to \mathbb{R}$ given by $f(x) = 7x$ and $V = (-1, 14) \cup (17, \infty)$, then

$$\begin{aligned} f^*((-1,35)) &= \{x \in [0,10] \ : \ 7x \in (-1,14) \cup (17,\infty)\} \\ &= \{x \in [0,10] \ : \ x \in (-1/7,2) \cup (17/7,\infty)\} \\ &= [0,10] \cap ((-1/7,2) \cup (17/7,\infty) \\ &= [0,2) \cup (17/7,10] \end{aligned}$$

  which is not an open set.

What it does suggest is that if V is an open set, then $f^*(V) = X \cap U$, where U is an open set.

Theorem 6.2.11 (Topological Characterisation of Continuity)**.** *Consider a function $f : X \to \mathbb{R}$. Then f is continuous if and only if for every open set $V \subseteq \mathbb{R}$, we have $f^*(V) = X \cap U$ for some open set $U \subseteq \mathbb{R}$.*

Proof. In this proof, we will use the formulation of continuity from Theorem 6.2.2 (iii).

($\Rightarrow$) Assume f is continuous and consider any open set $V \subseteq \mathbb{R}$. We wish to show that, for some open set U, we have $f^*(V) = X \cap U$. This is equivalent to showing that for any $a \in f^*(V)$, there exists a δ-neighbourhood $B_\delta(a)$ of a such that $X \cap B_\delta(a) \subseteq f^*(V)$.

Consider any $a \in f^*(V)$, therefore $f(a) \in V$. Since V is open, there exists an ε-neighbourhood $B_\varepsilon(f(a))$ of $f(a)$ such that $B_\varepsilon(f(a)) \subseteq V$. Since f is continuous, for this ε-neighbourhood, there exists a δ-neighbourhood $B_\delta(a)$ of a such that, for every $x \in X$, if $x \in B_\delta(a)$, then $f(x) \in B_\varepsilon(f(a))$. That is, $X \cap B_\delta(a) \subseteq f^*(V)$.

($\Leftarrow$) Assume that for every open set $V \subseteq \mathbb{R}$, we have $f^*(V) = X \cap U$ for some open set $U \subseteq \mathbb{R}$, we wish to show that f is continuous. Consider any $a \in X$ and $\varepsilon > 0$. Since $B_\varepsilon(f(a))$ is an open set in $\mathbb{R}$, we have that $f^*(B_\varepsilon(f(a)) = X \cap U$ for some open set U. Since $f(a) \in B_\varepsilon(f(a))$, therefore $a \in f^*(B_\varepsilon(f(a)) = X \cap U$ and hence it follows from this that $a \in U$. Since U is open, there exists a $\delta > 0$ such that $B_\delta(a) \subseteq U$. Therefore if $x \in X$ and $x \in B_\delta(a)$, then $x \in X \cap B_\delta(a) \subseteq X \cap U = f^*(B_\varepsilon(f(a))$ and thus $f(x) \in B_\varepsilon(f(a))$. Therefore f is continuous, by Theorem 6.2.2 (iii). □

Remark 6.2.12. It is important that the *inverse image*, under a continuous function, of an open set is, effectively, an open set. The *direct image*, under a continuous function, of an open say may not be an open set.

Remark 6.2.13. We will often use the phrase *f is continuous on X*, by this we mean that the function f is a continuous function with domain X.

6.3. Continuity on Compact Domains

Definition 6.3.1 (Bounded Function)**.** Given a function $f : X \to \mathbb{R}$, we say that f is a bounded function if its range $f(X)$ is a bounded set.

That is, if there exists an $M \in \mathbb{R}$ such that $|f(x)| < M$, for every $x \in X$.

Theorem 6.3.2. *Suppose $f : X \to \mathbb{R}$ is continuous. If $K \subseteq X$ is compact, then $f_*(K)$ is compact.*

Proof. We will use the open cover definition of compactness to prove this theorem. Let $K \subseteq \mathbb{R}$ be compact, and consider an open cover $\{U_\alpha\}_{\alpha \in T}$ of $f(K)$. That is,

$$f_*(K) \subseteq \bigcup_{\alpha \in T} U_\alpha$$

Therefore, we have

$$K \subseteq f^* \left(\bigcup_{\alpha \in T} U_\alpha \right) = \bigcup_{\alpha \in T} f^*(U_\alpha)$$

By Theorem 6.2.11, $f^*(U_\alpha) = X \cap V_\alpha$ for some open set V_α, and hence $f^*(U_\alpha) \subseteq V_\alpha$. Thus,

$$K \subseteq \bigcup_{\alpha \in T} f^*(U_\alpha) \subseteq \bigcup_{\alpha \in T} V_\alpha$$

That is, $\{V_\alpha\}_{\alpha \in T}$ is an open cover of K. Since K is compact, this cover contains a finite subcover, say $\{V_1, \ldots, V_n\}$, i.e.,

$$K \subseteq V_1 \cup \cdots \cup V_n = \bigcup_{i=1}^{n} V_i$$

Let U_i be those open sets such that $f^*(U_i) = X \cap V_i$. Then, note that

$$K = K \cap X \subseteq \left(\bigcup_{i=1}^{n} V_i \right) \cap X = \bigcup_{i=1}^{n} (V_i \cap X) = \bigcup_{i=1}^{n} f^*(U_i)$$

and therefore

$$f_*(K) \subseteq \bigcup_{i=1}^{n} U_i$$

and hence $\{U_1, \ldots, U_n\}$ is a cover of $f_*(K)$.

We have thus proved that the arbitrary open cover $\{U_\alpha\}_{\alpha \in T}$ of $f_*(K)$ contains a finite subcover. Therefore $f_*(K)$ is compact. □

Remark 6.3.3. Keeping the Heine-Borel Theorem (Theorem 5.3.5) in mind, Theorem 6.3.2 says that given a continuous function $f : X \to \mathbb{R}$ and a *closed* and *bounded* set A, then $f_*(A)$ is also a closed and bounded set.

Discussion 6.3.4. Consider the continuous function $f : \mathbb{R} \to \mathbb{R},\ x \mapsto x$, we see that $f(\mathbb{R}) = \mathbb{R}$, i.e., the image of an unbounded set is unbounded. Therefore, the function is unbounded.

Consider instead the continuous function $f : (0,1) \to \mathbb{R}, x \mapsto 1/x$, note that the domain of the function is bounded but its range $f((0,1)) = (1,\infty)$ is unbounded. Therefore, the function is unbounded.

Theorem 6.3.2 tells us that if we take a continuous function whose domain is closed and bounded, then its range is closed and bounded. That is, the function is a bounded function. For example, consider

$$f : [2,3] \to \mathbb{R},\ x \mapsto 1/x$$

then $f([2,3]) = [1/3, 1/2]$, and hence f is a bounded function.

Corollary 6.3.5. *A continuous function on a compact set is bounded.*

Proof. Let $f : K \to \mathbb{R}$ be a continuous function such that K is a compact set. Since f is continuous, and K is compact, the range of the function $f(K)$ is also compact by Theorem 6.3.2. By the Heine-Borel Theorem (Theorem 5.3.5), $f(K)$ is closed and bounded. Since, in particular, $f(K)$ is bounded, thus f is bounded. □

Theorem 6.3.6 (Extreme Value Theorem)**.** *A continuous function on a compact set attains its supremum and infimum.*

Proof. Let $f : K \to \mathbb{R}$ be a continuous function such that K is a compact set. By Corollary 6.3.5, the set $f(K)$ is compact. By HW 6, Problem 6.3, we have that $\sup(f(K))$ and $\inf(f(K))$ exist, and $\sup(f(K)) \in f(K)$ and $\inf(f(K)) \in f(K)$. Therefore, there exist $x, y \in K$ such that $f(x) = \sup(f(K))$ and $f(y) = \inf(f(K))$. □

Example 6.3.7. Consider the function $f : [-2,1] \to \mathbb{R}$ given as $f(x) = x^2$, then the supremum of $f([-2,1])$ is 4, while its infimum is 0. Moreover, these are achieved, since $4 = f(-2)$, where $-2 \in [-2,1]$ and $0 = f(0)$, where $0 \in [-2,1]$.

Remark 6.3.8. The two assumption of the Extreme Value Theorem (Theorem 6.3.6) is that the given function is continuous and its domain is compact. Let us show that these two conditions are necessary.

Consider the function $f : (-2,1] \to \mathbb{R}$ given as $f(x) = x^2$, then the supremum of $f((-2,1])$ is 4, while its infimum is 0. But note that the supremum is not achieved, since $4 = f(-2)$, but $-2 \notin (-2,1]$.

