Undergraduate Mathematics

Ryan Joo Rui An

The mathematician does not study mathematics because it is useful; he studies it because he delights in it and he delights in it because it is beautiful.

— Henri Poincaré (1854–1912)
French mathematician and theoretical physicist

This is (still!) an incomplete draft. Please send corrections and comments to ryanjooruian18@gmail.com, or pull-request at https://github.com/Ryanjoo18/fons.

Typeset using LaTeX.

Last updated August 28, 2024.

About the Author

Ryan Joo Rui An is a dedicated and passionate high school student currently engaged in A Level studies in Singapore. With a strong foundation in mathematics, He has spent over 11 years honing his skills in various mathematics competitions. His journey began at an early age, where he developed a fascination with numbers while doing mental arithmetic. This early interest quickly blossomed into a deep commitment to mathematics, leading him to participate in numerous mathematics olympiads competitions.

The author's (not many) mathematics credentials include:

- Singapore Mathematics Olympiad 2022 – 2024: 3 Silver awards
- Singapore and Asian Schools Math Olympiad 2019 – 2023: 6 Gold awards, top in Singapore in 2022 – 2023, top in Malaysia in 2024
- Singapore International Mathematical and Computational Challenge 2024: Merit award
- Australian Mathematics Competition 2019 – 2023: 2 Prize awards, 2 High Distinction awards, best in school in 2023
- High School Mathematical Contest in Modeling 2023: Honourable mention
- Hua Lo Geng Secondary School Mathematics Competition 2019: 2nd place
- Chen Jingrun's Cup Secondary School Mathematics Competition 2019: 1st place

Outside of mathematics, the author has a keen interest in playing chess and programming.

This book is a culmination of the author's years of experience, dedication, and love for mathematics while he studies mathematics at the undergraduate level.

Preface

This book is divided into the following sections.

Part I covers ***preliminary topics***, which are crucial stepping stones for subsequent topics. This includes logic and methods of proofs in Chapter 1, and basic set theory in Chapter 2.

Part II covers ***linear algebra***, which follows [Axl15]. Chapter 3 gives an introduction to vector spaces and subspaces. Chapter 4 gives an overview of span, linear independence, bases and dimension. Chapter 5 goes through linear maps, kernel and image, matrices, invertibility and isomorphism, as well as products and quotients of vector spaces.

Part III covers ***abstract algebra***, which follows [DF04]. Chapter 6 covers group theory.

Part IV covers ***real analysis***, which follows [Rud53; Apo57].

Part V covers ***complex analysis***, which follows [Ahl79; Lan99]. Complex analysis can be viewed as an extension of real analysis, with various overlapping concepts.

Part VI covers ***topology***, which follows [Mun18].

The chapters in this book are structured in the following typical manner. Each chapter begins with a **theoretical portion**, which starts off with a couple of definitions, followed by theorems, lemmas and propositions built upon the definitions. Each chapter is ended by some **exercises** accompanied by solutions to them.

The reader is not assumed to have any mathematical prerequisites, although some experience with proofs may be helpful.

Problem Solving

Mathematics is about problem solving. In [Pól45], George Pólya outlined the following problem solving cycle.

1. **Understand the problem**

   Ask yourself the following questions:

   - Do you understand all the words used in stating the problem?
   - Is it possible to satisfy the condition? Is the condition sufficient to determine the unknown? Or is it insufficient? Or redundant? Or contradictory?
   - What are you asked to find or show? Can you restate the problem in your own words?
   - Draw a figure. Introduce suitable notation.
   - Is there enough information to enable you to find a solution?

2. **Devise a plan**

   A partial list of heuristics – good rules of thumb to solve problems – is included:

   - Guess and check
   - Look for a pattern
   - Make an orderly list
   - Draw a picture
   - Eliminate possibilities
   - Solve a simpler problem
   - Use symmetry
   - Use a model
   - Consider special cases
   - Work backwards
   - Use direct reasoning
   - Use a formula
   - Solve an equation
   - Be ingenious

3. **Execute the plan**

   This step is usually easier than devising the plan. In general, all you need is care and patience, given that you have the necessary skills. Persist with the plan that you have chosen. If it continues not to work discard it and choose another. Don't be misled, this is how mathematics is done, even by professionals.

- Carrying out your plan of the solution, check each step. Can you see clearly that the step is correct? Can you prove that it is correct?

4. **Check and expand**

   Pólya mentions that much can be gained by taking the time to reflect and look back at what you have done, what worked, and what didn't. Doing this will enable you to predict what strategy to use to solve future problems.

   Look back reviewing and checking your results. Ask yourself the following questions:

   - Can you check the result? Can you check the argument?
   - Can you derive the solution differently? Can you see it at a glance?
   - Can you use the result, or the method, for some other problem?

Building on Pólya's problem solving strategy, Schoenfeld [Sch92] came up with the following framework for problem solving, consisting of four components:

1. **Cognitive resources**: the body of facts and procedures at one's disposal.
2. **Heuristics**: 'rules of thumb' for making progress in difficult situations.
3. **Control**: having to do with the efficiency with which individuals utilise the knowledge at their disposal. Sometimes, this is referred to as metacognition, which can be roughly translated as 'thinking about one's own thinking'.
   (a) These are questions to ask oneself to monitor one's thinking.
      - What (exactly) am I doing? [Describe it precisely.] Be clear what I am doing NOW. Why am I doing it? [Tell how it fits into the solution.]
      - Be clear what I am doing in the context of the BIG picture – the solution. Be clear what I am going to do NEXT.

   (b) Stop and reassess your options when you
      - cannot answer the questions satisfactorily [probably you are on the wrong track]; OR
      - are stuck in what you are doing [the track may not be right or it is right but it is at that moment too difficult for you].

   (c) Decide if you want to
      - carry on with the plan,
      - abandon the plan, OR
      - put on hold and try another plan.
4. **Belief system**: one's perspectives regarding the nature of a discipline and how one goes about working on it.

Reading Skills

A mathematics book is not a storybook; it is imperative that you know how to read a mathematics book properly.

Read 0 Don't read the book, read the Wikipedia article about the subject. Learn about the big questions asked in the subject, and the basics of the theorems that answer them. Often the most important ideas are those that can be stated concisely, so you should be able to remember them once you are engaging the book.

Read 1 Let your eyes jump from definition to lemma to theorem without reading the proofs in between unless something grabs your attention or bothers you. If the book has exercises, see if you can do the first one of each chapter or section as you go.

Read 2 Read the book but this time read the proofs. But don't worry if you don't get all the details. If some logical jump doesn't make complete sense, feel free to ignore it at your discretion as long as you understand the overall flow of reasoning.

Read 3 Read through the lens of a sceptic. Work through all of the proofs with a fine toothed comb, and ask yourself every question you think of. You should never have to ask yourself "why" you are proving what you are proving at this point, but you have a chance to get the details down.

Study Skills

The Faculty of Mathematics of the University of Cambridge has produced a leaflet called "Study Skills in Mathematics". The Faculty also has guidance notes intended to help students prepare for exams.

Similarly, the Mathematical Institute of the University of Oxford has a study guide and thoughts on preparing for exams.

Contents

Part I

Preliminaries

1 Mathematical Reasoning and Logic

§1.1 *Mathematical Terminology*

It is useful to be familiar with the following terminology.

- A ***definition*** is a precise and unambiguous description of the meaning of a mathematical term. It characterises the meaning of a word by giving all the properties and only those properties that must be true.
- A ***theorem*** is a true mathematical statement that can be proven mathematically. In a mathematical paper, the term theorem is often reserved for the most important results.
- A ***lemma*** is a minor result whose sole purpose is to help in proving a theorem. It is a stepping stone on the path to proving a theorem. Very occasionally lemmas can take on a life of their own.
- A ***corollary*** is a result in which the (usually short) proof relies heavily on a given theorem.
- A ***proposition*** is a proven and often interesting result, but generally less important than a theorem.
- A ***conjecture*** is a statement that is unproved, but is believed to be true.
- An ***axiom*** is a statement that is assumed to be true without proof. These are the basic building blocks from which all theorems are proven.
- An ***identity*** is a mathematical expression giving the equality of two (often variable) quantities.
- A ***paradox*** is a statement that can be shown, using a given set of axioms and definitions, to be both true and false. Paradoxes are often used to show the inconsistencies in a flawed theory.

A ***proof*** is a sequence of true statements, without logical gaps, that is a logical argument establishing some conclusion.

§1.2 *Zeroth-order Logic*

A ***proposition*** is a sentence which has exactly one truth value, i.e. it is either true or false, but not both and not neither. A proposition is denoted by uppercase letters such as P and Q. If the proposition P depends on a variable x, it is sometimes helpful to denote it by $P(x)$.

We can do some algebra on propositions, which include

(i) ***equivalence***, denoted by $P \iff Q$, which means P and Q are logically equivalent statements;

(ii) ***conjunction***, denoted by $P \wedge Q$, which means "P and Q";

(iii) ***disjunction***, denoted by $P \vee Q$, which means "P or Q";

(iv) ***negation***, denoted by $\neg P$, which means "not P".

Here are some useful properties when handling logical statements. You can easily prove all of them using truth tables.

Proposition 1.1 (Double negation law)**.**

$$P \iff \neg(\neg P)$$

Proposition 1.2 (Commutative property)**.**

$$P \wedge Q \iff Q \wedge P, \quad P \vee Q \iff Q \vee P$$

Proposition 1.3 (Associative property for conjunction)**.**

$$(P \wedge Q) \wedge R \iff P \wedge (Q \wedge R)$$

Proposition 1.4 (Associative property for disjunction)**.**

$$(P \vee Q) \vee R \iff P \vee (Q \vee R)$$

Proposition 1.5 (Distributive property for conjunction across disjunction)**.**

$$P \wedge (Q \vee R) \iff (P \wedge Q) \vee (P \wedge Q)$$

Proposition 1.6 (Distributive property for disjunction across conjunction)**.**

$$P \vee (Q \wedge R) \iff (P \vee Q) \wedge (P \vee R)$$

Proposition 1.7 (De Morgan's laws)**.**

$$\neg(P \vee Q) \iff (\neg P \wedge \neg Q)$$
$$\neg(P \wedge Q) \iff (\neg P \vee \neg Q)$$

Exercise 1
Negate the statement $1 < x < 2$.

Solution.

$$\begin{aligned} \neg(1 < x < 2) &\iff \neg[(x > 1) \wedge (x < 2)] \\ &\iff \neg(x > 1) \vee \neg(x < 2) \\ &\iff (x \leq 1) \vee (x \geq 2) \end{aligned}$$

where the last step follows from the trichotomy axiom of real numbers. □

Exercise 2
Negate the statement

$$n = 2 \quad \text{or} \quad n \text{ is odd.}$$

Solution.

$$\begin{aligned} \neg[(n = 2) \vee (n \text{ is odd})] &\iff \neg(n = 2) \wedge \neg(n \text{ is odd}) \\ &\iff (n \neq 2) \wedge (n \text{ is even}) \end{aligned}$$

where we have used the fact that every integer is either even or odd, but not both. □

§1.2.1 *If, only if*

Implication is denoted by $P \implies Q$, which means "P implies Q", i.e. if P holds then Q also holds. It is equivalent to saying "If P then Q". The only case when $P \implies Q$ is false is when the hypothesis P is true and the conclusion Q is false.

$P \implies Q$ is known as a ***conditional statement***. P is known as the ***hypothesis***, Q is known as the ***conclusion***.

Statements of this form are probably the most common, although they may sometimes appear quite differently. The following all mean the same thing:

(i) if P then Q;

(ii) P implies Q;

(iii) P only if Q;

(iv) P is a sufficient condition for Q;

(v) Q is a necessary condition for P.

The ***converse*** of $P \implies Q$ is given by $Q \implies P$; both are not logically equivalent.

The ***inverse*** of $P \implies Q$ is given by $\neg P \implies \neg Q$, i.e. the hypothesis and conclusion of the statement are both negated.

The ***contrapositive*** of $P \implies Q$ is given by $\neg Q \implies \neg P$; both are logically equivalent.

To prove $P \implies Q$, start by assuming that P holds and try to deduce through some logical steps that Q holds too. Alternatively, start by assuming that Q does not hold and show that P does not hold (that is, we prove the contrapositive).

§1.2.2 *If and only if, iff*

Bidirectional implication is denoted by $P \iff Q$, which means both $P \implies Q$ and $Q \implies P$. We can read this as "P if and only if Q". The letters "iff" are also commonly used to stand for "if and only if".

$P \iff Q$ is true exactly when P and Q have the same truth value.

$P \iff Q$ is known as a ***biconditional statement***.

These statements are usually best thought of separately as "if" and "only if" statements.

To prove $P \iff Q$, prove the statement in both directions, i.e. prove both $P \implies Q$ and $Q \implies P$. Remember to make very clear, both to yourself and in your written proof, which direction you are doing.

§1.3 *First-order Logic*

The ***universal quantifier*** is denoted by $\forall$, which means "for all" or "for every". An universal statement has the form $\forall x \in X, P(x)$.

The ***existential quantifier*** is denoted by $\exists$, which means "there exists". An existential statement has the form $\exists x \in X, P(x)$, where X is known as the ***domain***.

These are versions of De Morgan's laws for quantifiers:

$$\neg \forall x \in X, P(x) \iff \exists x \in X, \neg P(x)$$

$$\neg \exists x \in X, P(x) \iff \forall x \in X, \neg P(x)$$

Exercise 3

Negate the statement

$$\text{for all real numbers } x, \text{ if } x > 2, \text{ then } x^2 > 4$$

Solution. In logical notation, this statement is $(\forall x \in \mathbf{R})[x > 2 \implies x^2 > 4]$.

$$\begin{aligned}\neg\{(\forall x \in \mathbf{R})[x > 2 \implies x^2 > 4]\} &\iff (\exists x \in \mathbf{R})\neg[x > 2 \implies x^2 > 4] \\ &\iff (\exists x \in \mathbf{R})\neg[(x \leq 2) \vee (x^2 > 4)] \\ &\iff (\exists x \in \mathbf{R})[(x > 2) \wedge (x^2 \leq 4)]\end{aligned}$$

□

Exercise 4

Negate surjectivity.

Solution. If $f : X \to Y$ is not surjective, then it means that there exists $y \in Y$ not in the image of X, i.e. for all x in X we have $f(x) \neq y$.

$$\begin{aligned}\neg\forall y \in Y, \exists x \in X, f(x) = y &\iff \exists y \in Y, \neg(\exists x \in X, f(x) = y) \\ &\iff \exists y \in Y, \forall x \in X, \neg(f(x) = y) \\ &\iff \exists y \in Y, \forall x \in X, f(x) \neq y\end{aligned}$$

□

To prove a statement of the form $\forall x \in X$ s.t. $P(x)$, start the proof with "Let $x \in X$." or "Suppose $x \in X$ is given." to address the quantifier with an arbitrary x; provided no other assumptions about x are made during the course of proving $P(x)$, this will prove the statement for all $x \in X$.

To prove a statement of the form $\exists x \in X$ s.t. $P(x)$, there is not such a clear steer about how to continue: you may need to show the existence of an x with the right properties; you may need to demonstrate logically that such an x must exist because of some earlier assumption, or it may be that you can show constructively how to find one; or you may be able to prove by contradiction, supposing that there is no such x and consequently arriving at some inconsistency.

Remark. Read from left to right, and as new elements or statements are introduced they are allowed to depend on previously introduced elements but cannot depend on things that are yet to be mentioned.

Remark. To avoid confusion, it is a good idea to keep to the convention that the quantifiers come first, before any statement to which they relate.

§1.4 *Proofs*

A ***direct proof*** of $P \implies Q$ is a series of valid arguments that start with the hypothesis P and end with the conclusion Q. It may be that we can start from P and work directly to Q, or it may be that we make use of P along the way.

A ***proof by contrapositive*** of $P \implies Q$ is to prove instead $\neg Q \implies \neg P$.

A ***disproof by counterexample*** is to providing a counterexample in order to refute or disprove a conjecture. The counterexample must make the hypothesis a true statement, and the conclusion a false statement. In seeking counterexamples, it is a good idea to keep the cases you consider simple, rather than searching randomly. It is often helpful to consider "extreme" cases; for example, something is zero, a set is empty, or a function is constant.

A ***proof by cases*** is to first dividing the situation into cases which exhaust all the possibilities, and then show that the statement follows in all cases.

§1.4.1 *Proof by contradiction*

A ***proof by contradiction*** of P involves first supposing P is false, i.e. $\neg P$; to prove $P \implies Q$ by contradiction, suppose that Q is false, i.e. $P \wedge \neg Q$. Then show through some logical reasoning that this leads to a contradiction or inconsistency. We may arrive at something that contradicts the hypothesis P, or something that contradicts the initial supposition that Q is not true, or we may arrive at something that we know to be universally false.

Exercise 5 (Irrationality of $\sqrt{2}$)
Prove that $\sqrt{2}$ is irrational.

Proof. We prove by contradiction. Suppose otherwise, that $\sqrt{2}$ is rational. Then $\sqrt{2} = \dfrac{a}{b}$ for some $a, b \in \mathbf{Z}, b \neq 0$, a, b coprime.

Squaring both sides gives

$$a^2 = 2b^2.$$

Since RHS is even, LHS must also be even. Hence it follows that a is even. Let $a = 2k$ where $k \in \mathbf{Z}$. Substituting $a = 2k$ into the above equation and simplifying it gives us

$$b^2 = 2k^2.$$

This means that b^2 is even, from which follows again that b is even.

This contradicts the assumption that a, b coprime. Hence proven. □

Exercise 6 (Euclid's theorem)
There are infinitely many prime numbers.