Suppose $f : [0,4] \to \mathbb{R}$ given as

$$f(x) = \begin{cases} 2x & \text{if } x \in [0,2) \\ 2 & \text{if } x \in [2,4] \end{cases}$$

Note that f is continuous, since f is not continuous at 2. We have that $f(2) = 2$. Now consider the sequence $(2 - 1/n)_n$, here $2 - 1/n \in [0, 4]$ for every n, and note that $2 - 1/n \to 2$. Since $2 - 1/n < 2$, we have $f(2 - 1/n) = 4 - 2/n \to 4 \neq f(2)$.

We remark that the supremum of $f([0, 4])$ is 4 but note that the supremum is not achieved, since for no $x \in [0, 4]$ do we have $4 = f(x)$.

Lecture 21

Lemma 6.3.9. *Let $f : X \to \mathbb{R}$ be a function. Suppose f is continuous at $a \in X$ and $f(a) > 0$. Then, there exists a $\delta > 0$ such that $f(x) > 0$ for every $x \in (a - \delta, a + \delta) \cap X$.*

Likewise, if $f(a) < 0$, then there exists a $\delta > 0$ such that $f(x) < 0$ for every $x \in (a - \delta, a + \delta) \cap X$.

Proof. Assume $f(a) > 0$ for $a \in X$, and f is continuous at a. Therefore, for $\varepsilon = f(a)/2 > 0$, we may find a $\delta > 0$ such that for every $x \in X$

$$\text{if } |x - a| < \delta, \text{ then } |f(x) - f(a)| < \varepsilon = f(a)/2$$

Therefore, for any $x \in (a - \delta, a + \delta)$ we have $|f(x) - f(a)| < \varepsilon = f(a)/2$, i.e.,

$$-f(a)/2 < f(x) - f(a) < f(a)/2$$

Hence,

$$f(x) > f(a) - f(a)/2 = f(a)/2 > 0$$

as we wanted to show.

In the case $f(a) < 0$, one may provide a similar proof by taking $\varepsilon = -f(a)/2$. □

Proposition 6.3.10. *If f is continuous on $[a, b]$, and $f(a) \cdot f(b) < 0$ (that is, $f(a)$ and $f(b)$ have different sign), then there there exists a $c \in (a, b)$ such that $f(c) = 0$.*

Proof. We assume without loss of generality that $f(b) < 0 < f(a)$. By Lemma 6.3.9, there exist δ_1, $\delta_2 > 0$ such that $f(u) > 0$ for every $u \in [a, b] \cap (a - \delta_1, a + \delta_1) = [a, a + \delta_1)$ and $f(v) < 0$ for every $v \in [a, b] \cap (b - \delta_2, b + \delta_2) = (b - \delta_2, b]$.

Consider the set

$$A = \{t \,:\, f(x) > 0 \text{ for all } x \in [a, t]\}$$

Note that any $u \in [a, a + \delta_1)$ is such that $u \in A$, and so $[a, a + \delta_1 \subseteq A$ and hence $A \neq \varnothing$. We claim that any $v \in (b - \delta_2, b]$ is an upper bound of A; suppose not, then there exists a $v_0 \in (b - \delta_2, b]$ and $t \in A$ such that $v_0 < t$, and therefore $v_0 \in [a, t]$. Since $t \in A$, we have $f(x) > 0$ for all $x \in [a, t]$, in particular, $f(v_0) > 0$ since $v_0 \in [a, t]$. This contradicts the fact that $v_0 \in (b - \delta_2, b]$, as for every $v \in [a, b] \cap (b - \delta_2, b + \delta_2) = (b - \delta_2, b]$ we have $f(v) < 0$. Thus any $v \in (b - \delta_2, b]$ is an upper bound of A.

Hence, $\sup(A)$ exists and let $c = \sup(A)$; necessarily $a < a + \delta_1/2 \leqslant c \leqslant b - \delta_2/2 < b$. That is, $c \in (a, b)$. We will prove that $f(c) = 0$.

Suppose $f(c) \neq 0$, then

CASE I. assume $f(c) > 0$. By Lemma 6.3.9, there exists a $\delta > 0$ such that $f(x) > 0$ for every $x \in (c - \delta, c + \delta)$; shrinking δ if necessary (see Remark 6.3.13), we may assume that

$$(c - \delta, c + \delta) \subseteq (a, b) \subseteq [a, b].$$

In particular, $f(x) > 0$ for every $x \in (c - \delta, c + \delta/2]$.

Since $c = \sup(A)$, therefore $c - \delta$ is not an upper bound of A, and therefore there exists a $t \in A$ such that $c - \delta < t$. Since $t \in A$, we have that $f(x) > 0$ for every $x \in [a, t]$, and hence $f(x) > 0$ for every $x \in [a, c - \delta]$ since $c - \delta < t$.

Now, we have that $f(x) > 0$ for every $x \in [a, c - \delta]$ and $f(x) > 0$ for every $x \in (c - \delta, c + \delta/2]$ and thus $f(x) > 0$ for every $x \in [a, c + \delta/2]$. Therefore, $c + \delta/2 \in A$, contradicting the fact that $c = \sup(A)$.

CASE II. assume $f(c) < 0$; a similar argument to the one in the previous case again provides us with a contradiction.

Thus, $f(c) = 0$, for this $c \in (a, b)$. □

Theorem 6.3.11 (Intermediate Value Theorem)**.** *If f is continuous on $[a, b]$, and let $\alpha \in \mathbb{R}$ be such that either $f(a) < \alpha < f(b)$ or $f(b) < \alpha < f(a)$, then there there exists a $c \in (a, b)$ such that $f(c) = \alpha$.*

Proof. We assume that $f(b) < \alpha < f(a)$, as the other case follows similarly. Let us define a new function $g : [a, b] \to \mathbb{R}$ as follows

$$g(x) = f(x) - \alpha$$

Therefore g is a continuous function on $[a, b]$ and $g(b) < 0 < g(a)$. By Proposition 6.3.10, there exists a $c \in (a, b)$ such that $g(c) = 0$. That is,

$$0 = g(c) = f(c) - \alpha$$

Hence $f(c) = \alpha$. □

Example 6.3.12. The Intermediate Value Theorem (Theorem 6.3.11) gives us another way to construct n-roots of (positive) real numbers. Consider any $a \in \mathbb{R}$, where $a > 0$, and we wish to construct $a^{1/n}$. If $a = 1$, then $a^{1/n} = 1$, so assume $a \neq 1$. Then, we have $0 < a < 1$ or $a > 1$.

Assume $0 < a < 1$, and consider the function

$$f : [0, 1] \to \mathbb{R},\ x \mapsto x^n.$$

Note that $f(0) = 0$ and $f(1) = 1$, and so $f(0) < a < f(1)$. By the Intermediate Value Theorem (Theorem 6.3.11), there exists a $c \in (0, 1)$ such that $a = f(c) = c^n$. That is, $c = a^{1/n}$.

If $a > 1$, then $1/a < 1$ and hence we know $(1/a)^{1/n}$ exists. Then $a^{1/n} = 1/(1/a)^{1/n}$.

Remark 6.3.13 ("shrinking δ")**.** Suppose we have an open set U, and a point $c \in U$. Further, assume that you have considered a δ-neighbourhood of c, $(c - \delta, c + \delta)$, independent of the open set U.

Then we may "shrink δ" to assume that $(c - \delta, c + \delta) \subseteq U$. This means the following: we first observe that since $(c - \delta, c + \delta)$ and U are open sets, so is $(c - \delta, c + \delta) \cap U$. Moreover, since $c \in (c - \delta, c + \delta) \cap U$, by definition of open sets, we may find a δ' such that

$$(c - \delta', c + \delta') \subseteq (c - \delta, c + \delta) \cap U$$

The "shrunk δ" refers to the δ', as considering $(c - \delta', c + \delta')$, we have a δ'-neighbourhood of c such that $(c - \delta', c + \delta') \subseteq U$.

6.4. Uniform Continuity

Lecture 22

Discussion 6.4.1.

- Consider the function $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = 7x$. To prove f is continuous at each $a \in \mathbb{R}$; you would consider any $\varepsilon > 0$ and some scratch work will tell you that we may take $\delta = \varepsilon/7$. Then, for any $x \in \mathbb{R}$, if $|x - a| < \delta = \varepsilon/7$, then

$$|f(x) - f(a)| = |7x - 7a| = 7 \cdot |x - a| < 7 \cdot \frac{\varepsilon}{7} = \varepsilon$$

  Hence f is a continuous function.

- Consider now the function $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = 10x$. To prove f is continuous at each $a \in \mathbb{R}$, for any $\varepsilon > 0$ you take, taking $\delta = \varepsilon/10$ will make the proof of continuity at a work.

We make two observations:

(a) steeper the slope of these linear polynomial functions, the smaller δ is;

(b) our choice of δ only depends on ε but not a.

Consider, now, the following function.

- Consider the function $f : (0, \infty) \to \mathbb{R}$, where $f(x) = x^2$. To prove f is continuous at each $a \in (0, \infty)$; you would consider any $\varepsilon > 0$ and some scratch work (see Example 6.1.4) will tell you that we may take $\delta = \min(1, \varepsilon/(2a + 1))$. Then, for any $x \in (0, \infty)$, if $|x - a| < \delta$, then

$$\begin{aligned} |f(x) - f(a)| = \left|x^2 - a^2\right| &= |x + a| \cdot |x - a| \\ &< (2a + 1) \cdot |x - a| && \text{since } |x - a| < \delta \leqslant 1 \\ &< (2a + 1) \cdot \frac{\varepsilon}{2a + 1} && \text{since } |x - a| < \delta \leqslant \frac{\varepsilon}{2a + 1} \\ &= \varepsilon \end{aligned}$$

  Hence f is a continuous function.

In this example, we note that δ now depends on both ε and a. Therefore as a gets larger, our choice of δ in turn gets smaller; the function also gets steeper for larger values of a.

This is in contrast with the linear polynomial functions in our previous examples where the choice of δ only depends on ε.

This is another instance, where the behaviour of the function (rule) is sensitive to the choice of domain.

- Consider the function $f : [0, 4] \to \mathbb{R}$, where $f(x) = x^2$. To prove f is continuous at each $a \in [0, 4]$; you would consider any $\varepsilon > 0$ and we claim that $\delta = \varepsilon/8$ works for our proof of continuity of f at a. For any $x \in [0, 4]$, if $|x - a| < \delta = \varepsilon/8$, then

$$\begin{aligned} |f(x) - f(a)| = \left|x^2 - a^2\right| = |x + a| \cdot |x - a| &\leqslant (|x| + |a|) \cdot |x - a| \\ &\leqslant (4 + 4) \cdot |x - a| = 8 \cdot |x - a| < 8 \cdot \frac{\varepsilon}{8} = \varepsilon \end{aligned}$$

  Hence f is a continuous function. Moreover, our δ only depends on ε, and so works for any choice of $a \in [0, 4]$.

In our examples, we have seen two different situations while proving continuity: our choice of δ usually depends on ε, but may or may not depend on the point a at which we are checking continuity. Moreover, this may also depend on the domain of our function.

We distinguish the case where δ only depends on ε and not on any point in the domain of the function; we will call such a function *uniformly continuous*. Our first attempt of the definition is

> Consider a function $f : X \to \mathbb{R}$. We say f is *uniformly continuous* if
> *for all* $\varepsilon > 0$, *there exists a* $\delta > 0$ *such that for all* $a \in X$ *and for every* $x \in X$,
> *if* $|x - a| < \delta$, *then* $|f(x) - f(a)| < \varepsilon$.

Note that in this definition, points a and x are on equal footing. This motivates us to give the following definition of uniform continuity.

Definition 6.4.2 (Uniform Continuity)**.** For a function $f : X \to \mathbb{R}$, we say f is uniformly continuous if *for all* $\varepsilon > 0$, *there exists a* $\delta > 0$ *such that for all* $x, y \in X$, *if* $|x - y| < \delta$, *then* $|f(x) - f(y)| < \varepsilon$.

Example 6.4.3.

(a) As seen in Discussion 6.4.1, the function $f : [0, 4] \to \mathbb{R}$ defined as $f(x) = x^2$ is uniformly continuous. Since for any $\varepsilon > 0$, we may take $\delta = \varepsilon/8$ to give the proof of uniform continuity.

(b) Using Discussion 6.4.1 as an inspiration, the function $g : \mathbb{R} \to \mathbb{R}$ defined as $g(x) = ax + b$, where $a \neq 0$, is uniformly continuous. Since for any $\varepsilon > 0$, we may take $\delta = \varepsilon/|a|$ to give the proof of uniform continuity.

Example 6.4.4. Below is a table of continuous function (on a domain), and the δ that you could take if you tried to prove it was continuous at a. For each, we note whether it was uniformly continuous or not.

$f(x)$ on domain	δ	Uniformly continuous?
x^2 on $(0, \infty)$	$\min(1, \varepsilon/(2a+1))$	No
x^2 on $[0, 4]$	$\varepsilon/8$	Yes
$\sqrt{x}$ on $(0, \infty)$	ε^2	Yes
e^x on $(0, \infty)$	$\min(1, \varepsilon/(e^a(a^a - 1)))$	No
$5x + 2$ on $\mathbb{R}$	$\varepsilon/5$	No
$1/x$ on $(0, \infty)$	$\min(a^2\varepsilon/2, \varepsilon/2)$	No
$1/x$ on $(3, \infty)$	9ε	Yes
$\sin(x)$ on $\mathbb{R}$	ε	Yes
$\cos(x)$ on $\mathbb{R}$	ε	Yes

We also remark that being uniformly continuous is a stronger condition than being continuous.

Example 6.4.5. The function $f : (0, \infty) \to \mathbb{R}$ defined as $f(x) = x^2$ is *not* uniformly continuous.

Proof. Assume, for the sake of contradiction, that f is uniformly continuous. Then, for $\varepsilon = 1$, there exists a $\delta > 0$ such that for every $x, y \in (0, \infty)$, if

$$\text{if } \ |x - y| < \delta, \quad \text{then } \ |f(x) - f(y)| < \varepsilon$$

Now, let $y = 1/\delta > 0$ and $x = y + \delta/2$. Note that $|x - y| = \delta/2$ and therefore $0 < |x - y| < \delta$; hence we must have $|f(x) - f(y)| < \varepsilon = 1$. But observe that

$$|f(x) - f(y)| = |x^2 - y^2| = |x - y|\,|x + y| = \frac{\delta}{2} \cdot (2y + \delta/2) > \frac{\delta}{2} \cdot 2y = \delta \cdot y = 1,$$

a contradiction. Thus, f is not uniformly continuous. □

Proposition 6.4.6. *If $f : K \to \mathbb{R}$ is continuous, and K is compact, then f is uniformly continuous on K.*

Proof. Suppose f is continuous and K is compact. Fix $\varepsilon > 0$, then $\varepsilon/2 > 0$; for every $a \in K$, since f is continuous at a, there exists some $\delta(a) > 0$ such that whenever $x \in K$

$$\text{if } \ |x - a| < \delta(a), \quad \text{then } \ |f(x) - f(a)| < \varepsilon/2 \tag{$*$}$$

Now, observe that

$$\mathcal{U} = \left\{ \left(a - \frac{\delta(a)}{2}, a + \frac{\delta(a)}{2} \right) \right\}_{a \in K}$$

is an open cover of K, and hence by compactness of K this open cover contains a finite subcover

$$\mathcal{F} = \left\{ \left(a_1 - \frac{\delta(a_1)}{2}, a_1 + \frac{\delta(a_1)}{2} \right), \left(a_2 - \frac{\delta(a_2)}{2}, a_2 + \frac{\delta(a_2)}{2} \right), \ldots, \left(a_n - \frac{\delta(a_n)}{2}, a_n + \frac{\delta(a_n)}{2} \right) \right\}$$

of K.

To prove f is uniformly continuous, we wish to find a $\delta > 0$ such that whenever $|x - y| < \delta$, we have $|f(x) - f(y)| < \varepsilon$. Set

$$\delta = \min \left(\frac{\delta(a_1)}{2}, \frac{\delta(a_2)}{2}, \ldots, \frac{\delta(a_n)}{2} \right)$$

and consider $x, y \in K$ such that $|x - y| < \delta$.

Since $x \in K$, and $\mathcal{F}$ is a finite subcover of K, we get that

$$x \in \left(a_k - \frac{\delta(a_k)}{2}, a_k + \frac{\delta(a_k)}{2} \right)$$

for some $1 \leqslant k \leqslant n$.

Hence, $|x - a_k| < \dfrac{\delta(a_k)}{2} < \delta(a_k)$, and since $|x - y| < \delta \leqslant \dfrac{\delta(a_k)}{2}$, we get

$$|y - a_k| = |y - a_k + (x - x)| \leqslant |x - a_k| + |x - y| < \frac{\delta(a_k)}{2} + \frac{\delta(a_k)}{2} = \delta(a_k)$$

Thus, we have

$$|x - a_k| < \delta(a_k) \quad \text{and} \quad |y - a_k| < \delta(a_k)$$

From $(*)$ we conclude that

$$|f(x) - f(a_k)| < \frac{\varepsilon}{2} \quad \text{and} \quad |f(y) - f(a_k)| < \frac{\varepsilon}{2}$$

Therefore,

$$\begin{aligned} |f(x) - f(y)| &= |f(x) - f(y) + (f(a_k) - f(a_k))| \\ &\leqslant |f(x) - f(a_k)| + |f(y) - f(a_k)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

Hence f is uniformly continuous. □

Example 6.4.7. The function $f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$ is *not* uniformly continuous (this is a consequence of Example 6.4.5). But by Theorem 6.4.6, the function

$$g : [a, b] \to \mathbb{R},\ x \mapsto x^2$$

is continuous for any closed interval $[a, b]$.