Proof. Suppose otherwise, that only finitely many prime numbers exist. List them as $p_1, \ldots, p_n$. The number $N = p_1 p_2 \cdots p_n + 1$ is divisible by a prime p, yet is coprime to $p_1, \ldots, p_n$. Therefore, p does not belong to our list of all prime numbers, a contradiction. Hence the initial assumption was false, proving that there are infinitely many primes. □

To ***prove uniqueness***, we can either assume $\exists x, y \in S$ such that $P(x) \wedge P(y)$ is true and show $x = y$, or argue by assuming that $\exists x, y \in S$ are distinct such that $P(x) \wedge P(y)$, then derive a contradiction. $\exists!$ denotes "there exists a unique". To prove uniqueness and existence, we also need to show that $\exists x \in S$ s.t. $P(x)$ is true.

§1.4.2 *Proof of existence*

To prove existential statements, we can adopt two approaches:

1. Constructive proof (direct proof):

   To prove statements of the form $\exists x \in X$ s.t. $P(x)$, find or construct ***a specific example*** for x. To prove statements of the form $\forall y \in Y$, $\exists x \in X$ s.t. $P(x, y)$, construct example for x ***in terms of*** y (since x is dependent on y).

   In both cases, you have to justify that your example x

   (a) belongs to the domain X, and

   (b) satisfies the condition P.

2. Non-constructive proof (indirect proof):

   Use when specific examples are not easy or not possible to find or construct. Make arguments why such objects have to exist. May need to use proof by contradiction. Use definition, axioms or results that involve existential statements.

Exercise 7

Prove that we can find 100 consecutive positive integers which are all composite numbers.

Proof. We can prove this existential statement via constructive proof.

Our goal is to find integers $n, n+1, n+2, \ldots, n+99$, all of which are composite.

Take $n = 101! + 2$. Then n has a factor of 2 and hence is composite. Similarly, $n + k = 101! + (k + 2)$ has a factor $k + 2$ and hence is composite for $k = 1, 2, \ldots, 99$.

Hence the existential statement is proven. □

Exercise 8

Prove that for all rational numbers p and q with $p < q$, there is a rational number x such that $p < x < q$.

Proof. We prove this by construction. Our goal is to find such a rational x ***in terms of*** p ***and*** q.

We take the average. Let $x = \dfrac{p+q}{2}$ which is a rational number.

Since $p < q$,

$$x = \frac{p+q}{2} < \frac{q+q}{2} = q \implies x < q$$

Similarly,

$$x = \frac{p+q}{2} > \frac{p+p}{2} = p \implies p < x$$

Hence we have shown the existence of rational number x such that $p < x < q$.

Remark. For this type of question, there are two parts to prove: firstly, x satisfies the given statement; secondly, x is within the domain (for this question we do not have to prove x is rational since $\mathbf{Q}$ is closed under addition).

□

Exercise 9

Prove that for all rational numbers p and q with $p < q$, there is an irrational number r such that $p < r < q$.

Proof. We prove this by construction. Similarly, our goal is to find an irrational r in terms of p and q.

Note that we cannot simply take $r = \dfrac{p+q}{2}$; a simple counterexample is the case $p = -1, q = 1$ where $r = 0$ is clearly not irrational.

Since p lies in between p and q, let $r = p + c$ where $0 < c < q - p$. Since $c < q - p$, we have $c = \dfrac{q-p}{k}$ for some $k > 1$; to make c irrational, we take k to be irrational.

Take $r = p + \dfrac{q-p}{\sqrt{2}}$. We need to show r is irrational and $p < r < q$.

Part 1: $p < r < q$

Since $q < p$, $r = p +$ (positive number) $> p$. On the other hand, $\dfrac{q-p}{\sqrt{2}} < q - p$ so $r < p + (q - p) = q$.

Part 2: r is irrational

We prove by contradiction. Suppose r is rational. We have $\sqrt{2} = \dfrac{q-p}{r-p}$. Since p, q, r are all rational (and $r - p \neq 0$), RHS is rational. This implies that LHS is rational, i.e. $\sqrt{2}$ is rational, a contradiction. □

Non-constructive proof

> **Exercise 10**
> Prove that every integer greater than 1 is divisible by a prime.

Proof. If n is prime, then we are done as $n \mid n$.

If n is not prime, then n is composite. So n has a divisor d_1 such that $1 < d_1 < n$. If d_1 is prime then we are done as $d_1 \mid n$. If d_1 is not prime then d_1 is composite, has divisor d_2 such that $1 < d_2 < n$.

If d_2 is prime, then we are done as $d_2 \mid d_1$ and $d_1 \mid n$ imply $d_2 \mid n$. If d_2 is not prime then d_2 is composite, has divisor d_3 such that $1 < d_3 < d_2$.

Continuing in this manner after k times, we will get

$$1 < d_k < d_{k-1} < \cdots < d_2 < d_1 < n$$

where $d_i \mid n$ for all i.

This process must stop after finite steps, as there can only be a finite number of d_i's between 1 and n. On the other hand, the process will stop only if there is a d_i which is a prime.

Hence we conclude that there must be a divisor d_i of n that is prime. □

Remark. This proof is also known as ***proof by infinite descent***, a method which relies on the well-ordering principle of the positive integers.

> **Exercise 11**
> Prove that the equation $x^2 + y^2 = 3z^2$ has no solutions (x, y, z) in integers where $z \neq 0$.

Proof. Suppose we have a solution (x, y, z). Without loss of generality, we may assume that $z > 0$. By the least integer principle, we may also assume that our solution has z minimal. Taking remainders modulo 3, we see that

$$x^2 + y^2 \equiv 0 \pmod 3$$

Recalling that squares may only be congruent to 0 or 1 modulo 3, we conclude that

$$x^2 \equiv y^2 \equiv 0 \implies x \equiv y \equiv 0 \pmod 3$$

Writing $x = 3a$ and $y = 3b$ we obtain

$$9a^2 + 9b^2 = 3z^2 \implies 3(a^2 + b^2) = z^2 \implies 3 \mid z^2 \implies 3 \mid z$$

Now let $z = 3c$ and cancel 3's to obtain

$$a^2 + b^2 = 3c^2.$$

We have therefore constructed another solution $(a, b, c) = \left(\frac{x}{3}, \frac{y}{3}, \frac{z}{3}\right)$ to the original equation. However $0 < c < z$ contradicts the minimality of z. □

§1.4.3 *Proof by mathematical induction*

Induction is an extremely powerful method of proof used throughout mathematics. It deals with infinite families of statements which come in the form of lists. The idea behind induction is in showing how each statement follows from the previous one on the list – all that remains is to kick off this logical chain reaction from some starting point.

Theorem 1.8 (Principle of Mathematical Induction)**.** Let $P(n)$ be a family of statements indexed by $\mathbf{Z}^+$. Suppose that

(i) (**base case**) $P(1)$ is true and

(ii) (**inductive step**) for all $k \in \mathbf{Z}^+$, $P(k) \implies P(k+1)$.

Then $P(n)$ is true for all $n \in \mathbf{Z}^+$.

Using logic notation, this is written as

$$\{P(1) \wedge (\forall n \in \mathbf{Z}^+)[P(k) \implies P(k+1)]\} \implies (\forall n \in \mathbf{Z}^+)P(n)$$

Remark. Induction is often visualised like toppling dominoes. The inductive step (ii) corresponds to placing each domino sufficiently close that it will be hit when the previous one falls over, and base case (i) corresponds to knocking over the first one.

$$P(1) \implies P(2) \implies \cdots \implies P(k) \implies P(k+1) \implies \cdots$$

To practise, we go through a classic proof using induction.

Exercise 12

Prove that for any $n \in \mathbf{Z}^+$,

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

Proof. Let $P(n) : \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

Clearly $P(1)$ holds. Now suppose $P(k)$ holds for some $k \in \mathbf{Z}^+$, $k \geq 1$; that is,

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2}.$$

Adding $k+1$ to both sides,

$$\begin{aligned}\sum_{i=1}^{k+1} i &= \frac{k(k+1)}{2} + (k+1) \\ &= \frac{(k+1)(k+2)}{2} \\ &= \frac{(k+1)[(k+1)+1]}{2}\end{aligned}$$

thus $P(k+1)$ is true.

Since $P(1)$ true and $P(k) \implies P(k+1)$ for all $k \in \mathbf{Z}^+$, $k \geq 1$,
$P(n)$ is true for all $n \in \mathbf{Z}^+$. □

A corollary of induction is if the family of statements holds for $n \geq N$, rather than necessarily $n \geq 0$:

Corollary 1.9. Let N be an integer and let $P(n)$ be a family of statements indexed by integers $n \geq N$. Suppose that

(i) (**base case**) $P(N)$ is true and

(ii) (**inductive step**) for all $k \geq N$, $P(k) \implies P(k+1)$.

Then $P(n)$ is true for all $n \geq N$.

Proof. This follows directly by applying the above theorem to the statement $Q(n) = P(n+N)$ for $n \in N$. □

Strong induction

Another variant on induction is when the inductive step relies on some earlier case(s) but not necessarily the immediately previous case. This is known as ***strong induction***:

Theorem 1.10 (Strong Form of Induction)**.** Let $P(n)$ be a family of statements indexed by the natural numbers. Suppose that

(i) (**base case**) $P(1)$ is true and

(ii) (**inductive step**) for all $m \in \mathbf{Z}^+$, if for integers k with $1 \leq k \leq m$, $P(k)$ is true then $P(m+1)$ is true.

Then $P(n)$ is true for all $n \in \mathbf{N}$.

Using logic notation, this is written as

$$\{P(1) \wedge (\forall m \in \mathbf{Z}^+)[P(1) \wedge P(2) \wedge \cdots \wedge P(m) \implies P(m+1)]\} \implies (\forall n \in \mathbf{Z}^+)P(n)$$

Proof. We can this it to an instance of "normal" induction by defining a related family of statements $Q(n)$.

Let $Q(n)$ be the statement "$P(k)$ holds for $k = 0, 1, \ldots, n$". Then the conditions for the strong form are equivalent to

(i) $Q(0)$ holds and

(ii) for any n, if $Q(n)$ is true then $Q(n+1)$ is also true.

It follows by induction that $Q(n)$ holds for all n, and hence $P(n)$ holds for all n. □

The following example illustrates how the strong form of induction can be useful:

Example 1.11 (Fundamental Theorem of Arithmetic)
Every natural number greater than 1 may be expressed as a product of one or more prime numbers.

Proof. Let $P(n)$ be the statement that n may be expressed as a product of prime numbers.

Clearly $P(2)$ holds, since 2 is itself prime.

Let $n \geq 2$ be a natural number and suppose that $P(m)$ holds for all $m < n$.

- If n is prime then it is trivially the product of the single prime number n.
- If n is not prime, then there must exist some $r, s > 1$ such that $n = rs$. By the inductive hypothesis, each of r and s can be written as a product of primes, and therefore $n = rs$ is also a product of primes.

Thus, whether n is prime or not, we have have that $P(n)$ holds. By strong induction, $P(n)$ is true for all natural numbers. That is, every natural number greater than 1 may be expressed as a product of one or more primes. □

Cauchy induction

Theorem 1.12 (Cauchy Induction)**.** Let $P(n)$ be a family of statements indexed by $\mathbf{Z}^+_{\geq 2}$. Suppose that

(i) (**base case**) $P(2)$ is true and

(ii) (**inductive step**) for all $k \in \mathbf{Z}^+$, $P(k) \implies P(2k)$ and $P(k) \implies (k-1)$.

Then $P(n)$ is true for all $n \in \mathbf{Z}^+_{\geq 2}$.

Exercise 13

Using Cauchy Induction, prove the AM–GM Inequality for n variables, which states that for positive reals $a_1, a_2, \ldots a_n$,

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 a_2 \cdots a_n}.$$

Proof. Let $P(n)$ be $\frac{a_1+a_2+\cdots+a_n}{n} \geq \sqrt[n]{a_1 a_2 \cdots a_n}$.

Base case $P(2)$ is true because

$$\frac{a_1 + a_2}{2} \geq \sqrt{a_1 a_2} \iff (a_1 + a_2)^2 \geq 4a_1 a_2 \iff (a_1 - a_2)^2 \geq 0$$

Next we show that $P(n) \implies P(2n)$, i.e. if AM–GM holds for n variables, it also holds for $2n$ variables:

$$\frac{a_1 + a_2 + \cdots + a_{2n}}{2n} = \frac{\frac{a_1+a_2+\cdots+a_n}{n} + \frac{a_{n+1}+a_{n+2}+\cdots+a_{2n}}{n}}{2}$$

$$\frac{\frac{a_1+a_2+\cdots+a_n}{n} + \frac{a_{n+1}+a_{n+2}+\cdots+a_{2n}}{n}}{2} \geq \frac{\sqrt[n]{a_1 a_2 \cdots a_n} + \sqrt[n]{a_{n+1} a_{n+2} \cdots a_{2n}}}{2}$$

$$\frac{\sqrt[n]{a_1 a_2 \cdots a_n} + \sqrt[n]{a_{n+1} a_{n+2} \cdots a_{2n}}}{2} \geq \sqrt{\sqrt[n]{a_1 a_2 \cdots a_n}\sqrt[n]{a_{n+1} a_{n+2} \cdots a_{2n}}}$$

$$\sqrt{\sqrt[n]{a_1 a_2 \cdots a_n}\sqrt[n]{a_{n+1} a_{n+2} \cdots a_{2n}}} = \sqrt[2n]{a_1 a_2 \cdots a_{2n}}$$

The first inequality follows from n-variable AM–GM, which is true by assumption, and the second inequality follows from 2-variable AM–GM, which is proven above.

Finally we show that $P(n) \implies P(n-1)$, i.e. if AM–GM holds for n variables, it also holds for $n-1$ variables. By n-variable AM–GM, $\frac{a_1+a_2+\cdots+a_n}{n} \geq \sqrt[n]{a_1 a_2 \cdots a_n}$ Let $a_n = \frac{a_1+a_2+\cdots+a_{n-1}}{n-1}$ Then we have

$$\frac{a_1 + a_2 + \cdots + a_{n-1} + \frac{a_1+a_2+\cdots+a_{n-1}}{n-1}}{n} = \frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1}$$

So,

$$\frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1} \geq \sqrt[n]{a_1 a_2 \cdots a_{n-1} \cdot \frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1}}$$

$$\Rightarrow \left(\frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1}\right)^n \geq a_1 a_2 \cdots a_{n-1} \cdot \frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1}$$

$$\Rightarrow \left(\frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1}\right)^{n-1} \geq a_1 a_2 \cdots a_{n-1}$$

$$\Rightarrow \frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1} \geq \sqrt[n-1]{a_1 a_2 \cdots a_{n-1}}$$

By Cauchy Induction, this proves the AM–GM inequality for n variables. □

Other variations

Apart from proving $P(n)$ indexed by $\mathbf{Z}^+$, we can also use PMI to prove statements of the form

- $(\forall n \in \mathbf{Z})P(n)$
  Base case: $P(0)$
  Inductive step: $(\forall k \in \mathbf{Z}_{\geq 0})P(k) \implies P(k+1)$ and $(\forall k \in \mathbf{Z}_{\leq 0})P(k) \implies P(k-1)$
  $$\cdots \Longleftarrow P(-n) \Longleftarrow \cdots \Longleftarrow P(-1) \Longleftarrow P(0) \implies P(1) \implies \cdots \implies P(n) \implies \cdots$$
- $(\forall n \in \mathbf{Q})P(n)$
  Base case: $P(0)$
  Inductive step: $P(x) \implies P(-x)$ and $P\left(\frac{a}{b}\right) \implies P\left(\frac{a+1}{b}\right)$ and $P\left(\frac{a}{b}\right) \implies P\left(\frac{a}{b+1}\right)$

A more generalised version

Definition 1.13. A binary relation $\preceq$ on X that satisfies the following conditions is called a ***well-ordering*** on X:

(i) for every $a, b \in X$, $a \preceq b$ or $b \preceq a$,

(ii) every non-empty subset S of X contains a least element wrt $\preceq$.

Theorem 1.14 (Well-ordering principle)**.** Let $(X, \preceq)$ be a well-ordered set, with the least element x_0. Then $P(x)$ holds for all $x \in X$ if the following conditions hold:

(i) (**base case**) $P(x_0)$ holds

(ii) (**inductive step**) $\forall x' \prec x$, $P(x') \implies P(x)$

§1.4.4 *Pigeonhole principle*

Theorem 1.15 (Pigeonhole principle)**.** If $kn+1$ objects ($k \geq 1$ not necessarily finite) are distributed among n boxes, one of the boxes will contain at least $k+1$ objects.

> **Exercise 14** (IMO 1972)
> Prove that every set of 10 two-digit integer numbers has two disjoint subsets with the same sum of elements.

Solution. Let S be the set of 10 numbers. It has $2^{10} - 2 = 1022$ subsets that differ from both S and the empty set. They are the "pigeons".

If $A \subset S$, the sum of elements of A cannot exceed $91 + 92 + \cdots + 99 = 855$. The numbers between 1 and 855, which are all possible sums, are the "holes".