7. Differentiation

Lecture 23

Remark 7.0.1. In this section, the term *interval* will refer to sets of the form

$$(a,b),\ (a,b],\ [a,b),\ [a,b],\ (a,\infty),\ [a,\infty),\ (-\infty,b),\ (-\infty,b],\ (-\infty,\infty)=\mathbb{R}$$

7.1. The Derivative

Definition 7.1.1 (Derivative)**.** Consider a function $f : I \to \mathbb{R}$, where I is an interval, and any $c \in I$. For $x \neq c$, we may consider the *difference quotient at c*

$$q(x) = \frac{f(x)-f(c)}{x-c}$$

We say f is differentiable at c if the limit

$$\lim_{x\to c} q(x) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c}$$

exists. In which case, the limit is called the derivative of f at c and denoted $f'(c)$.

We say f is *differentiable on* $X \subseteq I$ if f is differentiable at every point in X.

Let $D = \{c \in I \ : \ f \text{ is differentiable at } c\}$, that is, it is the set of all points in I at which f is differentiable. Then f is differentiable on D and the derivative of f is the function

$$f' : D \to \mathbb{R},\ x \mapsto f'(x)$$

Remark 7.1.2. Suppose $f : I \to \mathbb{R}$, where I is an interval, is differentiable at $c \in I$. Then, the following limit exists

$$f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c}$$

We may let $h = x - c$, and we claim that

$$f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c} = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h}$$

For any $\varepsilon > 0$, we may find a $\delta > 0$ such that

$$\text{if } 0 < |x-c| < \delta, \quad \text{then } \left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| < \varepsilon \tag{$*$}$$

Since $h = x - c$, we may re-write $(*)$ as

$$\text{if } 0 < |h-0| < \delta, \quad \text{then } \left| \frac{f(c+h)-f(c)}{h} - f'(c) \right| < \varepsilon$$

which precisely says that

$$f'(c) = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h}$$

as we claimed.

Remark 7.1.3 (Notation of Derivatives)**.** Given a function f, we may also write

$$\frac{df}{dx}(x)$$

for its derivative. That is, $\frac{df}{dx}(x) = f'(x)$

Example 7.1.4.

- Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$. Then, f is differentiable at 2.

  Let us consider the difference quotient at 2

  $$\begin{aligned}\frac{f(2+h)-f(2)}{h} &= \frac{(2+h)^3 - 2^3}{h} \\ &= \frac{(2^3 + 3\cdot 2^2 \cdot h + 3\cdot 2 \cdot h^2 + h^3) - 2^3}{h} \\ &= \frac{12h + 6h^2 + h^3}{h} \\ &= 12 + 6h + h^2\end{aligned}$$

  To show f is differentiable at 2, we consider the limit

  $$\lim_{h\to 0} \frac{f(2+h)-f(2)}{h} = \lim_{h\to 0}(12 + 6h + h^2) = 12 + 6\cdot 0 + 0^2 = 12$$

  as $12 + 6h + h^2$ is a polynomial in h, and so continuous at 0.

  Thus, f is differentiable at 2 and $f'(2) = 12$.

- Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$. Then, f is differentiable at any $c \in \mathbb{R}$.

  Let us consider the difference quotient at c

  $$\begin{aligned}\frac{f(c+h)-f(2)}{h} &= \frac{(c+h)^3 - 2^3}{h} \\ &= \frac{(c^3 + 3\cdot c^2 \cdot h + 3\cdot c \cdot h^2 + h^3) - c^3}{h} \\ &= \frac{3c^2h + 3ch^2 + h^3}{h} \\ &= 3c^2 + 3ch + h^2\end{aligned}$$

  To show f is differentiable at c, we consider the limit

  $$\lim_{h\to 0} \frac{f(c+h)-f(c)}{h} = \lim_{h\to 0}(3c^2 + 3ch + h^2) = 3c^2 + 3c\cdot 0 + 0^2 = 3c^2$$

  as $12 + 6h + h^2$ is a polynomial in h, and so continuous at 0.

  Thus, f is differentiable at any $c \in \mathbb{R}$ and its derivative is

  $$f' : \mathbb{R} \to \mathbb{R},\ x \mapsto 3x^2$$

- Consider the constant function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = k$. Then, f is differentiable at any $c \in \mathbb{R}$.

  Let us consider the difference quotient at c

  $$\begin{aligned} \frac{f(c+h) - f(2)}{h} &= \frac{k-k}{h} \\ &= 0 \end{aligned}$$

  Consider the limit

  $$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h} = \lim_{h\to 0} 0 = 0$$

  Thus, f is differentiable at any $c \in \mathbb{R}$ and its derivative is

  $$f' : \mathbb{R} \to \mathbb{R},\ x \mapsto 0$$

Example 7.1.5. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$. Then, f is differentiable at c if and only if $c \neq 0$.

Proof. ($\Rightarrow$) We prove the contrapositive of the statement: f is not differentiable at 0.

Let us consider the difference quotient at 0

$$\begin{aligned} \frac{f(0+h) - f(0)}{h} &= \frac{|0+h| - |0|}{h} \\ &= \frac{|h|}{h} \end{aligned}$$

To show f is not differentiable at 0, we wish to show that the following limit

$$\lim_{h\to 0} \frac{|h|}{h}$$

does not exist.

Define the function $g : \mathbb{R} \setminus \{0\} \to \mathbb{R},\ h \mapsto |h|/h$, and consider sequences (a_n) and (b_n) in $\mathbb{R} \setminus \{0\}$ where

$$a_n = 1/n \quad \text{and} \quad b_n = -1/n$$

Note that $a_n \to 0$ and $b_n \to 0$, but

$$\begin{aligned} g(a_n) &= \frac{|1/n|}{1/n} & g(b_n) &= \frac{|-1/n|}{-1/n} \\ &= \frac{1/n}{1/n} & g(b_n) &= \frac{1/n}{-1/n} \\ &= 1 & &-1 \end{aligned}$$

So,

$$\lim_{n\to\infty} g(a_n) = 1 \neq -1 = \lim_{n\to\infty} g(b_n)$$

and therefore

$$\lim_{h\to 0} g(h) = \lim_{h\to 0} \frac{|h|}{h}$$

does not exist. Thus, f is not differentiable at 0.

($\Leftarrow$) We wish to prove that f is differentiable at any $c \neq 0$.

CASE I. $c > 0$. We claim that

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = 1$$

Consider any $\varepsilon > 0$. Since $c > 0$, we may find a $\delta > 0$ such that $(c - \delta, c + \delta) \subseteq (0, \infty)$ (e.g., you may take $\delta = c/2$). Therefore, for any $x \in \mathbb{R}$ such that $0 < |x - c| < \delta$, we note that

$$\begin{aligned}
\left| \frac{f(x) - f(c)}{x - c} - 1 \right| &= \left| \frac{|x| - |c|}{x - c} - 1 \right| \\
&= \left| \frac{x - c}{x - c} - 1 \right| \\
&= 0 \\
&< \varepsilon
\end{aligned}$$

Thus, we have proved that

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = 1$$

That is, $f'(c) = 1$ whenever $c > 0$.

CASE II. $c < 0$. We claim that

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = -1$$

A proof for this is similar to the proof given in the previous case.

Hence, f is differentiable at any $c \neq 0$ and its derivative is

$$f' : \mathbb{R} \setminus \{0\} \to \mathbb{R},\ x \mapsto \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

□

Theorem 7.1.6. *Consider a function $f : I \to \mathbb{R}$, where I is an interval. If f is differentiable at $c \in I$, then f is continuous at $c \in I$.*

Proof. To prove f is continuous at $c \in I$, we wish to prove that $\lim_{x \to c} f(x) = f(c)$; this is equivalent to proving

$$\lim_{x \to c} f(x) - f(c) = 0$$

Note that

$$\begin{aligned}
\lim x \to c f(x) - f(c) &= \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \cdot (x - c) \\
&= \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \cdot \lim_{x \to c} (x - c) \\
&= f'(c) \cdot 0 \\
&= 0
\end{aligned}$$

Hence f is continuous at c. □

Remark 7.1.7. Note that the converse of Theorem 7.1.6 is not true. Consider the function

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto |x|\,.$$

Then f is continuous at 0, i.e.

$$\lim_{x\to 0} |x| = 0$$

One can show this by noting that for any sequence (a_n) such that $a_n \to 0$, we have $f(a_n) = |a_n| \to |0| = f(0)$. But, as we saw in Example 7.1.4, f is not differentiable at 0.