Because the number of "pigeons" exceeds the number of "holes", there will be two "pigeons" in the same "hole". Specifically, there will be two subsets with the same sum of elements. Deleting the common elements, we obtain two disjoint sets with the same sum of elements. □

> **Exercise 15** (Putnam 2006)
> Prove that for every set $X = \{x_1, x_2, \ldots, x_n\}$ of n real numbers, there exists a nonempty subset S of X and an integer m such that
> $$\left| m + \sum_{x \in S} s \right| \leq \frac{1}{n+1}.$$

Solution. Recall that the fractional part of a real number x is $x - \lfloor x \rfloor$. Let us look at the fractional parts of the numbers $x_1, x_1 + x_2, \ldots, x_1 + x_2 + \cdots + x_n$. If any of them is either in the interval $\left[0, \frac{1}{n+1}\right]$ or $\left[\frac{n}{n+1}, 1\right]$, then we are done. If not, we consider these n numbers as the "pigeons" and the $n-1$ intervals $\left[\frac{1}{n+1}, \frac{2}{n+1}\right], \left[\frac{2}{n+1}, \frac{3}{n+1}\right], \ldots, \left[\frac{n-1}{n+1}, \frac{n}{n+1}\right]$ as the "holes". By the pigeonhole principle, two of these sums, say $x_1 + x_2 + \cdots + x_k$ and $x_1 + x_2 + \cdots + x_{k+m}$, belong to the same interval. But then their difference $x_{k+1} + \cdots + x_{k+m}$ lies within a distance of $\frac{1}{n+1}$ of an integer, and we are done. $\square$

Exercises

Problem 1. Use the Unique Factorisation Theorem to prove that, if a positive integer n is not a perfect square, then $\sqrt{n}$ is irrational.

[The Unique Factorisation Theorem states that every integer $n > 1$ has a unique standard factored form, i.e. there is exactly one way to express $n = p_1^{k_1} p_2^{k_2} \cdots p_t^{k_t}$ where $p_1 < p_2 < \cdots < p_t$ are distinct primes and $k_1, k_2, \ldots, k_t$ are some positive integers.]

Proof. Prove by contradiction.

Suppose n is not a perfect square and $\sqrt{n}$ is rational.

Then $\sqrt{n} = \frac{a}{b}$ for some integers a and b. Squaring both sides and clearing denominator gives

$$nb^2 = a^2. \qquad (*)$$

Consider the standard factored forms of n, a and b:

$$n = p_1^{k_1} p_2^{k_2} \cdots p_t^{k_t}$$

$$a = q_1^{e_1} q_2^{e_2} \cdots q_u^{e_u} \implies a^2 = q_1^{2e_1} q_2^{2e_2} \cdots q_u^{2e_u}$$

$$b = r_1^{f_1} r_2^{f_2} \cdots r_v^{f_v} \implies b^2 = r_1^{2f_1} r_2^{2f_2} \cdots r_v^{2f_v}$$

i.e. the powers of primes in the standard factored form of a^2 and b^2 are all even integers.

This means the powers k_i of primes p_i in the standard factored form of n are also even by Unique Factorisation Theorem (UFT):

Note that all p_i appear in the standard factored form of a^2 with even power $2c_i$, because of $(*)$. By UFT, p_i must also appear in the standard factored form of nb^2 with the same even power $2c_i$.

If $p_i \nmid b$, then $k_i = 2c_i$ which is even. If $p_i \mid b$, then p_i will appear in b^2 with even power $2d_i$. So $k_i + 2d_i = 2c_i$, and hence $k_i = 2(c_i - d_i)$, which is again even.

Hence $n = p_1^{k_1} p_2^{k_2} \cdots p_t^{k_t} = \left(p_1^{\frac{k_1}{2}} p_2^{\frac{k_2}{2}} \cdots p_t^{\frac{k_t}{2}} \right)^2$.

Since $\frac{k_i}{2}$ are all integers, $p_1^{\frac{k_1}{2}} p_2^{\frac{k_2}{2}} \cdots p_t^{\frac{k_t}{2}}$ is an integer and n is a perfect square. This contradicts the given hypothesis that n is not a perfect square.

So we conclude that when a positive integer n is not a perfect square, then $\sqrt{n}$ is irrational. □

Problem 2. Prove that for every pair of irrational numbers p and q such that $p < q$, there is an irrational x such that $p < x < q$.

Proof. Consider the average of p and q: $p < \dfrac{p+q}{2} < q$.

If $\dfrac{p+q}{2}$ is irrational, take $x = \dfrac{p+q}{2}$ and we are done.

If $\dfrac{p+q}{2}$ is rational, call it r, take the average of p and r: $p < \dfrac{p+r}{2} < r < q$. Since p is irrational and r is rational, $\dfrac{p+r}{2}$ is irrational. In this case, we take $x = \dfrac{3p+q}{4}$. □

Problem 3. Given n real numbers $a_1, a_2, \ldots, a_n$. Show that there exists an a_i $(1 \le i \le n)$ such that a_i is greater than or equal to the mean (average) value of the n numbers.

Proof. Prove by contradiction.

Let $\bar{a}$ denote the mean value of the n given numbers. Suppose $a_i < \bar{a}$ for all a_i. Then

$$\bar{a} = \frac{a_1 + a_2 + \cdots + a_n}{n} < \frac{\bar{a} + \bar{a} + \cdots + \bar{a}}{n} = \frac{n\bar{a}}{n} = \bar{a}.$$

We derive $\bar{a} < \bar{a}$, which is a contradiction.

Hence there must be some a_i such that $a_i > \bar{a}$. □

Problem 4. Prove that the following statement is false: there is an irrational number a such that for all irrational number b, ab is rational.

Thought process: prove the negation of the statement: for every irrational number a, there is an irrational number b such that ab is irrational.

Proving technique: constructive proof (note that we can consider multiple cases and construct more than one b)

Proof. Given an irrational number a, let us consider $\dfrac{\sqrt{2}}{a}$.

Case 1: $\dfrac{\sqrt{2}}{a}$ is irrational.

Take $b = \dfrac{\sqrt{2}}{a}$. Then $ab = \sqrt{2}$ which is irrational.

Case 2: $\dfrac{\sqrt{2}}{a}$ is rational.

Then the reciprocal $\dfrac{a}{\sqrt{2}}$. Since $\sqrt{6}$ is irrational, the product $\left(\dfrac{a}{\sqrt{2}}\right)\sqrt{6} = a\sqrt{3}$ is irrational. Take $b = \sqrt{3}$, which is irrational. Then $ab = a\sqrt{3}$ which is irrational. □

Problem 5. Prove that there are infinitely many prime numbers that are congruent to 3 modulo 4.

Proof. Prove by contradiction.

Suppose there are only finitely many primes that are congruent to 3 modulo 4. Let $p_1, p_2, \ldots, p_m$ be the list of all the primes that are congruent to 3 modulo 4.

We construct an integer M by $M = (p_1 p_2 \cdots p_m)^2 + 2$.

We have the following observation:

(i) $M \equiv 3 \mod 4$.

(ii) Every p_i divides $M - 2$.

(iii) None of the p_i divides M. [Otherwise, together with (ii), this will imply p_i divides 2, which is impossible.]

(iv) M is not a prime number. [Otherwise, by (i), M is a prime number congruent to 3 modulo 4. But $M \neq p_i$ for all $1 \leq i \leq m$. This contradicts the assumption that $p_1, p_2, \ldots, p_m$ are all the prime numbers congruent to 3 modulo 4.]

From the above discussion, we know that M is a composite number by (iv). So it has a prime factorization $M = q_1 q_2 \cdots q_k$.

Since M is odd, all these prime factors q_j must be odd, and hence q_j must be congruent to either 1 or 3 modulo 4.

By (iii), q_j cannot be any of the p_i. So all q_j must be congruent to 1 modulo 4. Then M, which is the product of q_j, must also be congruent to 1 modulo 4.

This contradicts (i) that M is congruent to 3 modulo 4.

Hence we conclude that there must be infinitely many primes that are congruent to 3 modulo 4. □

Problem 6. Prove that, for any positive integer n, there is a perfect square m^2 (m is an integer) such that $n \leq m^2 \leq 2n$.

Proof. Prove by contradiction.

Suppose otherwise, that $n > m^2$ and $(m+1)^2 > 2n$ so that there is no square between n and $2n$, then

$$(m+1)^2 > 2n > 2m^2.$$

Since we are dealing with integers and the inequalities are strict, we get

$$(m+1)^2 \geq 2m^2 + 2$$

which simplifies to

$$0 \geq m^2 - 2m + 1 = (m-1)^2$$

The only value for which this is possible is $m = 1$, but you can eliminate that easily enough. □

Problem 7. Prove that for every positive integer $n \geq 4$,

$$n! > 2^n.$$

Proof. Let $P(n) : n! > 2^n$

Base case: $P(4)$

LHS: $4! = 4 \times 3 \times 2 \times 1 = 24$, RHS: $2^4 = 16 < 24$

So $P(4)$ is true.

Inductive step: $P(k) \implies P(k+1)$ for all $k \in \mathbf{Z}^+_{\geq 4}$

$$\begin{aligned} k! &> 2^k \\ (k+1)k! &> 2^k(k+1) \\ &> 2^k 2 \quad \text{since from } k \geq 4,\ k+1 \geq 5 > 2 \\ &= 2^{k+1} \end{aligned}$$

hence proven $P(k) \implies P(k+1)$ for integers $k \geq 4$.

By PMI, we have proven $P(n)$ for all integers $n \geq 4$. □

Problem 8. Prove by mathematical induction, for $n \geq 2$,

$$\sqrt[n]{n} < 2 - \frac{1}{n}.$$

Proof. Let $P(n) : \sqrt[n]{n} < 2 - \dfrac{1}{n}$ for $n \geq 2$.

Base case: $P(2)$

When $n = 2$, $\sqrt{2} = 1.41\cdots < 2 - \dfrac{1}{2} = 1.5$ which is true. Hence $P(2)$ is true.

Inductive step: $P(k) \implies P(k+1)$ for all $k \in \mathbf{Z}^+_{\geq 2}$

Assume $P(k)$ is true for $k \geq 2, k \in \mathbf{Z}^+$, i.e.

$$\sqrt[k]{k} < 2 - \frac{1}{k} \implies k < \left(2 - \frac{1}{k}\right)^k$$

We want to prove that $P(k+1)$ is true, i.e.

$$k+1 < \left(2 - \frac{1}{k+1}\right)^{k+1}$$

Since $k > 2$, we have

$$\begin{aligned}\left(2-\frac{1}{k+1}\right)^{k+1} &> \left(2-\frac{1}{k}\right)^{k+1} \quad \because k > 2\\ &= \left(2-\frac{1}{k}\right)^{k}\left(2-\frac{1}{k}\right)\\ &> k\left(2-\frac{1}{k}\right) \quad \text{[by inductive hypothesis]}\\ &= 2k-1 = k+k-1 > k-1 \because k > 2\end{aligned}$$

Hence $P(k+1)$ is true.

Since $P(2)$ is true and $P(k) \implies P(k+1)$, by mathematical induction $P(n)$ is true. □

Problem 9. Prove that for all integers $n \geq 3$,

$$\left(1+\frac{1}{n}\right)^{n} < n$$

Proof. **Base case:** $P(3)$

On the LHS, $\left(1+\frac{1}{3}\right)^{3} = \frac{64}{27} = 2\frac{10}{27} < 3$. Hence $P(3)$ is true.

Inductive step: $P(k) \implies P(k+1)$ for all $k \in \mathbf{Z}^{+}_{\geq 3}$

Our inductive hypothesis is

$$\left(1+\frac{1}{k}\right)^{k} < k$$

Multiplying both sides by $\left(1+\frac{1}{k}\right)$ (to get a $k+1$ in the power),

$$\left(1+\frac{1}{k}\right)^{k}\left(1+\frac{1}{k}\right) = \left(1+\frac{1}{k}\right)^{k+1} < k\left(1+\frac{1}{k}\right) = k+1$$

Since $k < k+1 \iff \frac{1}{k} > \frac{1}{k+1}$,

$$\left(1+\frac{1}{k}\right)^{k+1} > \left(1+\frac{1}{k+1}\right)^{k+1}$$

The rest of the proof follows easily. □

A sequence of integers F_i, where integer $1 \leq i \leq n$, is called the ***Fibonacci sequence*** if and only if it is defined recursively by $F_1 = 1$, $F_2 = 1$, $F_n = F_{n-1} + F_{n-2}$ for $n > 2$.

Problem 10. Let a_i where integer $1 \leq i \leq n$ be a sequence of integers defined recursively by initial conditions $a_1 = 1$, $a_2 = 1$, $a_3 = 3$ and the recurrence relation $a_n = a_{n-1} + a_{n-2} + a_{n-3}$ for $n > 3$.

For all $n \in \mathbf{Z}^{+}$, prove that

$$a_n \leq 2^{n-1}.$$

Proof. Let $P(n) : a_n \leq 2^{n-1}$.

Given the recurrence relation, it could be possible to use $P(k), P(k+1), P(k+2)$ to prove $P(k+3)$ for all $k \in \mathbf{Z}^{+}$.

Base case: $P(1), P(2), P(3)$

$P(1) : a_1 = 1 \leq 2^{1-1} = 1$ is true.

$P(2) : a_2 = 1 \leq 2^{2-1} = 2$ is true.

$P(3) : a_3 = 3 \leq 2^{3-1} = 4$ is true.

Inductive step: $P(k) \wedge P(k+1) \wedge P(k+2) \implies P(k+3)$ for all $k \in \mathbf{Z}^+$

By inductive hypothesis, for $k \in \mathbf{Z}^+$ we have $a_k \leq 2^k, a_{k+1} \leq 2^{k+1}, a_{k+2} \leq 2^{k+2}$.

$$\begin{aligned} a_{k+3} &= a_k + a_{k+1} + a_{k+2} \quad \text{[start from recurrence relation]} \\ &\leq 2^k + 2^{k+1} + 2^{k+2} \quad \text{[use inductive hypothesis]} \\ &= 2^k(1 + 2 + 2^2) \\ &< 2^k(2^3) \quad \text{[approximation, since } 1 + 2 + 2^2 < 2^3\text{]} \\ &= 2^{k+3} \end{aligned}$$

which is precisely $P(k+3) : a_{k+3} \leq 2^{k+3}$. □

Problem 11. For $m, n \in \mathbf{N}$, prove that

$$F_{n+m+1} = F_n F_m + F_{n+1} F_{m+1}.$$

Proof. For $n \in \mathbf{N}$, take $P(n) : F_{n+m+1} = F_n F_m + F_{n+1} F_{m+1}$ for all $m \in \mathbf{N}$ in the cases $k = n$ and $k = n + 1$.

So we are using induction to progress through n and dealing with m simultaneously at each stage.

To verify $P(0)$, we note that

$$F_{m+1} = F_0 F_m + F_1 F_{m+1}$$

and

$$F_{m+2} = F_1 F_m + F_2 F_{m+1}$$

for all m, as $F_0 = 0$ and $F_1 = F_2 = 1$.

For the inductive step we assume $P(n)$, i.e. that for all $m \in \mathbf{N}$,

$$\begin{aligned} F_{n+m+1} &= F_n F_m + F_{n+1} F_{m+1}, \\ F_{n+m+2} &= F_{n+1} F_m + F_{n+2} F_{m+1}. \end{aligned}$$

To prove $P(n+1)$ it remains to show that for all $m \in \mathbf{N}$,

$$F_{n+m+3} = F_{n+2} F_m + F_{n+3} F_{m+1}.$$

From our $P(n)$ assumptions and the definition of the Fibonacci numbers,

$$\begin{aligned} \text{LHS of (5)} &= F_{n+m+3} \\ &= F_{n+m+2} + F_{n+m+1} \\ &= F_n F_m + F_{n+1} F_{m+1} + F_{n+1} F_m + F_{n+2} F_{m+1} \\ &= (F_n + F_{n+1}) Fm + (F_{n+1} + F_{n+2}) F_{m+1} \\ &= F_{n+2} F_m + F_{n+3} F_{m+1} = \text{RHS of (5)}. \end{aligned}$$

□

2 Set Theory

§2.1 *Basics*

A ***set*** S can be loosely defined as a collection of objects. For a set S, we write $x \in S$ to mean that x is an ***element*** of S, and $x \notin S$ if otherwise. A set can be defined in terms of some property $P(x)$ that the elements $x \in S$ satisfy, denoted by the following ***set builder notation***:

$$\{x \in S \mid P(x)\}$$

Some basic sets (of numbers) you should be familiar with:

- $\mathbf{N} = \{0, 1, 2, 3, \dots\}$ denotes the natural numbers (non-negative integers).
- $\mathbf{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ denotes the integers.
- $\mathbf{Q} = \{\frac{p}{q} \mid p, q \in \mathbf{Z}, q \neq 0\}$ denotes the rational numbers.
- $\mathbf{R}$ denotes the real numbers, which can be expressed in terms of decimal expansion.
- $\mathbf{C} = \{x + yi \mid x, y \in \mathbf{R}\}$ denotes the of complex numbers.

The ***empty set*** is the set with no elements, denoted by $\emptyset$.

A is a ***subset*** of B if every element of A is in B, denoted by $A \subseteq B$.

$$A \subseteq B \iff \forall x, x \in A \implies x \in B$$

Proposition 2.1 ($\subseteq$ is transitive)**.** If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

Proof. Let $x \in A$. Since $A \subseteq B$ and $x \in A$, $x \in B$. Since $B \subseteq C$ and $x \in B$, $x \in C$. Hence $A \subseteq C$. □

A is a ***proper subset*** of B if $A \subseteq B$ and $A \neq B$, denoted by $A \subset B$.

Using this definition, we have the relationship

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R}$$

- A and B are ***equal*** if and only if they contain the same elements, denoted by $A = B$.
  To prove that A and B are equal, we simply need to prove that $A \subseteq B$ and $A \subseteq B$.

  Proof. We have

  $$\begin{aligned} A = B &\iff (\forall x)[x \in A \iff x \in B] \\ &\iff (\forall x)[(x \in A \implies x \in B) \wedge (x \in B \implies x \in A)] \\ &\iff \{(\forall x)[x \in A \implies x \in B]\} \wedge (\forall x)[x \in B \implies x \in A)] \\ &\iff (A \subseteq B) \wedge (B \subseteq A) \end{aligned}$$

  □

- Some frequently occurring subsets of the real numbers are known as ***intervals***, which can be visualised as sections of the real line:
  - Open interval
    $$(a, b) = \{x \in \mathbf{R} \mid a < x < b\}$$
  - Closed interval
    $$[a, b] = \{x \in \mathbf{R} \mid a \leq x < b\}$$
  - Half open interval
    $$(a, b] = \{x \in \mathbf{R} \mid a < x \leq b\}$$
- The ***power set*** $\mathcal{P}(A)$ of A is the set of all subsets of A (including the set itself and the empty set).
- An ***ordered pair*** is denoted by (a, b), where the order of the elements matters. Two pairs (a_1, b_1) and (a_2, b_2) are equal if and only if $a_1 = a_2$ and $b_1 = b_2$.

  Similarly, we have ordered triples (a, b, c), quadruples (a, b, c, d) and so on. If there are n elements it is called an n-tuple.
-

The ***Cartesian product*** of sets A and B, denoted by $A \times B$, is the set of all ordered pairs with the first element of the pair coming from A and the second from B:

$$A \times B = \{(a, b) \mid a \in A, b \in B\} \tag{2.1}$$

More generally, we define $A_1 \times A_2 \times \cdots \times A_n$ to be the set of all ordered n-tuples $(a_1, a_2, \ldots, a_n)$, where $a_i \in A_i$ for $1 \leq i \leq n$. If all the A_i are the same, we write the product as A^n.

Example 2.2
$\mathbf{R}^2$ is the Euclidean plane, $\mathbf{R}^3$ is the Euclidean space, and $\mathbf{R}^n$ is the n-dimensional Euclidean space.

$$\begin{aligned} \mathbf{R} \times \mathbf{R} = \mathbf{R}^2 &= \{(x, y) \mid x, y \in \mathbf{R}\} \\ \mathbf{R} \times \mathbf{R} \times \mathbf{R} = \mathbf{R}^3 &= \{(x, y, z) \mid x, y, z \in \mathbf{R}\} \\ \mathbf{R}^n &= \{(x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in \mathbf{R}\} \end{aligned}$$

We now disuss the algebra of sets. Given $A \subset S$ and $B \subset S$.

The ***union*** $A \cup B$ is the set consisting of elements that are in A or B (or both):

$$A \cup B = \{x \in S \mid x \in A \vee x \in B\}$$

The ***intersection*** $A \cap B$ is the set consisting of elements that are in both A and B:

$$A \cap B = \{x \in S \mid x \in A \wedge x \in B\}$$

A and B are ***disjoint*** if both sets have no element in common:

$$A \cap B = \emptyset$$

More generally, we can take unions and intersections of arbitrary numbers of sets, even infinitely many. If we have a family of subsets $\{A_i \mid i \in I\}$, where I is an ***indexing set***, we write

$$\bigcup_{i \in I} A_i = \{x \mid \exists i \in I\, (x \in A_i)\}$$

and

$$\bigcap_{i \in I} A_i = \{x \mid \forall i \in I\, (x \in A_i)\}$$

The ***complement*** of A, denoted by A^c, is the set containing elements that are not in A:

$$A^c = \{x \in S \mid x \notin A\}$$

The ***set difference***, or complement of B in A, denoted by $A \setminus B$, is the subset consisting of those elements that are in A and not in B:

$$A \setminus B = \{x \in A \mid x \notin B\}$$

Note that $A \setminus B = A \cap B^c$.

Proposition 2.3 (Double Inclusion)**.** Let $A \subset S$ and $B \subset S$. Then

$$A = B \iff A \subseteq B \text{ and } B \subseteq A \tag{2.2}$$

Proof.

($\Longrightarrow$) If $A = B$, then every element in A is an element in B, so certainly $A \subseteq B$, and similarly $B \subseteq A$.

($\Longleftarrow$) Suppose $A \subseteq B$, and $B \subseteq A$. Then for every element $x \in S$, if $x \in A$ then $A \subseteq B$ implies that $x \in B$, and if $x \notin A$ then $B \subseteq A$ means $x \notin B$. So $x \in A$ if and only if $x \in B$, and therefore $A = B$. □

Proposition 2.4 (Distributive Laws)**.** Let $A \subset S$, $B \subset S$ and $C \subset S$. Then

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \tag{2.3}$$

$$(A \cap B) \cap C = (A \cup C) \cap (B \cup C) \tag{2.4}$$

Proof. For the first one, suppose x is in the LHS, that is $x \in A \cup (B \cap C)$. This means that $x \in A$ or $x \in B \cap C$ (or both). Thus either $x \in A$ or x is in both B and C (or x is in all three sets). If $x \in A$ then $x \in A \cup B$ and $x \in A \cup C$, and therefore x is in the RHS. If x is in both B and C then similarly x is in both $A \cup B$ and $A \cup C$. Thus every element of the LHS is in the RHS, which means we have shown $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$.

Conversely suppose that $x \in (A \cup B) \cap (A \cup C)$. Then x is in both $A \cup B$ and $A \cup C$. Thus either $x \in A$ or, if $x \notin A$, then $x \in B$ and $x \in C$. Thus $x \in A \cup (B \cap C)$. Hence $(A \cup B) \cap (A \cup C) \subseteq A \cup (B \cap C)$.

By double inclusion, $(A \cup B) \cap (A \cup C) = A \cup (B \cap C)$.

The proof of the second one follows similarly and is left as an exercise. □

Proposition 2.5 (De Morgan's Laws)**.** Let $A \subset S$ and $B \subset S$. Then

$$(A \cup B)^c = A^c \cap B^c \tag{2.5}$$

$$(A \cap B)^c = A^c \cup B^c \tag{2.6}$$

Proof. For the first one, suppose $x \in (A \cup B)^c$. Then x is not in either A or B. Thus $x \in A^c$ and $x \in B^c$, and therefore $x \in A^c \cap B^c$.

Conversely, suppose $x \in A^c \cap B^c$. Then $x \notin A$ and $x \notin B$, so x is in neither A nor B, and therefore $x \in (A \cup B)^c$.

By double inclusion, the first result holds. The second result follows similarly and is left as an exercise. □

De Morgan's laws extend naturally to any number of sets, so if $\{A_i \mid i \in I\}$ is a family of subsets of S, then

$$\left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c \quad \text{and} \quad \left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c$$

Exercise 16

Prove the following:

1. $\left(\bigcup_{i \in I} A_i\right) \cup B = \bigcup_{i \in I} (A_i \cup B)$
2. $\left(\bigcap_{i \in I} A_i\right) \cup B = \bigcap_{i \in I} (A_i \cup B)$
3. $\left(\bigcup_{i \in I} A_i\right) \cup \left(\bigcup_{j \in J} B_j\right) = \bigcup_{(i,j) \in I \times J} (A_i \cup B_j)$

4. $\left(\bigcap_{i\in I} A_i\right) \cup \left(\bigcap_{j\in J} B_j\right) = \bigcap_{(i,j)\in I\times J}(A_i \cup B_j)$

Exercise 17

Let $S \subset A \times B$. Express the set A_S of all elements of A which appear as the first entry in at least one of the elements in S.

(A_S here may be called the projection of S onto A.)

§2.2 *Relations*

Definition 2.6 (Relation)**.** R is a ***relation*** between A and B if and only if $R \subseteq A \times B$.

$a \in A$ and $b \in B$ are ***related*** if $(a, b) \in R$, denoted aRb.

Remark. A relation is a set of ordered pairs.

Visually speaking, a relation is uniquely determined by a simple bipartite graph over A and B. On the bipartite graph, this is usually represented by an edge between a and b.

Definition 2.7 (Binary relation)**.** A ***binary relation*** in A is a relation between A and itself, i.e. $R \subseteq A \times A$.

A and B are the ***domain*** and ***range*** of R respectively, denoted by $\operatorname{dom} R$ and $\operatorname{ran} R$ respectively, if and only if $A \times B$ is the smallest Cartesian product of which R is a subset.

Example 2.8
Given $R = \{(1, a), (1, b), (2, b), (3, b)\}$, then $\operatorname{dom} R = \{1, 2, 3\}$ and $\operatorname{ran} R = \{a, b\}$.

In many cases we do not actually use R to write the relation because there is some other conventional notation:

Example 2.9

- The "less than or equal to" relation $\leq$ on the set of real numbers is $\{(x, y) \in \mathbf{R}^2 \mid x \leq y\}$. We write $x \leq y$ if (x, y) is in this set.
- The "divides" relation $\mid$ on $\mathbf{N}$ is $\{(m, n) \in \mathbf{N}^2 : m \text{ divides } n\}$. We write $m \mid n$ if (m, n) is in this set.
- For a set S, the "subset" relation $\subseteq$ on $\mathcal{P}(S)$ is $\{(A, B) \in \mathcal{P}(S)^2 \mid A \subseteq B\}$. We write $A \subseteq B$ if (A, B) is in this set.

We now discuss some properties of relations. Let A be a set, R a relation on A, $x, y, z \in A$. We say that

- R is ***reflexive*** if xRx for all $x \in A$;
- R is ***symmetric*** if $xRy \implies yRx$;
- R is ***anti-symmetric*** if xRy and $yRx \implies x = y$;
- R is ***transitive*** if xRy and $yRz \implies xRz$.

Example 2.10 (Less than or equal to)
The relation $\leq$ on R is reflexive, anti-symmetric, and transitive, but not symmetric.

Definition 2.11. A ***partial order*** on a non-empty set A is a relation $\leq$ on A satisfying

(i) reflexivity,

(ii) anti-symmetry,

(iii) transitivity.

A ***total ordering*** on A is a partial ordering on A such that if for every $x, y \in A$, either xRy or yRx (or both).

A ***well ordering*** on A is a total ordering on A such that every non-empty subset of A has a minimal element, i.e. for each non-empty $B \subseteq A$ there exists some $s \in B$ such that $s \leq b$ for all $b \in B$.

Example 2.12 (Less than)
The relation $<$ on R is not reflexive, symmetric, or anti-symmetric, but it is transitive.

Example 2.13 (Not equal to)
The relation $\neq$ on R is not reflexive, anti-symmetric or transitive, but it is symmetric.

Exercise 18
Congruence modulo n Let $n \geq 2$ be an integer, and define R on $\mathbf{Z}$ by saying aRb if and only if $a - b$ is a multiple of n. Prove that R is reflexive, symmetric and transitive.

Proof.

(i) Reflexivity: For any $a \in \mathbf{Z}$ we have aRa as 0 is a multiple of n.

(ii) Symmetry: If aRb then $a - b = kn$ for some integer k. So $b - a = -kn$, and hence bRa.

(iii) Transitivity: If aRb and bRc then $a - b = kn$ and $b - c = ln$ for integers k, l. So then $a - c = (a - b) + (b - c) = (k + l)n$, and hence aRc.

□

Definition 2.14. Let the non-empty set A be partially ordered by $\leq$.

- A subset $B \subseteq A$ is called a ***chain*** if for all $x, y \in B$, either $x \leq y$ or $y \leq x$.
- An ***upper bound*** for a subset $B \subseteq A$ is an element $u \in A$ such that $b \leq u$ for all $b \in B$.
- A ***maximal element*** of A is an element $m \in A$ such that $m \leq x$ for any $x \in A$, then $m = x$.

Lemma 2.15 (Zorn's lemma)**.** If A is a non-empty partially ordered set in which every chain has an upper bound, then A has a maximal element.

It is a non-trivial result that Zorn's lemma is independent of the usual (Zermelo–Fraenkel) axioms of set theory in the sense that if the axioms of set theory are consistent, then so are these axioms together with Zorn's lemma; and if the axioms of set theory are consistent, then so are these axioms together with the negation of Zorn's lemma.

Lemma 2.16 (Axiom of choice)**.** The Cartesian product of any non-empty collection of non-empty sets is non-empty. In other words, if I is any non-empty (indexing) set and A_i is a non-empty set for all $i \in I$, then there exists a choice function from I to $\bigcup_{i \in I} A_i$.

Lemma 2.17 (Well-ordering principle)**.** Every non-empty set A has a well-ordering.

Theorem 2.18. Assuming the usual (Zermelo–Fraenkel) axioms of set theory, the following are equivalent:

(i) Zorn's lemma

(ii) Axiom of choice

(iii) Well-ordering principle

Proof. This follows from elementary set theory. We refer the reader to *Real and Abstract Analysis* by Hewitt and Stromberg, Section 3. □

One important type of relation is an equivalence relation. An equivalence relation is a way of saying two objects are, in some particular sense, "the same".

Definition 2.19 (Equivalence relation)**.** A binary relation R on A is an ***equivalence relation*** if it is reflexive, symmetric and transitive.

Notation. We use the symbol $\sim$ to denote the equivalence relation R in $A \times A$: whenever $(a, b) \in R$ we denote $a \sim b$.

An equivalence relation provides a way of grouping together elements that can be viewed as being the same:

Definition 2.20 (Equivalence class)**.** Given an equivalence relation $\sim$ on a set A, and given $x \in A$, the ***equivalence class*** of x is

$$[x] := \{y \in A \mid y \sim x\}.$$

Example 2.21 (Congruence modulo n)
For the equivalence relation of congruence modulo n, the equivalence class of 1 is the set $1 = \{\dots, -n+1, 1, n+1, 2n+1, \dots\}$; that is, all the integers that are congruent to 1 modulo n.

Properties of equivalence classes:

- Every two equivalence classes are disjoint
- The union of equivalence classes form the entire set

You can translate these properties into the point of view from the elements: Every element belongs to one and only one equivalence class.

- No element belongs to two distinct classes
- All elements belong to an equivalence class

Definition 2.22. The ***set of equivalence classes*** (quotient sets) are the set of all equivalence classes, denoted by $A/\sim$.

Grouping the elements of a set into equivalence classes provides a partition of the set, which we define as follows:

Definition 2.23 (Partition)**.** A ***partition*** of a set A is a collection of subsets $\{A_i \subseteq A \mid i \in I\}$, where I is an indexing set, with the property that

(i) $A_i \neq \emptyset$ for all $i \in I$ (all the subsets are non-empty)

(ii) $\bigcup_{i \in I} Ai = A$ (every member of A lies in one of the subsets)

(iii) $A_i \cap A_j = \emptyset$ for every $i \neq j$ (the subsets are disjoint)

The subsets are called the ***parts*** of the partition.

Example 2.24 (Odd and even natural numbers)
$\{\{n \in \mathbf{N} \mid n \text{ is divisible by } 2\}, \{n \in \mathbf{N} \mid n+1 \text{ is divisible by } 2\}\}$ forms a partition of the natural numbers, into evens and odds.

§2.3 *Functions*

Definition 2.25 (Function)**.** A ***function*** $f : X \to Y$ is a mapping of every element of X to some element of Y.

X and Y are known as the ***domain*** and ***codomain*** of f respectively.

Remark. The definition requires that a unique element of the codomain is assigned for every element of the domain. For example, for a function $f : \mathbf{R} \to \mathbf{R}$, the assignment $f(x) = \frac{1}{x}$ is not sufficient as it fails at $x = 0$. Similarly, $f(x) = y$ where $y^2 = x$ fails because $f(x)$ is undefined for $x < 0$, and for $x > 0$ it does not return a unique value; in such cases, we say the the function is ***ill-defined***. We are interested in the opposite; functions that are ***well-defined***.

Definition 2.26. Given a function $f : X \to Y$, the ***image*** (or range) of f is

$$f(X) = \{f(x) \mid x \in X\} \subseteq Y$$

More generally, given $A \subseteq X$, the image of A under f is

$$f(A) = \{f(x) \mid x \in A\} \subseteq Y$$

Given $B \subseteq Y$, the ***pre-image*** of B under f is

$$f^{-1}(B) = \{x \mid f(x) \in B\} \subseteq X$$

Remark. Beware the potentially confusing notation: for $x \in X$, $f(x)$ is a single element of Y, but for $A \subseteq X$, $f(A)$ is a set (a subset of Y). Note also that $f^{-1}(B)$ should be read as "the pre-image of B" and not as "f-inverse of B"; the pre-image is defined even if no inverse function exists (in which case f^{-1} on its own has no meaning; we discuss invertibility of a function below).

Exercise 19

Prove the following statements:

(a) $f(A \cup B) = f(A) \cup f(B)$

(b) $f(A_1 \cup \cdots \cup A_n) = f(A_1) \cup \cdots \cup f(A_n)$

(c) $f(\bigcup_{\lambda \in A} A_\lambda) = \bigcup_{\lambda \in A} f(A_\lambda)$

(d) $f(A \cap B) \subset f(A) \cap f(B)$

(e) $f^{-1}(f(A)) \supset A$

(f) $f(f^{-1}(A)) \subset A$

(g) $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$

(h) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$

(i) $f^{-1}(A_1 \cup \cdots \cup A_n) = f^{-1}(A_1) \cup \cdots \cup f^{-1}(A_n)$

(j) $f^{-1}(\bigcup_{\lambda \in A} A_\lambda) = \bigcup_{\lambda \in A} f^{-1}(A_\lambda)$

If a function is defined on some larger domain than we care about, it may be helpful to restrict the domain:

Definition 2.27 (Restriction)**.** Given a function $f : X \to Y$ and a subset $A \subseteq X$, the ***restriction*** of f to A is the map $f|_A : A \to Y$ defined by $f|_A(x) = f(x)$ for all $x \in A$.