7.2. Affine Approximation

Definition 7.2.1 (Affine Approximation)**.** Let $f : I \to \mathbb{R}$ be a function such that f is differentiable at $c \in I$. Define a linear function

$$a : \mathbb{R} \to \mathbb{R},\ x \mapsto f(c) + f'(c)(x - c)$$

The graph of $a(x)$ is a line that passes through the point $(c, f(c))$ on the graph of $f(x)$.

This function is called the affine approximation of f near c

Remark 7.2.2. For a function $f : I \to \mathbb{R}$ differentiable at $c \in I$, the affine approximation a near c is the "best linear approximation" amongst all lines passing through the point $(c, f(c))$. We give this phrase meaning by making the following observations.

- Observe that
$$\lim_{x\to c} f(x) - a(x) = \lim_{x\to c} f(x) - f(c) - f'(c)(x - c) = 0$$
since f is continuous at c. Therefore *"$a(x)$ get closer to $f(x)$ as x approaches c."*

- Now, observe that
$$\begin{aligned}\lim_{x\to c} \frac{a(x) - f(x)}{x - c} &= \lim_{x\to c} \frac{f(x) - f(c) - f'(c)(x - c)}{x - c} \\ &= \lim_{x\to c} \frac{f(x) - f(c)}{x - c} - f'(c) \\ &= f'(c) - f'(c) \\ &= 0\end{aligned}$$
Since $a(x) - f(x) \to 0$ and $x - c \to 0$ as x approaches c, for
$$\frac{a(x) - f(x)}{x - c} \to 0 \quad \text{as } x \text{ approaches } c$$
we may say that *"$a(x)$ get closer to $f(x)$ faster than x gets closer to c, as x approaches c."*

- Consider any other linear function

$$p(x) = f(c) + m(x - c)$$

passing through $(c, f(c))$ with $m \neq f'(c)$. First note that

$$\lim_{x \to c} f(x) - p(x) = \lim_{x \to c} f(x) - f(c) - m(x - c) = 0$$

since f is continuous at c. Now, observe that

$$\begin{aligned}
\lim_{x \to c} \frac{a(x) - f(x)}{p(x) - f(x)} &= \lim_{x \to c} \frac{f(x) - f(c) - f'(c)(x - c)}{f(x) - f(c) - m(x - c)} \\
&= \lim_{x \to c} \frac{\dfrac{f(x) - f(c)}{x - c} - f'(c)}{\dfrac{f(x) - f(c)}{x - c} - m} \\
&= \lim_{x \to c} \frac{f'(c) - f'(c)}{f'(c) - m} && \text{since } m \neq f'(c) \\
&= 0
\end{aligned}$$

Therefore, we may say that

"$a(x)$ get closer to $f(x)$ faster than $p(x)$ gets closer to $f(x)$, as x approaches c."

In this way, $a(x)$ is the "best linear approximation" of $f(x)$ near c.

Example 7.2.3. For $f(x) = x^3$, we have computed that $f'(2) = 12$ in Example 7.1.4. The best linear approximation of f near $c = 2$ has equation

$$a(x) = f(2) + f'(2)(x - 2) = 8 + 12(x - 2) = 12x - 16.$$

7.3. Differentiability Rules

Lecture 24

Remark 7.3.1 (Notation)**.** Suppose f and g are function and $k \in \mathbb{R}$. Then we define

- $(f + g)(x) := f(x) + g(x)$
- $(f - g)(x) := f(x) - g(x)$
- $(fg)(x) := f(x) \cdot g(x)$
- $\left(\dfrac{f}{g}\right)(x) := \dfrac{f(x)}{g(x)}$
- $(kf)(x) := k \cdot f(x)$

Proposition 7.3.2 (Sum & Difference Rules)**.** *Let $f, g : I \to \mathbb{R}$ be function, where I is an interval, that are differentiable at $c \in I$. Then,*

(i) $(f + g)'(c) = f'(c) + g'(c)$

(ii) $(f - g)'(c) = f'(c) - g'(c)$

Proof. We prove (i) and (ii) together. Note that

$$\begin{aligned}\lim_{x\to c}\frac{(f\pm g)(x)-(f\pm g)(c)}{x-c} &= \lim_{x\to c}\frac{(f(x)\pm g(x))-(f(c)\pm g(c))}{x-c}\\ &= \lim_{x\to c}\frac{(f(x)-f(c))\pm(g(x)-g(c))}{x-c}\\ &= \lim_{x\to c}\frac{f(x)-f(c)}{x-c}\pm\lim_{x\to c}\frac{g(x)-g(c)}{x-c}\\ &= f'(c)\pm g'(c)\end{aligned}$$

Therefore the limit exists, and $(f\pm g)'(c) = f'(c)\pm g'(c)$. □

Theorem 7.3.3 (Product Rule)**.** *Let $f, g : I \to \mathbb{R}$ be function, where I is an interval, that are differentiable at $c \in I$. Then,*

$$(fg)'(c) = f'(c)g(c) + f(c)g'(c)$$

Proof. Let us first consider the difference quotient of fg at c

$$\begin{aligned}\frac{(fg)(x)-(fg)(c)}{x-c} &= \frac{f(x)g(x)-f(c)g(c)}{x-c}\\ &= \frac{f(x)g(x)-f(c)g(c)+(f(x)g(c)-f(x)g(c))}{x-c}\\ &= \frac{(f(x)g(c)-f(c)g(c))+(f(x)g(x)-f(x)g(c))}{x-c}\\ &= \frac{f(x)g(c)-f(c)g(c)}{x-c}+\frac{f(x)g(x)-f(x)g(c)}{x-c}\\ &= \frac{f(x)-f(c)}{x-c}\cdot g(c)+f(x)\cdot\frac{g(x)-g(c)}{x-c}\end{aligned}$$

Now, we note that

$$\begin{aligned}\lim_{x\to c}\frac{(fg)(x)-(fg)(c)}{x-c} &= \lim_{x\to c}\frac{f(x)-f(c)}{x-c}\cdot g(c)+f(x)\cdot\frac{g(x)-g(c)}{x-c}\\ &= \lim_{x\to c}\frac{f(x)-f(c)}{x-c}\cdot g(c)+\lim_{x\to c}f(x)\cdot\frac{g(x)-g(c)}{x-c}\\ &= \lim_{x\to c}\frac{f(x)-f(c)}{x-c}\cdot\lim_{x\to c}g(c)+\lim_{x\to c}f(x)\cdot\lim_{x\to c}\frac{g(x)-g(c)}{x-c}\\ &= f'(c)g(c)+f(c)g'(c)\end{aligned}$$

Therefore the limit exists, and $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$. □

Corollary 7.3.4 (Constant Rule)**.** *Let $f : I \to \mathbb{R}$ be function, where I is an interval, that is differentiable at $c \in I$. Then,*

$$(kf)'(c) = k\cdot f'(x)$$

for any $k \in \mathbb{R}$

Proof. Apply the Product Rule to f and the constant function k. □

Theorem 7.3.5 (Quotient Rule)**.** *Let $f, g : I \to \mathbb{R}$ be function, where I is an interval, that are differentiable at $c \in I$, such that $g(c) \neq 0$. Then,*

$$\left(\frac{f}{g}\right)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{g(c)^2}$$

Proof. For g, define the function $(1/g)(x) = 1/g(x)$. Therefore, the quotient rule is the product rule applied to the functions f and $1/g$. We will therefore prove first that

$$\left(\frac{1}{g}\right)'(c) = -\frac{g'(c)}{g(c)^2}$$

and the quotient rule then follows from this.

We leave it to the reader to finish this proof. □

Corollary 7.3.6 (Power Rule)**.** *Consider the function $p_n(x) = x^n$, where $n \in \mathbb{Z}$ such that $n \neq 0$. Then*

$$p_n'(x) = nx^{n-1}$$

(assume $x \neq 0$ when $n < 0$).

Proof. We consider two cases, the case of $n > 0$ follows the Product Rule and Principle of Mathematical Induction.

The case of $n < 0$ follows from applying the previous case to $-n > 0$ and the Quotient Rule. □

Theorem 7.3.7 (Chain Rule)**.** *Let $f : I \to \mathbb{R}$ and $g : J \to \mathbb{R}$ be functions, where I, J are interval and $f(I) \subseteq J$. Then, if f is differentiable at $c \in I$ and g is differentiable at $f(c) \in J$, then*

$$(g \circ f)'(c) = g'(f(c)) \cdot f'(c).$$

Proof. Let's start by defining an auxiliary function on J

$$\phi(y) = \begin{cases} \dfrac{g(y) - g(f(c))}{y - f(c)} & y \neq f(c) \\ g'(f(c)) & y = f(c) \end{cases}$$

Since g is differentiable at $f(c)$, then

$$\lim_{y \to f(c)} \phi(y) = \lim_{y \to f(c)} \frac{g(y) - g(f(c))}{y - f(c)} = g'(f(c)) = \phi(f(c))$$

and therefore ϕ is continuous at $f(c)$.