The restriction is almost the same function as the original f – just the domain has changed.

Another rather trivial but nevertheless important function is the identity map:

Definition 2.28 (Identity map)**.** Given a set X, the ***identity*** $\mathrm{id}_X : X \to X$ is defined by $\mathrm{id}_X(x) = x$ for all $x \in X$.

Notation. If the domain is unambiguous, the subscript may be removed.

Definition 2.29 (Injectivity)**.** $f : X \to Y$ is ***injective*** if each element of Y has at most one element of X that maps to it.

$$\forall x_1, x_2 \in X,\ f(x_1) = f(x_2) \implies x_1 = x_2$$

Definition 2.30 (Surjectivity)**.** $f : X \to Y$ is ***surjective*** if every element of Y is mapped to at least one element of X.

$$\forall y \in Y,\ \exists x \in X \text{ s.t. } f(x) = y$$

Definition 2.31 (Bijectivity)**.** $f : X \to Y$ is ***bijective*** if it is both injective and surjective: each element of Y is mapped to a unique element of X.

Notation. Given two sets X and Y , we will write $X \sim Y$ to denote the existence of a bijection from X to Y . One easily checks that $\sim$ is transitive, i.e. if $X \sim Y$ and $Y \sim Z$, then $X \sim Z$.

Theorem 2.32 (Cantor–Schroder–Bernstein)**.** If $f : A \to B$ and $g : B \to A$ are both injections, then $A \sim B$.

Proof. □

§2.3.1 *Composition and invertibility*

Definition 2.33 (Composition)**.** Given two functions $f : X \to Y$ and $g : Y \to Z$, the ***composition*** $g \circ f : X \to Z$ is defined by

$$(g \circ f)(x) = g(f(x)) \quad (\forall x \in X)$$

The composition of functions is not commutative. However, composition is associative, as the following results shows:

Proposition 2.34 (Associativity)**.** Let $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$ be three functions. Then

$$f \circ (g \circ h) = (f \circ g) \circ h.$$

Proof. Let $x \in X$. Then, by the definition of composition, we have

$$(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x))) = (f \circ g)(h(x)) = ((f \circ g) \circ h)(x).$$

□

Proposition 2.35 (Composition preserves injectivity)**.** If $f : X \to Y$ is injective and $g : Y \to Z$ is injective, then $g \circ f : X \to Z$ is injective.

Proof. Let $f : X \to Y$ and $g : Y \to Z$ be arbitrary injective functions. We want prove that the function $g \circ f : X \to Z$ is also injective.

To do so, we will prove $\forall x, x' \in X$ that

$$(g \circ f)(x) = (g \circ f)(x') \implies x = x'$$

Suppose that $(g \circ f)(x) = (g \circ f)(x')$. Expanding out the definition of $g \circ f$, this means that $g(f(x)) = g(f(x'))$.

Since g is injective and $g(f(x)) = g(f(x'))$, we know $f(x) = f(x')$.

Similarly, since f is injective and $f(x) = f(x')$, we know that $x = x'$, as required. □

Proposition 2.36. f is injective if and only if for any set Z and any functions $g_1, g_2 : Z \to X$ we have $f \circ g_1 = f \circ g_2 \implies g_1 = g_2$.

Proof. ($\implies$) If f is injective, we ultimately wish to show that $g_1 = g_2$, so in order to do this we consider all possible inputs $z \in Z$, hoping to show that $g_1(z) = g_2(z)$.

But this is quite simple because we are given that $f \circ g_1 = f \circ g_2$ and that f is injective, so

$$f \circ g_1(z) = f \circ g_2(z) \implies g_1(z) = g_2(z)$$

($\impliedby$) We specifically pick $Z = \{1\}$, basically some random one-element set.

Then $\forall x, y \in X$, we define

$$g_1 : Z \to X, g_1(1) = x$$
$$g_2 : Z \to Y, g_2(1) = y$$

Then

$$f(x) = f(y) \implies f(g_1(1)) = f(g_2(1)) \implies g_1(1) = g_2(1) \implies x = y$$

□

Proposition 2.37 (Composition preserves surjectivity)**.** If $f : X \to Y$ is surjective and $g : Y \to Z$ is surjective, then $g \circ f : X \to Z$ is surjective.

Proof. Let $f : X \to Y$ and $g : Y \to Z$ be arbitrary surjective functions. We want to prove that the function $g \circ f : X \to Z$ is subjective.

To do so, we want to prove that for any $z \in Z$, there is some $x \in X$ such that $(g \circ f)(x) = z$. Equivalently, we want to prove that for any $z \in Z$, there is some $x \in X$ such that $g(f(x)) = z$.

Consider any $z \in Z$. Since $g : Y \to Z$ is surjective, there is some $y \in Y$ such that $g(y) = z$. Similarly, since $f : X \to Y$ is surjective, there is some $x \in X$ such that $f(x) = y$. This means that there is some $x \in X$ such that $g(f(x)) = g(y) = z$, as required. □

Proposition 2.38. f is surjective if and only if for any set Z and any functions $g_1, g_2 : Y \to Z$ we have $g_1 \circ f = g_2 \circ f \implies g_1 = g_2$.

Proof.

($\implies$) Suppose that f is surjective. Again, we wish to show that $g_1 = g_2$, so we need to consider every possible input y in Y. Then, since f is injective, we can always pick $x \in X$ such that $f(x) = y$.

Then

$$g_1 \circ f = g_2 \circ f \implies g_1 \circ f(x) = g_2 \circ f(x) \implies g_1(y) = g_2(y)$$

On the other hand, if f is not surjective, then there exists $y \in Y$ such that for all $x \in X$ we have $f(x) \neq y$. We then aim to construct set Z and $g_1, g_2 : Y \to Z$ such that

(i) $g_1(y) \neq g_2(y)$

(ii) $\forall y' \neq y, g_1(y') = g_2(y')$

Because if this is satisfied, then $\forall x \in X$, since $f(x) \neq y$ we have from (ii) that $g_1(f(x)) = g_2(f(x))$; thus $g_1 \circ f = g_2 \circ f$, and yet from (i) we have $g_1 \neq g_2$.

($\impliedby$) We construct $Z = Y \cup \{1, 2\}$ for some random $1, 2 \notin Y$.

Then we define

$$g_1 : Y \to Z, g_1(y) = 1, g_1(y') = y' \qquad\qquad g_2 : Y \to Z, g_2(y) = 2, g_2(y') = y'$$

Then when y is not in the image of f, these two functions will satisfy $g_1 \circ f = g_2 \circ f$ but not $g_1 = g_2$.

So conversely, if for any set Z and any functions $g_i : Y \to Z$ we have $g_1 \circ f = g_2 \circ f \implies g_1 = g_2$, such a value y that is in the codomain but not in the range of f cannot appear, and hence f must be surjective. □

The following proposition addresses the extent to which composition of functions preserves injectivity and surjectivity:

Proposition 2.39. Let $f : X \to Y$ and $g : Y \to Z$ be functions.

(i) If f and g are injective then so is $g \circ f$. Conversely, if $g \circ f$ is injective, then f is injective, but g need not be.

(ii) If f and g are surjective then so is $g \circ f$. Conversely, if $g \circ f$ is surjective, then g is surjective, but f need not be.

Proof. For the first part of (i), suppose $(g \circ f)(x_1) = (g \circ f)(x_2)$ for some $x_1, x_2 \in X$. From the injectivity of g we know that $g(f(x_1)) = g(f(x_2))$ implies $f(x_1) = f(x_2)$, and then from the injectivity of f we know that this implies $x_1 = x_2$. So $g \circ f$ is injective.

For the second part of (i), suppose $f(x_1) = f(x_2)$ for some $x_1, x_2 \in X$. Then applying g gives $g(f(x_1)) = g(f(x_2))$, and by the injectivity of $g \circ f$ this means $x_1 = x_2$. So f is injective. To see that g need not be injective, a counterexample is $X = Z = \{0\}, Y = \mathbf{R}$, with $f(0) = 0$ and $g(y) = 0$ for all $y \in \mathbf{R}$. □

Recalling that id_X is the identity map on X, we can define invertibility:

Definition 2.40 (Invertibility)**.** A function $f : X \to Y$ is ***invertible*** if there exists $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. g is known as the ***inverse*** of f, denoted by $g = f^{-1}$.

Remark. Note that directly from the definition, if f is invertible then f^{-1} is also invertible, and $(f^{-1})^{-1} = f$.

Proposition 2.41 (Uniqueness of inverse)**.** If $f : X \to Y$ is invertible then its inverse is unique.

Proof. Let g_1 and g_2 be two functions for which $g_i \circ f = \mathrm{id}_X$ and $f \circ g_i = \mathrm{id}_Y$. Using the fact that composition is associative, and the definition of the identity maps, we can write

$$g_1 = g_1 \circ \mathrm{id}_Y = g_1 \circ (f \circ g_2) = (g_1 \circ f) \circ g_2 = \mathrm{id}_X \circ g_2 = g_2$$

□

The following result shows how to invert the composition of invertible functions:

Proposition 2.42. Let $f : X \to Y$ and $g : Y \to Z$ be functions. If f and g are invertible, then $g \circ f$ is invertible, and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof. Making repeated use of the fact that function composition is associative, and the definition of the inverses f^{-1} and g^{-1}, we note that

$$\begin{aligned}(f^{-1} \circ g^{-1}) \circ (g \circ f) &= ((f^{-1} \circ g^{-1}) \circ g) \circ f \\ &= (f^{-1} \circ (g^{-1} \circ g)) \circ f \\ &= (f^{-1} \circ \mathrm{id}_Y) \circ f \\ &= f^{-1} \circ f \\ &= \mathrm{id}_X\end{aligned}$$

and similarly,

$$\begin{aligned}(g \circ f) \circ (f^{-1} \circ g^{-1}) &= g \circ (f \circ (f^{-1} \circ g^{-1})) \\ &= g \circ ((f \circ f^{-1}) \circ g^{-1}) \\ &= g \circ (\mathrm{id}_Y \circ g^{-1}) \\ &= g \circ g^{-1} \\ &= \mathrm{id}_Z\end{aligned}$$

which shows that $f^{-1} \circ g^{-1}$ satisfies the properties required to be the inverse of $g \circ f$. □

The following result provides an important and useful criterion for invertibility:

Theorem 2.43. A function $f : X \to Y$ is invertible if and only if it is bijective.

Proof.

($\implies$) Suppose f is invertible, so it has an inverse $f^{-1} : Y \to X$. To show f is injective, suppose that for some $x_1, x_2 \in X$ we have $f(x_1) = f(x_2)$. Then applying f^{-1} to both sides and noting that by definition $f^{-1} \circ f = \mathrm{id}_X$, we see that $x_1 = f^{-1}(f(x_1)) = f^{-1}(f(x_2)) = x_2$. So f is injective. To show that f is surjective, let $y \in Y$, and note that $f^{-1}(y) \in X$ has the property that $f(f^{-1}(y)) = y$. So f is surjective. Therefore f is bijective.

($\impliedby$) Suppose f is bijective, we aim to show that there is a well-defined $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. Since f is surjective, we know that for any $y \in Y$, there is an $x \in X$ such that $f(x) = y$. Furthermore, since f is injective, we know that this x is unique. So for each $y \in Y$ there is a unique $x \in X$ such that $f(x) = y$. This recipe provides a well-defined function $g(y) = x$, for which we have $g(f(x)) = x$ for any $x \in X$ and $f(g(y)) = y$ for any $y \in Y$. So g satisfies the property required to be an inverse of f and therefore f is invertible. □

It is also possible to define left-inverse and right-inverse functions as functions that partially satisfy the definition of the inverse:

Definition 2.44. A function $f : X \to Y$ is ***left invertible*** if there exists a function $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$, and is ***right invertible*** if there exists a function $h : Y \to X$ such that $f \circ h = \mathrm{id}_Y$.

As may be somewht apparent from the previous proof, being left- and right-invertible is equivalent to being injective and surjective, respectively. We leave this as an exercise to show.

§2.3.2 *Monotonicity*

Definition 2.45. $f : [a, b] \to \mathbf{R}$ is called

(1) ***increasing***, if any $a < x_1 \leq x_2 < b$, there is $f(x_1) \leq f(x_2)$;

(2) ***decreasing***, if any $a < x_1 \leq x_2 < b$, there is $f(x_1) \geq f(x_2)$;

f is ***monotonic*** if it is increasing or decreasing.

Suppose $f(x)$ is continuous in $[a, b]$. To locate the roots of $f(x) = 0$:

- If $f(a)$ and $f(b)$ have ***opposite*** signs, i.e. $f(a)f(b) < 0$, then there is an odd number of real roots (counting repeated) in $[a, b]$.

  Furthermore, if f is either strictly increasing or decreasing in $[a, b]$, then $f(x) = 0$ has ***exactly one real root*** in $[a, b]$.
- If $f(a)$ and $f(b)$ have ***same*** signs, i.e. $f(a)f(b) > 0$, then there is an even number of roots (counting repeated) in $[a, b]$.

§2.3.3 *Convexity and concavity*

Definition 2.46. A function f is ***convex*** if for all $x_1, x_2 \in D_f$ and $0 \leq t \leq 1$, we have

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

Note that equality holds when $x_1 = x_2$.

Definition 2.47. A function f is ***strictly convex*** if for all $x_1, x_2 \in D_f$ with $x_1 \neq x_2$ and $0 < t < 1$, we have

$$f(tx_1 + (1-t)x_2) < tf(x_1) + (1-t)f(x_2).$$

Definition 2.48. A function f is ***concave*** if for all $x_1, x_2 \in D_f$ and $0 \leq t \leq 1$, we have

$$f(tx_1 + (1-t)x_2) \geq tf(x_1) + (1-t)f(x_2).$$

Note that equality holds when $x_1 = x_2$.

Definition 2.49. A function f is ***strictly concave*** if for all $x_1, x_2 \in D_f$ with $x_1 \neq x_2$ and $0 < t < 1$, we have

$$f(tx_1 + (1-t)x_2) > tf(x_1) + (1-t)f(x_2).$$

§2.3.4 *Other functions*

Piecewise Functions

A function that has its domain divided into ***separate partitions*** and each partition of the domain given a different formula or rule is known as a ***piecewise funtion***, i.e. a function defined "piece-wise".

Definition 2.50 (Absolute value function)**.**

$$f(x) = |x| = \begin{cases} -x & x < 0, \\ x & x \geq 0. \end{cases}$$

Definition 2.51 (Floor function)**.** The ***floor function*** $f(x) = \lfloor x \rfloor$ is defined as the greatest integer smaller than or equal to x.

For $x \in \mathbf{R}$ and $n \in \mathbf{Z}$,

$$\lfloor x \rfloor = n \iff n \leq x < n + 1.$$

Definition 2.52 (Ceiling function)**.** The ceiling function $f(x) = \lceil x \rceil$ is the direct opposite of the floor function; it maps all real numbers in the domain to the smallest integer not smaller than it.

$$\lceil x \rceil = \begin{cases} \lfloor x \rfloor + 1 & x \notin \mathbf{Z} \\ \lfloor x \rfloor & x \in \mathbf{Z} \end{cases}$$

Exercise 20

Prove that

(a) $\lfloor \sqrt{x} \rfloor = \left\lfloor \sqrt{\lfloor x \rfloor} \right\rfloor$

(b) $\lceil \sqrt{x} \rceil = \left\lceil \sqrt{\lceil x \rceil} \right\rceil$

Solution.

(a)

$$\begin{aligned} & \lfloor \sqrt{x} \rfloor = n \\ \iff & n \le \sqrt{x} < n+1 \quad \text{[by definition of floor function]} \\ \iff & n^2 \le x < (n+1)^2 \quad \text{[square both sides]} \\ \iff & n^2 \le \lfloor x \rfloor \le x < (n+1)^2 \\ \iff & n \le \sqrt{\lfloor x \rfloor} < n+1 \quad \text{[take square root throughout]} \\ \iff & \left\lfloor \sqrt{\lfloor x \rfloor} \right\rfloor = n \quad \text{[by definition of floor function]} \end{aligned}$$

(b)

$$\begin{aligned} & \lceil \sqrt{x} \rceil = n+1 \\ \iff & n < \sqrt{x} \le n+1 \quad \text{[by definition of ceiling function]} \\ \iff & n^2 < x \le (n+1)^2 \quad \text{[square both sides]} \\ \iff & n^2 < x \le \lceil x \rceil \le (n+1)^2 \\ \iff & n < \sqrt{\lceil x \rceil} \le n+1 \quad \text{[take square root throughout]} \\ \iff & \left\lceil \sqrt{\lceil x \rceil} \right\rceil = n+1 \quad \text{[by definition of ceiling function]} \end{aligned}$$

□

Symmetrical Functions

There are special functions with some form of geometric symmetry.

- Even Functions
  f is ***even*** if $f(-x) = f(x)$ for every $x \in D_f$.
  The graph of an even function is symmetric about the y-axis.
- Odd Functions
  f is ***odd*** if $f(-x) = -f(x)$ for every $x \in D_f$.
  The graph of an odd function is symmetric about the origin.
- Periodic Functions
  f is ***periodic*** if $f(x+p) = f(x)$ for every $x \in D_f$, where p is a positive constant. The smallest such p is known as the period.