Furthermore, since f is differentiable at c, it is continuous at c. So, we have that f is continuous at c, and ϕ is continuous at $f(c)$, and thus $\phi \circ f$ is continuous at c; in particular,

$$\lim_{x \to c} \phi(f(x)) = \lim_{x \to c} (\phi \circ f)(x) = (\phi \circ f)(c) = \phi(f(c)) = g'(f(c))$$

Rewriting the above expression, we get the following expression which is valid on all of J.

$$g(y) - g(f(c)) = (y - f(c))\phi(y)$$

Now, for $y = f(x) \in f(I)$, we have

$$\frac{g(f(x)) - g(f(c))}{x - c} = \frac{(f(x) - f(c))\phi(f(x))}{x - c} = \phi(f(x)) \cdot \frac{f(x) - f(c)}{x - c}$$

Therefore,

$$\begin{aligned}\lim_{x\to a} \frac{g(f(x)) - g(f(c))}{x - c} &= \lim_{x\to c} \phi(f(x)) \cdot \frac{f(x) - f(c)}{x - c} \\ &= \lim_{x\to c} \phi(f(x)) \cdot \lim_{x\to c} \frac{f(x) - f(c)}{x - c} = g'(f(c)) \cdot f'(c)\end{aligned}$$

Therefore the limit exists, and $(g \circ f)'(c) = g'(f(c)) \cdot f'(c)$. □

7.4. Local Minimum and Maximums

Lecture 25

Definition 7.4.1. Consider a function $f : X \to \mathbb{R}$, and $c \in X$.

- f has a local maximum at c if there exists a $\delta > 0$ such that for all $x \in X$, whenever $|x - c| < \delta$, or equivalently $x \in (c - \delta, c + \delta)$, we have $f(x) \leqslant f(c)$.
- f has a local minimum at c if there exists a $\delta > 0$ such that for all $x \in X$, whenever $|x - c| < \delta$, or equivalently $x \in (c - \delta, c + \delta)$, we have $f(c) \leqslant f(x)$.

Remark 7.4.2. Recall that in this chapter, *interval* refers to one of the following sets (a, b), $(a, b]$, $[a, b)$, $[a, b]$, (a, ∞), $[a, \infty)$, $(-\infty, b)$, $(-\infty, b]$, $(-\infty, \infty) = \mathbb{R}$.

In what follows, many results will require us to consider an interval which also an open set, and therefore refers to one of the following sets

$$(a, b),\ (a, \infty),\ [a, \infty),\ (-\infty, \infty) = \mathbb{R}$$

Proposition 7.4.3. *Suppose $f : I \to \mathbb{R}$, where I is an interval and open, is differentiable at $c \in I$. If f has a local maximum, or minimum, at c, then $f'(c) = 0$.*

Proof. Since f is differentiable at $c \in I$, we know that

$$f'(c) = \lim_{x\to c} \frac{f(x) - f(c)}{x - c}$$

Suppose c is a local maximum; and therefore there exists a $\delta > 0$ such that for all $x \in I$ and $x \in (c - \delta, c + \delta)$, we have $f(x) \leqslant f(c)$. We may shrink $\delta > 0$ to assume that $(c - \delta, c + \delta) \subseteq I$, since I is open, and hence $f(x) \leqslant f(c)$ for every $x \in (c - \delta, c + \delta)$. Thus,

$$\frac{f(x) - f(c)}{x - c} \text{ is } \begin{cases} \leqslant 0 & \text{if } c < x < c + \delta \\ \geqslant 0 & \text{if } c - \delta < x < c \end{cases}$$

The sequential characterisation of limits tells us that if we consider a sequence (a_n) in I such that, $a_n \neq c$ for every n, and $a_n \to c$, then

$$\frac{f(a_n) - f(c)}{a_n - c} \to f'(c)$$

Consider the sequence (x_n), where $x_n = c - \delta/n$. Then $x_n \to c$ and $c - \delta < x_n < c$ for every n; therefore,

$$\frac{f(x_n) - f(c)}{x_n - c} \geqslant 0, \text{ for every } n, \quad \text{and} \quad \frac{f(x_n) - f(c)}{x_n - c} \to f'(c)$$

Hence, we conclude that $f'(c) \geqslant 0$.

Consider the sequence (y_n), where $y_n = c + \delta/n$. Then $y_n \to c$ and $c < y_n < c + \delta$ for every n; therefore,

$$\frac{f(y_n) - f(c)}{y_n - c} \leqslant 0, \text{ for every } n, \quad \text{and} \quad \frac{f(y_n) - f(c)}{y_n - c} \to f'(c)$$

Hence, we conclude that $f'(c) \leqslant 0$.

Thus, $f'(c) = 0$, as we wanted to show.

A similar argument tells us that $f'(c) = 0$, when c is a local minimum. □

Theorem 7.4.4 (Rolle's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous and differentiable on (a, b). If $f(a) = f(b)$, then there exists a $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. Since f is continuous on $[a, b]$, by the Extreme Value Theorem we may find u, $v \in [a, b]$ such that $f(u) = \sup(f([a, b]))$ and $f(v) = \sup(f([a, b]))$. Therefore u is a (local) maximum of f, while v is a (local) minimum of f.

If $u \in (a, b)$ or $v \in (a, b)$, then $f'(u) = 0$ or $f'(v) = 0$ respectively, by Proposition 7.4.3.

Assume now that u and v are end-points of $[a, b]$; without loss of generality let $u = a$ and $v = b$. Then, for all $x \in [a, b]$, we have

$$f(a) \geqslant f(x) \geqslant f(b)$$

Since $f(a) = f(b)$ by assumption, we observe that $f(a) = f(x) = f(b)$ for every $x \in [a, b]$. That is, if we let $k = f(a)$, we have that $f(x) = k$ for every $x \in [a, b]$. Therefore, f is a constant function, and hence $f'(x) = 0$ for every x. Thus $f'(c) = 0$, for any $c \in (a, b)$. □

Theorem 7.4.5 (Mean Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous and differentiable on (a, b). Then there exists some $c \in (a, b)$ such that*

$$f(b) - f(a) = f'(c) \cdot (b - a)$$

Proof. Assume $f(a) = f(b)$, then by Rolle's Theorem (Theorem 7.4.4) there exists a $c \in (a, b)$ such that $f'(c) = 0$. Hence,

$$f(b) - f(a) = 0 = f'(c) \cdot (b - a)$$

Now, assume that $f(a) \neq f(b)$. Consider the function

$$s(x) = \left(\frac{f(b) - f(a)}{b - a}\right)(x - a) + f(a)$$

"the secant line from $f(a)$ to $f(b)$". Define $g : [a, b] \to \mathbb{R}$ as

$$g(x) = f(x) - s(x)$$

Observe that as a linear function, $s(x)$ is continuous on $[a, b]$ and differentiable on (a, b) and

$$s'(x) = \frac{f(b) - f(a)}{b - a}.$$

Therefore g is continuous on $[a, b]$ and differentiable on (a, b), as a difference of functions with this property. Moreover, observe that

$$g(a) = f(a) - s(a) = 0 = f(b) - s(b) = g(b)$$

Hence, by Rolle's Theorem (Theorem 7.4.4) there exists a $c \in (a, b)$ such that $g'(c) = 0$. That is,

$$0 = g'(c) = f'(c) - s'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$$

Thus, we have found $c \in (a, b)$ such that $f(b) - f(a) = f'(c) \cdot (b - a)$. □

Corollary 7.4.6 (Sufficient Condition for Constantness)**.** *Let $f : I \to \mathbb{R}$, where I is an interval, be a differentiable function. If $f'(x) = 0$ for every $x \in I$, then f is constant on I.*

Proof. Let $f : I \to \mathbb{R}$, where I is an interval, be a differentiable function such that $f'(x) = 0$ for every $x \in I$. Consider any pair $x_1, x_2 \in I$ with $x_1 < x_2$. To show f is constant, it is enough to show $f(x_1) = f(x_2)$.

Since f is differentiable on I, it is differentiable on $[x_1, x_2]$. Hence, by the Mean Value Theorem (Theorem 7.4.5), there exists a $c \in (x_1, x_2) \subseteq I$ such that

$$f(x_1) - f(x_2) = f'(c) \cdot (x_1 - x_2) = 0$$

Thus $f(x_1) = f(x_2)$, and therefore f is a constant function. □

Corollary 7.4.7. *Let $f, g : I \to \mathbb{R}$, where I is an interval, be differentiable functions. If $f'(x) = g'(x)$ for every $x \in I$, then $f(x) = g(x) + C$ for some $C \in \mathbb{R}$.*

Proof. Define $h : I \to \mathbb{R}$, where $h = f - g$. Using the Sum Rule, we note that $f - g$ is differentiable on I and

$$h'(x) = f'(x) - g'(x) = 0$$

for all $x \in I$. Then, by Corollary 7.4.6, we conclude that h is constant. Therefore $h(x) = C$, for some $C \in \mathbb{R}$, for every $x \in I$. Hence $f(x) = g(x) + C$, as we wished to prove. □

Lecture 26

Definition 7.4.8 (Monotone Functions)**.** Consider $f : I \to \mathbb{R}$, where I is an interval, and $x_1, x_2 \in I$ such that $x_1 < x_2$. If

- $f(x_1) \leqslant f(x_2)$, then f is said to be monotonically increasing.
- $f(x_1) < f(x_2)$, then f is said to be strictly monotonically increasing.

- $f(x_1) \geqslant f(x_2)$, then f is said to be monotonically decreasing.
- $f(x_1) > f(x_2)$, then f is said to be strictly monotonically decreasing.

A function is said to be monotone (or monotonic) if it is either monotonically increasing *or* monotonically decreasing.

Corollary 7.4.9 (Sign of Derivatives dictates Monotonicity)**.** *Let $f : I \to \mathbb{R}$, where I is an interval, be a differentiable function.*

- *f is monotonically increasing if and if $f'(x) \geqslant 0$ for every $x \in I$. In particular, if $f'(x) > 0$ for every $x \in I$, then f is strictly monotonically increasing.*
- *f is monotonically decreasing if and if $f'(x) \leqslant 0$ for every $x \in I$. In particular, if $f'(x) < 0$ for every $x \in I$, then f is strictly monotonically decreasing.*

Proof. We will prove (i), the proof of (ii) is similar.

($\Rightarrow$) Assume f is monotonically increasing. Consider any $c \in I$, since f is monotonically increasing, for any $x \in I$ such that $x \neq c$, we have

$$\frac{f(x) - f(c)}{x - c} \geqslant 0, \quad \text{thus, } f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \geqslant 0$$

for all $c \in I$, as we wanted to show.

($\Leftarrow$) Assume $f'(x) \geqslant 0$ (resp. $f'(x) > 0$) for every $x \in I$. Consider any $x_1, x_2 \in I$ such that $x_1 < x_2$. Since f is differentiable on I, it is differentiable on $[x_1, x_2]$. Hence, by the Mean Value Theorem (Theorem 7.4.5), there exists a $c \in (x_1, x_2) \subseteq I$ such that

$$f(x_1) - f(x_2) = f'(c) \cdot (x_1 - x_2) \leqslant 0 \qquad \text{if } f'(x) \geqslant 0 \text{ for every } x \in I$$

$$f(x_1) - f(x_2) = f'(c) \cdot (x_1 - x_2) < 0 \qquad \text{if } f'(x) > 0 \text{ for every } x \in I$$

Thus $f(x_1) \leqslant f(x_2)$ (resp. $f(x_1) < f(x_2)$), and therefore f is a monotonically increasing (resp. strictly monotonically increasing) function. □

Remark 7.4.10. In the light of Corollary 7.4.9, it is worth mentioning that given a strictly monotonically increasing function $f : I \to \mathbb{R}$, one may still have some $c \in I$ such that $f'(c) = 0$. Therefore, the conclusion of Corollary 7.4.9, that is, $f'(x) \geqslant 0$ for every $x \in I$, cannot be relaxed for strictly monotonic increasing function.

For example, consider $f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^3$. This is a strictly increasing function, but $f'(0) = 0$.

7.5. L'Hôpital's Rule

Theorem 7.5.1 (Cauchy Mean Value Theorem)**.** *Let $f, g : [a, b] \to \mathbb{R}$ be continuous and differentiable on (a, b). Then there exists a $c \in (a, b)$ such that*

$$(f(b) - f(a)) \cdot g'(c) = (g(b) - g(a)) \cdot f'(c)$$

Proof. Skipped. □

Theorem 7.5.2 (L'Hôpital's Rule)**.** *For $a \in \mathbb{R}$, let I be an interval and open such that $a \in I$. Consider functions $f, g : I \to \mathbb{R}$ that are differentiable on I. Suppose also $g'(x) \neq 0$ on I. Then, if*

$$\lim_{x\to a} f(x) = 0 \quad \textit{and} \quad \lim_{x\to a} g(x) = 0,$$

then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)},$$

provided the limit $\lim_{x\to a} \dfrac{f'(x)}{g'(x)}$ *exists.*

Proof. Let

$$L = \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

Fix $\varepsilon > 0$. Since $\varepsilon/2 > 0$, there exists a $\delta > 0$ such that for all $x \in I$, whenever $x \in (a-\delta, a+\delta)$, we have

$$\left| \frac{f'(x)}{g'(x)} - L \right| < \frac{\varepsilon}{2}$$

Since I is an open interval, we may shrink δ and assume that $(a-\delta, a+\delta) \subseteq I$. Therefore, for any $x \in (a-\delta, a+\delta)$, we have

$$L - \frac{\varepsilon}{2} < \frac{f'(x)}{g'(x)} < L + \frac{\varepsilon}{2} \tag{1}$$

Consider any $t, s \in (a-\delta, a+\delta)$ such that $t < a < s$. Since f and g are differentiable on I, they are differentiable on $[t, s]$. Hence, by the Cauchy Mean Value Theorem (Theorem 7.5.1), there exists a $c \in (t, s) \subseteq (a-\delta, a+\delta)$ such that

$$\frac{f'(c)}{g'(c)} = \frac{f(t) - f(s)}{g(t) - g(s)} \tag{2}$$

(the condition $g'(x) \neq 0$ for every $x \in I$ guarantees that $g(t) \neq g(s)$)

Since $c \in (a-\delta, a+\delta)$, by (1) we observe that

$$L - \frac{\varepsilon}{2} < \frac{f'(c)}{g'(c)} < L + \frac{\varepsilon}{2}$$

Now, using (2) we obtain

$$L - \frac{\varepsilon}{2} < \frac{f(t) - f(s)}{g(t) - g(s)} < L + \frac{\varepsilon}{2} \tag{3}$$

for any $t, s \in (a-\delta, a+\delta)$ such that $t < a < s$. In particular, we have from (3)

$$L - \frac{\varepsilon}{2} < \frac{f(a-\delta/n) - f(s)}{g(a-\delta/n) - g(s)} < L + \frac{\varepsilon}{2}$$

since $a - \delta/n \in (a-\delta, a) \subseteq (a-\delta, a+\delta)$ for every n. Hence

$$\lim_{n\to\infty} \left(L - \frac{\varepsilon}{2} \right) < \lim_{n\to\infty} \frac{f(a-\delta/n) - f(s)}{g(a-\delta/n) - g(s)} < \lim_{n\to\infty} \left(L + \frac{\varepsilon}{2} \right)$$

Since $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$, we obtain

$$L - \varepsilon < L - \frac{\varepsilon}{2} \leqslant \frac{-f(s)}{-g(s)} \leqslant L + \frac{\varepsilon}{2} < L + \varepsilon$$

Thus, for any $s \in (a, a+\delta)$ we have

$$L - \varepsilon < \frac{f(s)}{g(s)} < L + \varepsilon \tag{4}$$

Similarly, we have from (3)

$$L - \frac{\varepsilon}{2} < \frac{f(t) - f(a+\delta/n)}{g(t) - g(a+\delta/n)} < L + \frac{\varepsilon}{2}$$

since $a + \delta/n \in (a, a+\delta) \subseteq (a-\delta, a+\delta)$ for every n. Hence

$$\lim_{n\to\infty}\left(L - \frac{\varepsilon}{2}\right) < \lim_{n\to\infty} \frac{f(t) - f(a+\delta/n)}{g(t) - g(a+\delta/n)} < \lim_{n\to\infty}\left(L + \frac{\varepsilon}{2}\right)$$

Since $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$, we obtain

$$L - \varepsilon < L - \frac{\varepsilon}{2} \leqslant \frac{f(t)}{g(t)} \leqslant L + \frac{\varepsilon}{2} < L + \varepsilon$$

Thus, for any $t \in (a-\delta, a)$ we have

$$L - \varepsilon < \frac{f(t)}{g(t)} < L + \varepsilon \tag{5}$$

Therefore, (4) and (5) tells us that for any $x \in I$ such that $0 < |x - a| < \delta$ we have

$$L - \varepsilon < \frac{f(x)}{g(x)} < L + \varepsilon, \quad \text{and hence,} \quad \left|\frac{f(x)}{g(x)} - L\right| < \varepsilon$$

Thus,

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L = \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

as we wanted to show. □

7.6. Inverse Functions

Lecture 27

Remark 7.6.1 (Inverse Functions)**.** Consider a function $f : X \to Y$ that is bijective, i.e., it is injective and surjective. Then, this is equivalent to the existence of a function $f^{-1} : Y \to X$ such that

$$(f^{-1} \circ f)(x) = x, \quad \text{for every } x \in X; \text{ and}$$
$$(f \circ f^{-1})(y) = y, \quad \text{for every } y \in Y$$

Given a function f,

- when and where can we construct f^{-1}?
- when we know f^{-1} exist, then when is it differentiable, and how to compute $(f^{-1})'$?