Exercise 21

For a triangle ABC with corresponding angles a, b and c, show that

$$\sin a + \sin b + \sin c \leq \frac{3\sqrt{3}}{2}$$

and determine when equality holds. (Hint: $y = \sin x$ is concave)

Solution. Since $f(x) = \sin x$ is strictly concave on $[0, \pi]$,

$$\begin{aligned}
&\frac{1}{3}f(a) + \frac{1}{3}f(b) + \frac{1}{3}f(c) \\
&= \frac{1}{3}f(a) + \frac{2}{3}\left(\frac{1}{2}f(b) + \frac{1}{2}f(c)\right) \\
&\leq \frac{1}{3}f(a) + \frac{2}{3}\left(f\left(\frac{b}{2} + \frac{c}{2}\right)\right) \quad \text{[Concavity Inequality]} \\
&\leq f\left(\frac{a}{3} + \frac{2}{3}\left(\frac{b+c}{2}\right)\right) \quad \text{[Concavity Inequality]} \\
&= f\left(\frac{a+b+c}{3}\right)
\end{aligned}$$

Hence

$$\sin a + \sin b + \sin c = f(a) + f(b) + f(c) \leq 3f\left(\frac{a+b+c}{3}\right) = 3\sin\frac{\pi}{3} = \frac{3\sqrt{3}}{2}.$$

Equality holds when $a = b = c$, i.e. when ABC is an equilateral triangle. □

§2.4 *Boundedness*

Let S be a set.

Definition 2.53 (Order)**.** An ***order*** on S is a relation, denoted by $<$, with the following properties:

(i) (**trichotomy**) $\forall x, y \in S$, one and only one of the statements

$$x < y, \quad x = y, \quad y < x$$

is true.

(ii) (**transitivity**) $\forall x, y, z \in S$, if $x < y$ and $y < z$, then $x < z$.

Notation. $x \leq y$ indicates that $x < y$ or $x = y$, without specifying which of these two is to hold. In other words, $x \leq y$ is the negation of $x > y$.

Definition 2.54 (Ordered set)**.** An ***ordered set*** is a set S in which an order is defined.

Example 2.55
Q is an ordered set if $r < s$ is defined to mean that $s - r$ is a positive rational number.

Definition 2.56. Suppose S is an ordered set, and $E \subset S$.

(1) $M \in S$ is an ***upper bound*** of E if $x \leq M$ for all $x \in E$.
E is ***bounded above*** if there exists an upper bound $M \in S$.

(2) $m \in S$ is a ***lower bound*** of E if $x \geq m$ for all $x \in E$.
E is ***bounded below*** if there exists a lower bound $m \in S$.

(3) E is ***bounded*** in S if it is bounded above and below.

Definition 2.57 (Supremum)**.** Suppose S is an ordered set, $E \subset S$, and E is bounded above. Suppose there exists $\alpha \in S$ with the following properties:

(i) α is an upper bound for E;

(ii) if $\beta < \alpha$ then β is not an upper bound of E, i.e. $\exists x \in S$ s.t. $x > \beta$ (least upper bound).

Then we call α the ***supremum*** of E, and we write $\alpha = \sup E$.

Definition 2.58 (Infimum)**.** Suppose there exists $\alpha \in S$ with the following properties:

(i) α is a lower bound for E;

(ii) if $\beta > \alpha$ then β is not a lower bound of E, i.e. $\exists x \in S$ s.t. $x < \beta$ (greatest lower bound).

Then we call α the ***infimum*** of E, and we write $\alpha = \inf E$.

Proposition 2.59 (Uniqueness of suprenum)**.** If E has a supremum, then it is unique.

Proof. Assume that M and N are suprema of E.

Since N is a supremum, it is an upper bound for E. Since M is a supremum, then it is the least upper bound and thus $M \leq N$.

Similarly, since M is a supremum, it is an upper bound for E; since N is a supremum, it is a least upper bound and thus $N \leq M$.

Since $N \leq M$ and $M \leq N$, thus $M = N$. Therefore, a supremum for a set is unique if it exists. □

Definition 2.60. An ordered set S is said to have the ***least-upper-bound property*** (l.u.b.) if the following is true: if non-empty $E \subset S$ is bounded above, then $\sup E$ exists in S.

We shall now show that there is a close relation between greatest lower bounds and least upper bounds, and that every ordered set with the least-upper-bound property also has the greatest-lower-bound property.

Theorem 2.61. Suppose S is an ordered set with the least-upper-bound property, $B \subset S$, B is not empty, and B is bounded below. Let L be the set of all lower bounds of B. Then

$$\alpha = \sup L$$

exists in S, and $\alpha = \inf B$.

In particular, $\inf B$ exists in S.

Proof. Since B is bounded below, L is not empty. Since L consists of exactly those $y \in S$ which satisfy the inequality $y \leq x$ for every $x \in B$, we see that every $x \in B$ is an upper bound of L. Thus L is bounded above. Our hypothesis about S thus implies that L has a supremum in S; call it α.

If $\gamma < \alpha$ then γ is not an upper bound of L, hence $\gamma \notin B$. It follows that $\alpha \leq x$ for every $x \in B$. Thus $\alpha \in L$.

If $\alpha < \beta$ then $\beta \notin L$, since α is an upper bound of L.

We have shown that $\alpha \in L$ but $\beta \notin L$ if $\beta > \alpha$. In other words, α is a lower bound of B, but β is not if $\beta > \alpha$. This means that $\alpha = \inf B$. □

Theorem 2.62 (Comparison Theorem)**.** Let $S, T \subset \mathbf{R}$ be non-empty sets such that $s \leq t$ for every $s \in S$ and $t \in T$. If T has a supremum, then so does S, and $\sup S \leq \sup T$.

Proof. Let $\tau = \sup T$. Since τ is a supremum for T, then $t \leq \tau$ for all $t \in T$. Let $s \in S$ and choose any $t \in T$. Then, since $s \leq t$ and $t \leq \tau$, then $s \leq t$. Thus, τ is an upper bound for S.

By the Completeness Axiom, S has a supremum, say $\sigma = \sup S$. We will show that $\sigma \leq \tau$. Notice that, by the above, τ is an upper bound for S. Since σ is the least upper bound for S, then $\sigma \leq \tau$. Therefore,

$$\sup S \leq \sup T.$$

□

Let's explore some useful properties of sup and inf.

Proposition 2.63. Let S, T be non-empty subsets of $\mathbf{R}$, with $S \subseteq T$ and with T bounded above. Then S is bounded above, and $\sup S \leq \sup T$.

Proof. Since T is bounded above, it has an upper bound, say b. Then $t \leq b$ for all $t \in T$, so certainly $t \leq b$ for all $t \in S$, so b is an upper bound for S.

Now S, T are non-empty and bounded above, so by completeness each has a supremum. Note that $\sup T$ is an upper bound for T and hence also for S, so $\sup T \geq \sup S$ (since $\sup S$ is the least upper bound for S). □

Proposition 2.64. Let $T \subseteq \mathbf{R}$ be non-empty and bounded below. Let $S = \{-t \mid t \in T\}$. Then S is non-empty and bounded above. Furthermore, $\inf T$ exists, and $\inf T = -\sup S$.

Proof. Since T is non-empty, so is S. Let b be a lower bound for T, so $t \geq b$ for all $t \in T$. Then $-t \leq -b$ for all $t \in T$, so $s \leq -b$ for all $s \in S$, so $-b$ is an upper bound for S.

Now S is non-empty and bounded above, so by completeness it has a supremum. Then $s \leq \sup S$ for all $s \in S$, so $t \geq -\sup S$ for all $t \in T$, so $-\sup S$ is a lower bound for T.

Also, we saw before that if b is a lower bound for T then $-b$ is an upper bound for S. Then $-b \geq \sup S$ (since $\sup S$ is the least upper bound), so $b \leq -\sup S$. So $-\sup S$ is the greatest lower bound.

So $\inf T$ exists and $\inf T = -\sup S$. □

Proposition 2.65 (Approximation Property)**.** Let $S \subseteq \mathbf{R}$ be non-empty and bounded above. For any $\epsilon > 0$, there is $s_\epsilon \in S$ such that $\sup S - \epsilon < s_\epsilon \leq \sup S$.

Proof. Take $\epsilon > 0$.

Note that by definition of the supremum we have $s \leq \sup S$ for all $s \in S$. Suppose, for a contradiction, that $\sup S - \epsilon \geq s$ for all $s \in S$.

Then $\sup S - \epsilon$ is an upper bound for S, but $\sup S - \epsilon < \sup S$, which is a contradiction.

Hence there is $s_\epsilon \in S$ with $\sup S - \epsilon < s_\epsilon$. □

Problem 12. Consider the set $\{\frac{1}{n} \mid n \in \mathbf{Z}^+\}$.

(a) Show that $\max S = 1$.

(b) Show that if d is a lower bound for S, then $d \leq 0$.

(c) Use (b) to show that $0 = \inf S$.

Proof. □

If we are dealing with rational numbers, the sup/inf of a set may not exist. For example, a set of numbers in $\mathbf{Q}$, defined by $\{[\pi \cdot 10^n]/10^n\}$. 3,3.1,3.14,3.141,3.1415,3.14159,... But this set does not have an infimum in $\mathbf{Q}$.

By ZFC, we have the Completeness Axiom, which states that any non-empty set $A \subset \mathbf{R}$ that is bounded above has a supremum; in other words, if A is a non-empty set of real numbers that is bounded above, there exists a $M \in \mathbf{R}$ such that $M = \sup A$.

Problem 13. Find, with proof, the supremum and/or infimum of $\{\frac{1}{n}\}$.

Proof. For the suprenum,

$$\sup\left\{\frac{1}{n}\right\} = \max\left\{\frac{1}{n}\right\} = 1.$$

For the infinum, for all positive a we can pick $n = [\frac{1}{a}] + 1$, then $a > \frac{1}{n}$. Hence

$$\inf\left\{\frac{1}{n}\right\} = 0.$$

□

Problem 14. Find, with proof, the supremum and/or infimum of $\{\sin n\}$.

Proof. The answer is easy to guess: ± 1

For the supremum, we need to show that 1 is the smallest we can pick, so for any $a = 1 - \epsilon < 1$ we want to find an integer n close enough to $2k\pi + \dfrac{\pi}{2}$ so that $\sin n > a$.

Whenever we want to show the approximations between rational and irrational numbers we should think of the **pigeonhole principle**.

$$2k\pi + \frac{\pi}{2} = 6k + (2\pi - 6)k + \frac{\pi}{2}$$

Consider the set of fractional parts $\{(2\pi - 6)k\}$. Since this an infinite set, for any small number δ there is always two elements $\{(2\pi - 6)a\} < \{(2\pi - 6)b\}$ such that

$$|\{(2\pi - 6)b\} - \{(2\pi - 6)a\}| < \epsilon$$

Then $\{(2\pi - 6)(b - a)\} < \delta$

We then multiply by some number m (basically adding one by one) so that

$$0 \leq \{(2\pi - 6) \cdot m(b - a)\} - \left(2 - \frac{\pi}{2}\right) < \delta$$

Picking $k = m(b - a)$ thus gives

$$\begin{aligned} 2k\pi + \frac{\pi}{2} &= 6k + (2\pi - 6)k + \frac{\pi}{2} \\ &= 6k + [(2\pi - 6)k] + 2 + (2\pi - 6)k - \left(2 - \frac{\pi}{2}\right) \end{aligned}$$

Thus $n = 6k + [(2\pi - 6)k] + 2$ satisfies $\left|2k\pi + \dfrac{\pi}{2} - n\right| < \delta$

Now we're not exactly done here because we still need to talk about how well $\sin n$ approximates to 1.

We need one trigonometric fact: $\sin x < x$ for $x > 0$. (This simply states that the area of a sector in the unit circle is larger than the triangle determined by its endpoints.)

$$\begin{aligned}\sin n &= \sin\left(n - \left(2k\pi + \frac{\pi}{2}\right) + \left(2k\pi + \frac{\pi}{2}\right)\right) \\ &= \cos\left(n - \left(2k\pi + \frac{\pi}{2}\right)\right) \\ &= \cos\theta\end{aligned}$$

$$1 - \sin n = 2\sin^2\frac{\theta}{2} = 2\sin^2\left|\frac{\theta}{2}\right| \leq \frac{\theta^2}{2} < \delta$$

Hence we simply pick $\delta = \epsilon$ to ensure that $1 - \sin n < \epsilon$, and we're done. □

§2.5 *Cardinality*

Definition 2.66. If there exists a bijective mapping of A onto B, we say that A and B can be put in ***1-1 correspondence***, or that A and B have the same ***cardinal number***, or, briefly, that A and B are ***equivalent***, denoted by $A \sim B$ (an equivalence relation).

Notation. For any positive integer n, let J_n be the set whose elements are the integers $1, 2, \ldots, n$. Let J be the set consisting of all positive integers.

Definition 2.67. For any set A, we say

- A is ***finite*** if $A \sim J_n$ for some n (the empty set is also considered to be finite)
- A is ***infinite*** if A is not finite.
- A is ***countable*** if $A \sim J$.
- A is ***uncountable*** if A is neither finite nor countable.
- A is ***at most countable*** if A is finite or countable.

For two finite sets A and B, we evidently have $A \sim B$ if and only if A and B contain the same number of elements.

For infinite sets, however, the idea of "having the same number of elements" becomes quite vague, whereas the notion of bijectivity retains its clarity.

Proposition 2.68. $2J = \{2n \mid n \in J\}$ is countable.

Proof. We can find the function $f : J \to 2J$ given by

$$f(n) = 2n$$

which is bijective. Thus there is a 1-1 correspondence between J and $2J$. □

Proposition 2.69. $\mathbf{Z}$ is countable.

Proof. Consider the following arrangement of the sets $\mathbf{Z}$ and J:

$$\begin{aligned} \mathbf{Z}: &\quad 0, 1, -1, 2, -2, 3, -3, \ldots \\ J: &\quad 1, 2, 3, 4, 5, 6, 7, \ldots \end{aligned}$$

We can even give an explicit formula for a bijective function $f : J \to \mathbf{Z}$:

$$f(n) = \begin{cases} \dfrac{n}{2} & \text{if } n \text{ is even,} \\ -\dfrac{n-1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

□

Proposition 2.70. Every infinite subset of a countable set A is countable.

Proof. Suppose $E \subset A$, and E is infinite. Arrange the elements $x \in A$ in a sequence $\{x_n\}$ of distinct elements.

Construct a sequence $\{n_k\}$ as follows: Let n_1 be the smallest positive integer such that $x_{n_1} \in E$. Having chosen $n_1, \ldots, n_{k-1}$ $(k = 2, 3, 4, \ldots)$, let n_k be the smallest integer greater than n_{k-1} such that $x_{n_k} \in E$.

Putting $f(k) = x_{n_k}$ $(k = 1, 2, 3, \ldots)$, we obtain a 1-1 correspondence between E and J. □

This shows that countable sets represent the "smallest" infinity: No uncountable set can be a subset of a countable set.

Proposition 2.71. Let $\{E_n \mid n \in J\}$ be a sequence of countable sets, and put

$$S = \bigcup_{n=1}^{\infty} E_n.$$

Then S is countable.

Proof. Let every set E_n be arranged in a sequence $\{x_{n_k}\}$ $(k = 1, 2, 3, \dots)$, and consider the infinite array

$$\begin{array}{llll} x_{11} & x_{12} & x_{13} & x_{14} & \cdots \\ x_{21} & x_{22} & x_{23} & x_{24} & \cdots \\ x_{31} & x_{32} & x_{33} & x_{34} & \cdots \\ x_{41} & x_{42} & x_{43} & x_{44} & \cdots \\ \vdots & & & & \end{array}$$

in which the elements of E_n form the n-th row. The array contains all elements of S. These elements can be arranged in a sequence

$$x_{11}, x_{21}, x_{12}, x_{31}, x_{22}, x_{13}, x_{41}, x_{32}, x_{23}, x_{14}, \dots$$

□

Proposition 2.72. Let A and B be finite sets. Then $|A \cup B| = |A| + |B| - |A \cap B|$.

Proof. The proof is left as an exercise. □

Proposition 2.73 (Subsets of a finite set)**.** If a set A is finite with $|A| = n$, then its power set has $|\mathcal{P}(A)| = 2^n$.

Proof. We use induction. For the initial step, note that if $|A| = 0$ then $A = \emptyset$ has no elements, so there is a single subset $\emptyset$, and therefore $|\mathcal{P}(A)| = 1 = 2^0$.

Now suppose that $n \geq 0$ and that $|P(S)| = 2^n$ for any set S with $|S| = n$. Let A be any set with $|A| = n + 1$. By definition, this means that there is an element a and a set $A_0 = A \setminus \{a\}$ with $|A_0| = n$. Any subset of A must either contain the element a or not, so we can partition $\mathcal{P}(A) = P(A_0) \cup \{S \cup \{a\} \mid S \in P(A_0)\}$. These two sets are disjoint, and each of them has cardinality $|P(A_0)| = 2^n$ by the inductive hypothesis. Hence $|\mathcal{P}(A)| = 2^n + 2^n = 2^{n+1}$.

Thus, by induction, the result holds for all n. □

Another way to see this is through combinatorics: Consider the process of creating a subset. We can do this systematically by going through each of the $|A|$ elements in A and making the yes/no decision whether to put it in the subset. Since there are $|A|$ such choices, that yields $2^{|A|}$ different combinations of elements and therefore $2^{|A|}$ different subsets.

Theorem 2.74 (Principle of Inclusion and Exclusion)**.** Let S_i be finite sets. Then

$$\left|\bigcup_{i=1}^{n} S_i\right| = \sum_{i=1} |S_i| - \sum_{1 \leq i < j \leq n} |S_i \cap S_j| + \sum_{1 \leq i < j < k \leq n} |S_i \cap S_j \cap S_k| + \cdots + (-1)^{n+1} \left|\bigcap_{i=1}^{n} S_i\right|. \tag{2.7}$$

Proof. By induction. □

The following more elegant proof was presented to the author by Dr. Ho Weng Kin during a H3 Mathematics lecture in 2024.