Theorem 7.6.2 (Inverse Function Theorem). *Let $f : I \to \mathbb{R}$ be a differentiable function, where I is an interval and open, such that f' is continuous and $f(I) = J$, where J is another interval and also open. Suppose either $f'(x) > 0$ or $f'(x) < 0$ for every $x \in I$. Then $f : I \to J$ is bijective, $f^{-1} : J \to I$ is differentiable and*

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

for every $y \in J$. Moreover, $(f^{-1})'$ is continuous.

Proof. Assume $f'(x) > 0$ for every $x \in I$, the other case follows similarly. Then by Corollary 7.4.9, we have that f is a strictly monotonically increasing function on I.

We will now show that f is an injective function; consider any $x_1, x_2 \in I$ such that $x_1 \neq x_2$. We may assume, without loss of generality, that $x_1 < x_2$. Therefore, since f is strictly increasing, we have $f(x_1) < f(x_2)$; hence, in particular, $f(x_1) \neq f(x_2)$.

Moreover, by the Intermediate Value Theorem (Theorem 6.3.11), we obtain

$$f((x_1, x_2)) = (f(x_1), f(x_2)).$$

Thus, f is injective, and since $f(I) = J$, we have

$$f : I \to J$$

is a bijective function. Therefore, we may consider the inverse function

$$f^{-1} : J \to I$$

Consider any $v \in J$ and let $u = f^{-1}(v)$, we will show that f^{-1} is continuous at v.

Since $f'(u) > 0$ and set $M := 2/f'(u)$. By assumption, since f' is continuous at u, for $\varepsilon = 1/M > 0$, there exists a $\delta > 0$ such that whenever $x \in (u - \delta, u + \delta)$, we have

$$\left|f'(x) - f'(u)\right| < \frac{1}{M}, \quad \text{and hence} \quad -\frac{1}{M} + f'(u) < f'(x) < f'(u) + \frac{1}{M}$$

Thus, for any $x \in (u - \delta, u + \delta)$

$$\left|f'(x)\right| \geqslant f'(x) > f'(u) - \frac{f'(u)}{2} = \frac{f'(u)}{2} = \frac{1}{M}$$

For every $x \in (u - \delta, u + \delta)$ and u, by the Mean Value Theorem (Theorem 7.4.5), there exists a c, between x and u, such that

$$|f(x) - f(u)| = \left|f'(c)\right| \cdot |x - u| > \frac{1}{M} \cdot |x - u|$$

For any $y \in f((u - \delta, u + \delta)) = (f(u - \delta), f(u + \delta))$, since $f^{-1}(y) \in (u - \delta, u + \delta)$ and therefore

$$\left|f^{-1}(y) - f^{-1}(v)\right| < M \cdot |y - v|$$

Moreover, observe that since $u - \delta < u < u + \delta$, we get $f(u - \delta) < f(u) = v < f(u + \delta)$.

Consider any $\varepsilon > 0$, and the open interval $(v - \varepsilon/M, v - \varepsilon/M)$. We observe that

$$v \in (v - \varepsilon/M, v + \varepsilon/M) \cap (f(u - \delta), f(u + \delta)),$$

and so $(v - \varepsilon/M, v + \varepsilon/M) \cap (f(u - \delta), f(u + \delta))$ is a non-empty open set. Therefore, there exists a δ' such that

$$(v - \delta', v + \delta') \subseteq (v - \varepsilon/M, v + \varepsilon/M) \cap (f(u - \delta), f(u + \delta))$$

Hence, for any $y \in (v - \delta', v + \delta')$

$$\begin{aligned}
\left| f^{-1}(y) - f^{-1}(v) \right| &< M \cdot |y - v| && \text{since } y \in (v - \delta', v + \delta') \subseteq (f(u - \delta), f(u + \delta)) \\
&< M \cdot \frac{\varepsilon}{M} && \text{since } y \in (v - \delta', v + \delta') \subseteq (v - \varepsilon/M, v + \varepsilon/M) \\
&= \varepsilon
\end{aligned}$$

Thus,

$$\lim_{y \to v} f^{-1}(y) = f^{-1}(v)$$

Therefore f^{-1} is continuous at v.

We will now show that f^{-1} is differentiable at v. We define the function

$$\rho(x) = \begin{cases} \dfrac{f(x) - f(u)}{x - u} - f'(u) & \text{if } x \neq u \\ 0 & \text{if } x = u \end{cases}$$

Note that ρ is continuous at any $x \neq u$. Since f is differentiable at $u = f^{-1}(v)$, we note that

$$\lim_{x \to u} \rho(x) = 0 = \rho(u),$$

and so it also continuous at u; i.e. ρ is a continuous function on I. We now observe that

$$f(x) - f(u) = (x - u) \cdot (f'(u) + \rho(x))$$

For $y = f(x) \neq v$, we have

$$\begin{aligned}
f^{-1}(y) - f^{-1}(v) = x - u &= \frac{f(x) - f(u)}{f'(u) + \rho(x)} \\
&= \frac{y - v}{f'(u) + \rho(x)} \\
&= \frac{1}{f'(u)} \cdot (y - v) + (y - v) \cdot \eta(y)
\end{aligned}$$

where

$$\eta(y) = -\frac{\rho(x)}{f'(u) + \rho(x)} = -\frac{\rho(f^{-1}(y))}{f'(u) + \rho(f^{-1}(u))}$$

Since f^{-1} is continuous at v, and ρ is continuous at $u = f^{-1}(v)$, hence $\rho \circ f^{-1}$ is continuous at v. Thus,

$$\lim_{y \to v} (\rho \circ f^{-1})(y) = (\rho \circ f^{-1})(v) = \rho(u) = 0$$

and thus $\lim_{y \to v} \eta(y) = 0$. Which tells us,

$$\lim_{y \to v} \frac{f^{-1}(y) - f^{-1}(v)}{y - v} = \lim_{y \to v} \frac{1}{f'(u)} + \eta(y) = \frac{1}{f'(u)}$$

Therefore the limit exists, and

$$(f^{-1})'(v) = \frac{1}{f'(u)} = \frac{1}{f'(f^{-1}(v))}$$

Furthermore, since f' is continuous, the above formulation of $(f^{-1})'$ in terms of f', tells us that $(f^{-1})'$ is also continuous. □

Theorem 7.6.3 (Local Inverse Function Theorem)**.** *Suppose f is defined and continuously differentiable in a neighbourhood of u. If $f'(u) \neq 0$, then there exists a $\delta > 0$ such that*

$$f : (u - \delta, u + \delta) \to f((u - \delta, u + \delta))$$

has an inverse that is continuously differentiable.

Proof. Since $f'(u) \neq 0$, by continuity of f', there exists a $\delta > 0$ such that for every $x \in (u - \delta, u + \delta)$ we have that $f'(x)$ and $f'(u)$ have the same sign (Lemma 6.3.9). We may now apply Theorem 7.6.2 to get our result. □

Example 7.6.4 (Example from Worksheet)**.** Consider the exponential function $\exp : \mathbb{R} \to \mathbb{R}$ as defined in Worksheet B. We observe that we have

$$\exp'(x) = \exp(x) > 0$$

for every $x \in \mathbb{R}$. Moreover, since $\exp(x) > 1 + x$ for every $x > 0$, we have that

$$\lim_{x \to \infty} \exp(x) = \infty$$

Therefore $\exp(\mathbb{R}) = (0, \infty)$.

Hence, by the Inverse Function Theorem 7.6.2, exp has an inverse function

$$\ln : (0, \infty) \to \mathbb{R}$$

that is continuously differentiable, and such that

$$\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}$$

We call ln the logarithm function, and it has the following properties.

- Since $\exp(0) = 1$, therefore $\ln(1) = 0$

- For any $x, y \in (0, \infty)$, we may find $a, b \in \mathbb{R}$ such that $x = \exp(a)$ and $y = \exp(b)$, and

$$\ln x + \ln y = a + b = \ln(\exp(a + b)) = \ln(\exp(a)\exp(b)) = \ln(xy)$$

Example 7.6.5. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f(x) = \begin{cases} x + x^2 \sin \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

then f is differentiable, and

$$f'(x) = \begin{cases} 1 + 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

f' is not continuous at 0. One can show that f does not have a local inverse, and hence the assumption for f' to be continuous is necessary for the Inverse Function Theorem (Theorem 7.6.2).

References

[Cum18] J. Cummings, *Real Analysis: A Long-Form Mathematics Textbook*, CreateSpace Independent Publishing Platform, 2018.

[Mon22] F. Monard, *Introductory Real Analysis - Lecture Notes* (unpublished), University of California, Santa Cruz, Winter 2022.