Proof. Let U be a finite set (interpreted as the universal set), and $S \subseteq U$. Define the characteristic/indicator function of S by

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

In other words,

$$x \in S \iff \chi_S(x) = 1$$

and equivalently,

$$x \notin S \iff \chi_S(x) = 0.$$

Let $S_1, S_2 \subseteq U$ be given. Then for any $x \in U$ it holds that

$$\chi_{S_1 \cap S_2}(x) = \chi_{S_1}(x) \cdot \chi_{S_2}(x)$$

where $\cdot$ denotes ordinary multiplication.

Similarly,

$$\chi_{S_1 \cup S_2}(x) = 1 - \big(1 - \chi_{S_1}(x)\big) \cdot \big(1 - \chi_{S_2}(x)\big).$$

In general, for any $x \in U$ it holds that

$$\chi_{S_1 \cup \cdots \cup S_n}(x) = 1 - \big(1 - \chi_{S_1}(x)\big) \cdots \big(1 - \chi_{S_n}(x)\big)$$

for any $S_1, \ldots, S_n \subset U$.

Since $x \in S$ if and only if $\chi_S(x) = 1$, it follows that

$$|S| = \sum_{x \in U} \chi_S(x).$$

To prove the PIE, we calculate

$$\begin{aligned}
&|S_1 \cup \cdots \cup S_n| \\
&= \sum_{x \in U} \chi_{S_1 \cup \cdots \cup S_n}(x) \\
&= \sum_{x \in U} 1 - \big(1 - \chi_{S_1}(x)\big) \cdots \big(1 - \chi_{S_n}(x)\big) \\
&= \big(\chi_{S_1}(x) + \cdots + \chi_{S_n}(x)\big) - \Big(\chi_{S_1}(x)\chi_{S_2}(x) + \cdots + \chi_{S_{n-1}}(x)\chi_{S_n}(x)\Big) + \cdots + (-1)^{n+1}\chi_{S_1}(x) \cdots \chi_{S_n}(x) \\
&= \big(\chi_{S_1}(x) + \cdots + \chi_{S_n}(x)\big) - \Big(\chi_{S_1 \cap S_2}(x) + \cdots + \chi_{S_{n-1} \cap S_n}(x)\Big) + \cdots + (-1)^{n+1}\chi_{S_1 \cap \cdots \cap S_n}(x) \\
&= \sum_{i=1}^{n} |S_i| - \sum_{J \subseteq \{1,\ldots,n\}, |J|=2} \left| \bigcap_{j \in J} S_j \right| + \cdots + (-1)^{k+1} \sum_{J \subseteq \{1,\ldots,n\}, |J|=k} \left| \bigcap_{j \in J} S_j \right| + \cdots + (-1)^{n+1} \left| \bigcap_{i=1}^{n} S_i \right|.
\end{aligned}$$

□

Theorem 2.75 (Cantor)**.** For a set A, $|A| < |\mathcal{P}(A)|$.

Proof. Define the function $f : A \to \mathcal{P}(A)$ by $f(x) = \{x\}$. Then, f is injective as $\{x\} = \{y\} \implies x = y$. Thus $|A| \leq |\mathcal{P}(A)|$. To finish the proof now all we need to show is that $|A| \neq |\mathcal{P}(A)|$. We will do so through contradiction. Suppose that $|A| = |\mathcal{P}(A)|$. Then, there exists a surjection $g : A \to \mathcal{P}(A)$. We define the set B as

$$B \coloneqq \{x \in A \mid x \notin g(x)\} \in \mathcal{P}(A)$$

Since g is surjective, there exists a $b \in A$ such that $g(b) = B$. There are two cases:

1. $b \in B$. Then $b \notin g(b) = B \implies b \notin B$.
2. $b \notin B$. Then $b \notin g(b) = B \implies b \in B$.

In either case we obtain a contradiction. Thus, g is not surjective so $|A| \neq |\mathcal{P}(A)|$. □

Corollary 2.76. For all $n \in \mathbf{N} \cup \{0\}$, $n < 2^n$.

Proof. This can be easily proven through induction. □

Exercises

Problem 15. Let A be the set of all complex polynomials in n variables. Given a subset $T \subset A$, define the *zeros* of T as the set

$$Z(T) = \{P = (a_1, \dots, a_n) \mid f(P) = 0 \text{ for all } f \in T\}$$

A subset $Y \in \mathbf{C}^n$ is called an algebraic set if there exists a subset $T \subset A$ such that $Y = Z(T)$.

Prove that the union of two algebraic sets is an algebraic set.

Proof. We would like to consider $T = \{f_1, f_2, \dots\}$ expressed as indexed sets $T = \{f_i\}$. Then $Z(T)$ can also be expressed as $\{P \mid \forall i, f_i(P) = 0\}$.

Suppose that we have two algebraic sets X and Y. Let $X = Z(S)$, $Y = Z(T)$ where S, T are subsets of A (basically, they are certain sets of polynomials). Then

$$X = \{P \mid \forall f \in S, f(P) = 0\}$$

$$Y = \{P \mid \forall g \in T, g(P) = 0\}$$

We imagine that for $P \in X \cap Y$, we have $f(P) = 0$ or $g(P) = 0$. Hence we consider the set of polynomials

$$U = \{f \cdot g \mid f \in S, g \in T\}$$

For any $P \in X \cup Y$ and for any $fg \in U$ where $f \in S$ and $f \in g$, either $f(P) = 0$ or $g(P) = 0$, hence $fg(P) = 0$ and thus $P \in Z(U)$.

On the other hand if $P \in Z(U)$, suppose otherwise that P is not in $X \cup Y$, then P is neither in X nor in Y. This means that there exists $f \in S, g \in T$ such that $f(P) \neq 0$ and $g(P) \neq 0$, hence $fg(P) \neq 0$. This is a contradiction as $P \in Z(U)$ implies $fg(P) = 0$. Hence we have $X \cup Y = Z(U)$ and thus $X \cup Y$ is an algebraic set.

Now the other direction is simpler and can actually be generalised: The intersection of arbitrarily many algebraic sets is algebraic.

The basic result is that if $X = Z(S)$ and $Y = Z(T)$ then $X \cap Y = Z(S \cup T)$. □

Problem 16 (Modular Arithmetic)**.** Define the ring of integers modulo n:

$$\mathbf{Z}/n\mathbf{Z} = \mathbf{Z}/\sim \text{ where } x \sim y \iff x - y \in n\mathbf{Z}.$$

The equivalence classes are called congruence classes modulo n.

(a) Define the sum of two congruence classes modulo n, $[x], [y] \in \mathbf{Z}/n\mathbf{Z}$, by

$$[x] + [y] = [x + y]$$

Show that the above definition is well-defined.

(b) Define the product of two congruence classes modulo n and show that such a definition is well-defined.

Solution.

(a) We often define such concepts by considering the **representatives** of the equivalence classes.

For example, here we define $[x] + [y] = [x + y]$ for two elements $[x]$ and $[y]$ in $\mathbf{Z}/n\mathbf{Z}$. So what we know here are the classes $[x]$ and $[y]$. But what exactly are x and y? They are just some element in the equivalence classes that was arbitrarily picked out. We then perform the sum $x + y$, and consequently, we used this to point towards the class $[x + y]$.

However, x and y are arbitrarily picked. We want to show that, regardless of which representatives are chosen from the equivalence classes $[x]$ and $[y]$, we will always obtain the same result.

In the definition itself, we have defined that, for the two representatives x and y we define $[x] + [y] = [x + y]$. So now, let's say that we take two other arbitrary representatives, $x' \in [x]$ and $y' \in [y]$. Then by definition, we have

$$[x] + [y] = [x' + y']$$

Thus, our goal is to show that $x' + y'] = [x + y]$. This expression means that the two sides of the equation are referring to the same equivalence class. Therefore, the expression above is completely equivalent to the condition.

$$x' + y' \sim x + y$$

We then check that this final expression is indeed true: Since $x' \in [x]$ and $y' \in [y]$, we have

$$\begin{aligned} &x' \sim x \text{ and } y' \sim y \\ &\implies x' - x, y' - y \in n\mathbf{Z} \\ &\implies (x' + y') - (x + y) = (x' - x) + (y' - y) \in n\mathbf{Z} \end{aligned}$$

(b) The product of two congruence classes is defined by

$$[x][y] = [xy]$$

For any other representatives x', y' we have

$$\begin{aligned} &x'y' - xy \\ &= x'y' - xy' + xy' - xy \\ &= (x' - x)y' + x(y' - y) \in n\mathbf{Z} \end{aligned}$$

Thus $[x'y'] = [xy]$ and the product is well-defined.

□

Problem 17. Let $A = \mathbf{R}$ and for any $x, y \in A$, $x \sim y$ if and only if $x - y \in \mathbf{Z}$. For any two equivalence classes $[x], [y] \in A/\sim$, define

$$[x] + [y] = [x + y] \text{ and } -[x] = [-x]$$

(a) Show that the above definitions are well-defined.

(b) Find a one-to-one correspondence $\phi : X \to Y$ between $X = A/\sim$ and $Y : |z| = 1$, i.e. the unit circle in $\mathbf{C}$, such that for any $[x_1], [x_2] \in X$ we have

$$\phi([x_1])\phi([x_2]) = \phi([x_1 + x_2])$$

(c) Show that for any $[x] \in X$,

$$\phi(-[x]) = \phi([x])^{-1}$$

Solution.

(a)

$$(x' + y') - (x + y) = (x' - x) + (y' - y) \in \mathbf{Z}$$

Thus $[x' + y'] = [x + y]$

$$(-x') - (-x) = -(x' - x) \in \mathbf{Z}$$

Thus $[-x'] = [-x]$.

(b) Complex numbers in the polar form: $z = re^{i\theta}$

Then the correspondence is given by $\phi([x]) = e^{2\pi ix}$

$$[x] = [y] \iff x - y \in \mathbf{Z} \iff e^{2\pi i(x-y)} = 1 \iff e^{2\pi ix} = e^{2\pi iy}$$

Hence this is a bijection.

Before that, we also need to show that ϕ is well-defined, which is almost the same as the above.

If we choose another representative x' then

$$\phi([x]) = e^{2\pi ix'} = e^{2\pi ix} \cdot e^{2\pi i(x'-x)} = e^{2\pi ix}$$

(c) You can either refer to the specific correspondence $\phi([x]) = e^{2\pi ix}$ or use its properties.

$$\phi(-[x])\phi([x]) = \phi([-x])\phi([x]) = \phi([-x + x]) = \phi([0]) = 1$$

□

Problem 18 (Complex Numbers)**.** Let $\mathbf{R}[x]$ denote the set of real polynomials. Define

$$\mathbf{C} = \mathbf{R}[x]/(x^2+1)\mathbf{R}[x]$$

where

$$f(x) \sim g(x) \iff x^2+1 \text{ divides } f(x)-g(x).$$

The complex number $a+bi$ is defined to be the equivalence class of $a+bx$.

(a) Define the sum and product of two complex numbers and show that such definitions are well-defined.

(b) Define the reciprocal of a complex number.

Part II

Linear Algebra

3 Vector Spaces

§3.1 *Definition of Vector Space*

Notation. $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.

Notation. $\mathbf{F}^n$ is the set of n-tuples whose elements belong to $\mathbf{F}$:

$$\mathbf{F}^n := \{(x_1, \ldots, x_n) \mid x_i \in \mathbf{F}\}$$

For $(x_1, \ldots, x_n) \in \mathbf{F}^n$ and $i = 1, \ldots, n$, we say that x_i is the i-th coordinate of $(x_1, \ldots, x_n)$.

Definition 3.1 (Vector space)**.** V is a ***vector space*** over $\mathbf{F}$ if the following properties hold:

(i) Addition is commutative: $u + v = v + u$ for all $u, v \in V$

(ii) Addition is associative: $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$
Multiplication is associative: $(ab)v = a(bv)$ for all $v \in V$, $a, b \in \mathbf{F}$

(iii) Additive identity: there exists $\mathbf{0} \in V$ such that $v + \mathbf{0} = v$ for all $v \in V$

(iv) Additive inverse: for every $v \in V$, there exists $w \in V$ such that $v + w = \mathbf{0}$

(v) Multiplicative identity: $1v = v$ for all $v \in V$

(vi) Distributive properties: $a(u + v) = au + av$ and $(a + b)v = av + bv$ for all $a, b, \in \mathbf{F}$ and $u, v \in V$

Notation. For the rest of this text, V denotes a vector space over $\mathbf{F}$.

> **Example 3.2**
> $\mathbf{R}^n$ is a vector space over $\mathbf{R}$, $\mathbf{C}^n$ is a vector space over $\mathbf{C}$.

Elements of a vector space are called ***vectors*** or ***points***.

The scalar multiplication in a vector space depends on $\mathbf{F}$. Thus when we need to be precise, we will say that V is a vector space over $\mathbf{F}$ instead of saying simply that V is a vector space. For example, $\mathbf{R}^n$ is a vector space over $\mathbf{R}$, and $\mathbf{C}^n$ is a vector space over $\mathbf{C}$. A vector space over $\mathbf{R}$ is called a ***real vector space***; a vector space over $\mathbf{C}$ is called a ***complex vector space***.

Proposition 3.3 (Uniqueness of additive identity)**.** A vector space has a unique additive identity.

Proof. Suppose otherwise, then $\mathbf{0}$ and $\mathbf{0}'$ are additive identities of V. Then

$$\mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}$$

where the first equality holds because $\mathbf{0}$ is an additive identity, the second equality comes from commutativity, and the third equality holds because $\mathbf{0}'$ is an additive identity. Thus $\mathbf{0}' = \mathbf{0}$. □

Proposition 3.4 (Uniqueness of additive inverse)**.** Every element in a vector space has a unique additive inverse.

Proof. Suppose otherwise, then for $v \in V$, w and w' are additive inverses of v. Then

$$w = w + \mathbf{0} = w + (v + w') = (w + v) + w' = \mathbf{0} + w' = w'.$$

Thus $w = w'$. □

Because additive inverses are unique, the following notation now makes sense.

Notation. Let $v, w \in V$. Then $-v$ denotes the additive inverse of v; $w - v$ is defined to be $w + (-v)$.

Proposition 3.5 (The number 0 times a vector)**.** For every $v \in V$, $0v = \mathbf{0}$.

Proof. For $v \in V$, we have

$$0v = (0 + 0)v = 0v + 0v.$$

Adding the additive inverse of $0v$ to both sides of the equation gives $\mathbf{0} = 0v$. □

Proposition 3.6 (A number times the vector 0)**.** For every $a \in \mathbf{F}$, $a\mathbf{0} = \mathbf{0}$.

Proof. For $a \in \mathbf{F}$, we have

$$a\mathbf{0} = a(\mathbf{0} + \mathbf{0}) = a\mathbf{0} + a\mathbf{0}.$$

Adding the additive inverse of $a\mathbf{0}$ to both sides of the equation gives $\mathbf{0} = a\mathbf{0}$. □

Now we show that if an element of V is multiplied by the scalar 1, then the result is the additive inverse of the element of V.

Proposition 3.7 (The number -1 times a vector)**.** For every $v \in V$, $(-1)v = -v$.

Proof. For $v \in V$, we have

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = \mathbf{0}.$$

Since $v + (-1)v = \mathbf{0}$, $(-1)v$ is the additive inverse of v. □

Example 3.8

$\mathbf{F}^\infty$ is defined to be the set of all sequences of elements of $\mathbf{F}$:

$$\mathbf{F}^\infty := \{(x_1, x_2, \dots) \mid x_i \in \mathbf{F}\}$$

Addition and scalar multiplication on $\mathbf{F}^\infty$ are defined as expected:

$$(x_1, x_2, \dots) + (y_1, y_2, \dots) = (x_1 + y_1, x_2 + y_2, \dots)$$
$$\lambda(x_1, x_2, \dots) = (\lambda x_1, \lambda x_2, \dots)$$

With these definitions, $\mathbf{F}^\infty$ becomes a vector space over $\mathbf{F}$, as you should verify. The additive identity in this vector space is $\mathbf{0} = (0, 0, \dots)$.

Our next example of a vector space involves a set of functions.

Notation. If S is a set, $\mathbf{F}^S := \{f \mid f : S \to \mathbf{F}\}$.

For $f, g \in \mathbf{F}^S$, the sum $f + g \in \mathbf{F}^S$ is the function defined by

$$(f + g)(x) = f(x) + g(x) \quad (\forall x \in S)$$

For $\lambda \in \mathbf{F}$, $f \in \mathbf{F}^S$, the product $\lambda f \in \mathbf{F}^S$ is the function defined by

$$(\lambda f)(x) = \lambda f(x) \quad (\forall x \in S)$$

Example 3.9
If $S = [0, 1]$ and $\mathbf{F} = \mathbf{R}$, then $\mathbf{R}^{[0,1]}$ is the set of real-valued functions on the interval $[0, 1]$.

Example 3.10
If S is a non-empty set, then $\mathbf{F}^S$ (with the operations of addition and scalar multiplication as defined above) is a vector space over $\mathbf{F}$.

Additive identity of $\mathbf{F}^S$ is the function $0 : S \to \mathbf{F}$ defined by

$$0(x) = 0 \quad (\forall x \in S)$$

For $f \in \mathbf{F}^S$, additive inverse of f is the function $-f : S \to \mathbf{F}$ defined by

$$(-f)(x) = -f(x) \quad (\forall x \in S)$$

It is easy to see that $\mathbf{F}^n$ and $\mathbf{F}^\infty$ are special cases of the vector space $\mathbf{F}^S$ because a list of length n of numbers in $\mathbf{F}$ can be thought of as a function from $\{1, 2, \ldots, n\}$ to $\mathbf{F}$ and a sequence of numbers in $\mathbf{F}$ can be thought of as a function from the set of positive integers to $\mathbf{F}$. In other words, we can think of $\mathbf{F}^n$ as $\mathbf{F}^{\{1,2,\ldots,n\}}$ and we can think of $\mathbf{F}^\infty$ as $\mathbf{F}^{\{1,2,\ldots\}}$.

§3.2 *Subspaces*

Definition 3.11 (Subspace)**.** $U \subset V$ is a ***subspace*** of V if U is also a vector space (with the same addition and scalar multiplication as on V).

Lemma 3.12 (Conditions for a subspace)**.** $U \subset V$ is a subspace of V if and only if U satisfies the following conditions:

(i) Additive identity: $\mathbf{0} \in U$

(ii) Closed under addition: $u + w \in U$ for all $u, w \in U$

(iii) Closed under scalar multiplication: $\lambda u \in U$ for all $\lambda \in \mathbf{F}$, $u \in U$

Proof. If U is a subspace of V, then U satisfies the three conditions above by the definition of vector space.

Conversely, suppose U satisfies the three conditions above. (i) ensures that the additive identity of V is in U. (ii) ensures that addition makes sense on U. (iii) ensures that scalar multiplication makes sense on U.

If $u \in U$, then $-u = (-1)u \in U$ by (iii). Hence every element of U has an additive inverse in U.

The other parts of the definition of a vector space, such as associativity and commutativity, are automatically satisfied for U because they hold on the larger space V. Thus U is a vector space and hence is a subspace of V. □

Remark. The three conditions in Lemma 3.12 usually enable us to determine quickly whether a given subset of V is a subspace of V.

Definition 3.13 (Sum of subsets)**.** Suppose $U_1, \ldots, U_n \subset V$. The ***sum*** of $U_1, \ldots, U_n$ is the set of all possible sums of elements of $U_1, \ldots, U_n$:

$$U_1 + \cdots + U_n \coloneqq \{u_1 + \cdots + u_n \mid u_i \in U_i\}.$$

Example 3.14

Suppose that $U = \{(x, 0, 0) \in \mathbf{F}^3 \mid x \in F\}$ and $W = \{(0, y, 0) \in \mathbf{F}^3 \mid y \in \mathbf{F}\}$. Then

$$U + W = \{(x, y, 0) \mid x, y \in \mathbf{F}\}.$$

Example 3.15

Suppose that $U = \{(x, x, y, y) \in \mathbf{F}^4 \mid x, y \in \mathbf{F}\}$ and $W = \{(x, x, x, y) \in \mathbf{F}^4 \mid x, y \in \mathbf{F}\}$. Then

$$U + W = \{(x, x, y, z) \in \mathbf{F}^4 \mid x, y, z \in \mathbf{F}\}.$$

The next result states that the sum of subspaces is a subspace, and is in fact the smallest subspace containing all the summands.

Proposition 3.16. Suppose $U_1, \ldots, U_n$ are subspaces of V. Then $U_1 + \cdots + U_n$ is the smallest subspace of V containing $U_1, \ldots, U_n$.

Proof. It is easy to see that $\mathbf{0} \in U_1 + \cdots + U_n$ and that $U_1 + \cdots + U_n$ is closed under addition and scalar multiplication. Hence $U_1 + \cdots + U_n$ is a subspace of V.

Clearly $U_1, \ldots, U_n$ are all contained in $U_1 + \cdots + U_n$ (to see this, consider sums $u_1 + \cdots + u_n$ where all except one of the u's are $\mathbf{0}$). Conversely, every subspace of V containing $U_1, \ldots, U_n$ contains $U_1 + \cdots + U_n$ (because subspaces must contain all finite sums of their elements). Thus $U_1 + \cdots + U_n$ is the smallest subspace of V containing $U_1, \ldots, U_n$. □

Remark. Sums of subspaces in the theory of vector spaces are analogous to unions of subsets in set theory. Given two subspaces of a vector space, the smallest subspace containing them is their sum. Analogously, given two subsets of a set, the smallest subset containing them is their union.

Definition 3.17 (Direct sum)**.** Suppose $U_1, \ldots, U_n$ are subspaces of V. The sum $U_1 + \cdots + U_n$ is called a ***direct sum*** if each element of $U_1 + \cdots + U_n$ can be written in only one way a sum of $u_1 + \cdots + u_n$, $u_i \in U_i$. In this case, we denote the sum as

$$U_1 \oplus \cdots \oplus U_n.$$

Example 3.18

Suppose that $U = \{(x, y, 0) \in \mathbf{F}^3 \mid x, y \in \mathbf{F}\}$ and $W = \{(0, 0, z) \in \mathbf{F}^3 \mid z \in \mathbf{F}\}$. Then $\mathbf{F}^3 = U \oplus W$.

Example 3.19

Suppose U_i is the subspace of $\mathbf{F}^n$ of those vectors whose coordinates are all 0 except for the i-th coordinate; that is, $U_i = \{(0, \ldots, 0, x, 0, \ldots, 0) \in \mathbf{F}^n \mid x \in \mathbf{F}\}$. Then $\mathbf{F}^n = U_1 \oplus \cdots \oplus U_n$.

Lemma 3.20 (Condition for direct sum)**.** Suppose $V_1, \ldots, V_n$ are subspaces of V, $W = V_1 + \cdots + V_n$. Then the following are equivalent:

(i) Any element in W can be uniquely expressed as the sum of vectors in $V_1, \ldots, V_n$.

(ii) If $v_i \in V_i$ satisfies $v_1 + \cdots + v_n = \mathbf{0}$, then $v_1 = \cdots = v_n = \mathbf{0}$.

(iii) For $k = 2, \ldots, n$, $(V_1 + \cdots + V_{k-1}) \cap V_k = \{\mathbf{0}\}$.

Proof.

(i) $\iff$ (ii) First suppose W is a direct sum. Then by the definition of direct sum, the only way to write $\mathbf{0}$ as a sum $u_1 + \cdots + u_n$ is by taking $u_i = \mathbf{0}$.

Now suppose that the only way to write $\mathbf{0}$ as a sum $v_1 + \cdots + v_n$ by taking $v_1 = \cdots = v_n = \mathbf{0}$. For $v \in V_1 + \cdots + V_n$, suppose that there is more than one way to represent v:

$$v = v_1 + \cdots + v_n$$
$$v = v_1' + \cdots + v_n'$$

for some $v_i, v_i' \in V_i$. Substracting the above two equations gives

$$\mathbf{0} = (v_1 - v_1') + \cdots + (v_n - v_n').$$

Since $v_i - v_i' \in V_i$, we have $v_i - v_i' = \mathbf{0}$ so $v_i = v_i'$. Hence there is only one unique way to represent $v_1 + \cdots + v_n$, thus W is a direct sum.

(ii) $\iff$ (iii) First suppose if $v_i \in V_i$ satisfies $v_1 + \cdots + v_n = \mathbf{0}$, then $v_1 = \cdots = v_n = \mathbf{0}$. Let $v_k \in (V_1 + \cdots + V_{k-1}) \cap V_k$. Then $v_k = v_1 + \cdots + v_{k-1}$ where $v_i \in V_i$ $(1 \leq i \leq k-1)$. Thus

$$\begin{aligned} v_1 + \cdots + v_{k-1} - v_k &= \mathbf{0} \\ v_1 + \cdots + v_{k-1} + (-v_k) + \mathbf{0} + \cdots + \mathbf{0} &= \mathbf{0} \end{aligned}$$

by taking $v_{k+1} = \cdots = v_n = \mathbf{0}$. Then $v_1 = \cdots = v_k = \mathbf{0}$.

Now suppose that for $k = 2, \ldots, n$, $(V_1 + \cdots + V_{k-1}) \cap V_k = \{\mathbf{0}\}$.

$$\begin{aligned} v_1 + \cdots + v_n &= \mathbf{0} \\ v_1 + \cdots + v_{n-1} &= -v_n \end{aligned}$$

where $v_1 + \cdots + v_{n-1} \in V_1 + \cdots + V_{n-1}$, $-v_n \in V_n$. Thus

$$v_1 + \cdots + v_{n-1} = -v_n \in (V_1 + \cdots + V_{n-1}) \cap V_n = \{\mathbf{0}\}$$

so $v_1 + \cdots + v_{n-1} = \mathbf{0}$, $v_n = \mathbf{0}$. Induction on n gives $v_1 = \cdots = v_{n-1} = v_n = \mathbf{0}$. □

Proposition 3.21. Suppose U and W are subspaces of V. Then $U + W$ is a direct sum if and only if $U \cap W = \{\mathbf{0}\}$.

Proof. First suppose that $U + W$ is a direct sum. If $v \in U \cap W$, then $\mathbf{0} = v + (-v)$, where $v \in U$, $-v \in W$. By the unique representation of $\mathbf{0}$ as the sum of a vector in U and a vector in W, we have $v = \mathbf{0}$. Thus $U \cap W = \{\mathbf{0}\}$.

Now suppose $U \cap W = \{\mathbf{0}\}$. To prove that $U + W$ is a direct sum, suppose $u \in U$, $w \in W$, and

$$0 = u + w.$$

$u = -w \in W$, thus $u \in U \cap W$, so $u = w = \mathbf{0}$. By Lemma 3.20, $U + W$ is a direct sum. □

Exercises

Problem 19. Suppose W is a vector space over $\mathbf{F}$, V_1 and V_2 are subspaces of W. Show that $V_1 \cap V_2$ is a vector space over $\mathbf{F}$ if and only if $V_1 \subset V_2$ or $V_2 \subset V_1$.

Solution. The backward direction is trivial. We focus on proving the forward direction.

Supppse otherwise, then $V_1 \setminus V_2 \neq \emptyset$ and $V_2 \setminus V_1 \neq \emptyset$. Pick $v_1 \in V_1 \setminus V_2$ and $v_2 \in V_2 \setminus V_1$. Then

$$\begin{aligned} v_1, v_2 \in V_1 \cup V_2 &\implies v_1 + v_2 \in V_1 \cup V_2 \\ &\implies v_2, v_1 + v_2 \in V_2 \\ &\implies v_1 = (v_1 + v_2) - v_2 \in V_2 \end{aligned}$$

which is a contradiction. □

Problem 20. Suppose W is a vector space over $\mathbf{F}$, V_1, V_2, V_3 are subspaces of W. Then $V_1 \cup V_2 \cup V_3$ is a vector space over $\mathbf{F}$ if and only if one of the V_i contains the other two.

Solution. We prove the forward direction. Suppose otherwise, then $v_1 \in V_1 \setminus (V_2 + V_3)$, $v_2 \in V_2 \setminus (V_1 + V_3)$, $v_3 \in V_3 \setminus (V_1 + V_2)$. Consider

$$\{v_1 + v_2 + v_3, v_1 + v_2 + 2v_3, v_1 + 2v_2 + v_3, v_1 + 2v_2 + 2v_3\} \subset V_1 \cup V_2 \cup V_3$$

Then

$$\begin{aligned} &(v_1 + v_2 + 2v_3) - (v_1 + v_2 + v_3) = v_3 \notin V_1 + V_2 \\ &\implies v_1 + v_2 + v_3 \notin V_1 + V_2 \quad \text{or} \quad v_1 + v_2 + 2v_3 \notin V_1 + V_2 \\ &\implies v_1 + v_2 + v_3 \in V_3 \quad \text{or} \quad v_1 + v_2 + 2v_3 \in V_3 \\ &\implies v_1 + v_2 \in V_3 \end{aligned}$$

Similarly,

$$\begin{aligned} &(v_1 + 2v_2 + 2v_3) - (v_1 + 2v_2 + v_3) = v_3 \notin V_1 + V_2 \\ &\implies v_1 + 2v_2 + v_3 \notin V_1 + V_2 \quad \text{or} \quad v_1 + 2v_2 + 2v_3 \notin V_1 + V_2 \\ &\implies v_1 + 2v_2 + v_3 \in V_3 \quad \text{or} \quad v_1 + 2v_2 + 2v_3 \in V_3 \\ &\implies v_1 + 2v_2 \in V_3 \end{aligned}$$

This implies $(v_1 + 2v_2) - (v_1 + v_2) = v_2 \in V_3$, a contradiction. □

4 Finite-Dimensional Vector Spaces

§4.1 *Span and Linear Independence*

Definition 4.1 (Linear combination)**.** A ***linear combination*** of vectors $\{v_1, \ldots, v_n\}$ in V is a vector of the form

$$a_1 v_1 + \cdots + a_n v_n$$

where $a_i \in \mathbf{F}$.

Definition 4.2 (Span)**.** The ***span*** of $\{v_1, \ldots, v_n\}$ is the set of all linear combinations of $v_1, \ldots, v_n$:

$$\operatorname{span}\{v_1, \ldots, v_n\} = \{a_1 v_1 + \cdots + a_n v_n \mid a_i \in \mathbf{F}\}.$$

The span of the empty list ( ) is defined to be $\{\mathbf{0}\}$.

We say that $v_1, \ldots, v_n$ ***spans*** V if $\operatorname{span}\{v_1, \ldots, v_n\} = V$.

Proposition 4.3. $\operatorname{span}\{v_1, \ldots, v_n\}$ in V is the smallest subspace of V containing $v_1, \ldots, v_n$.

Proof. First we show that $\operatorname{span}\{v_1, \ldots, v_n\}$ is a subspace of V.

(i) Additive identity $\mathbf{0} = 0v_1 + \cdots + 0v_n \in \operatorname{span}\{v_1, \ldots, v_n\}$

(ii) $(a_1 v_1 + \cdots + a_n v_n) + (c_1 v_1 + \cdots + c_n v_n) = (a_1 + c_1)v_1 + \cdots + (a_n + c_n)v_n \in \operatorname{span}\{v_1, \ldots, v_n\}$, so $\operatorname{span}\{v_1, \ldots, v_n\}$ is closed under addition.

(iii) $\lambda(a_1 v_1 + a_n v_n) = (\lambda a_1)v_1 + \cdots + (\lambda a_n)v_n \in \operatorname{span}\{v_1, \ldots, v_n\}$, so $\operatorname{span}\{v_1, \ldots, v_n\}$ is closed under scalar multiplication.

Thus $\operatorname{span}\{v_1, \ldots, v_n\}$ is a subspace of V.

Each v_i is a linear combination of $v_1, \ldots, v_n$:

$$v_i = 0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_n.$$

Thus $v_i \in \operatorname{span}\{v_1, \ldots, v_n\}$. Conversely, since subspaces are closed under scalar multiplication and addition, every subspace of V containing each v_i contains $\operatorname{span}\{v_1, \ldots, v_n\}$.

Hence $\operatorname{span}\{v_1, \ldots, v_n\}$ is the smallest subspace of V containing $v_1, \ldots, v_n$. □

Definition 4.4 (Finite-dimensional vector space)**.** V is ***finite-dimensional*** if there exists $v_1, \ldots, v_n$ that spans V; otherwise, it is infinite-dimensional.

> **Example 4.5**
> $\mathbf{F}^3$ is finite-dimensional because $\mathbf{F}^3 = \operatorname{span}\{(1,0,0), (0,1,0), (0,0,1)\}$; $\mathbf{F}^\infty$ is infinite-dimensional.

Otherwise mentioned, all subsequent vector spaces are finite-dimensional.

Definition 4.6 (Polynomial)**.** A function $p : \mathbf{F} \to \mathbf{F}$ is a ***polynomial*** with coefficients in $\mathbf{F}$ if there exist $a_i \in \mathbf{F}$ such that

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n$$

for all $z \in \mathbf{F}$.

We denote the set of all polynomials with coefficients in $\mathbf{F}$ by $\mathcal{P}(\mathbf{F})$.

A polynomial $p \in \mathcal{P}(\mathbf{F})$ is has degree n if there exist scalars $a_0, a_1, \ldots, a_n \in \mathbf{F}$ with $a_n \neq 0$ such that $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ for all $z \in \mathbf{F}$; if p has degree n, we write $\deg p = n$.

For non-negative integer n, $\mathcal{P}^n(\mathbf{F})$ denotes the set of all polynomials with coefficients in $\mathbf{F}$ and degree at most n.

Definition 4.7 (Linear independence)**.** $\{v_1, \ldots, v_n\}$ is ***linearly independent*** in V if the only choice of $a_1, \ldots, a_n \in \mathbf{F}$ that makes $a_1 v_1 + \cdots + a_n v_n = \mathbf{0}$ is $a_1 = \cdots = a_n = 0$; otherwise, it is ***linearly dependent***.

Lemma 4.8 will often be useful; it states that given a linearly dependent list of vectors, one of the vectors is in the span of the previous ones and furthermore we can throw out that vector without changing the span of the original list.

Lemma 4.8 (Linear dependence lemma)**.** Suppose $\{v_1, \ldots, v_n\}$ is linearly dependent in V. Then there exists v_k such that the following hold:

(i) $v_k \in \operatorname{span}\{v_1, \ldots, v_{k-1}\}$

(ii) $\operatorname{span}\{v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n\} = \operatorname{span}\{v_1, \ldots, v_n\}$

Proof. Since $\{v_1, \ldots, v_n\}$ is linearly dependent, there exists $a_1, \ldots, a_n \in \mathbf{F}$, not all 0, such that

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Let $k = \max\{1, \ldots, n\}$ such that $a_k \neq 0$. Then

$$v_k = -\frac{a_1}{a_k} v_1 - \cdots - \frac{a_{k-1}}{a_k} v_{k-1},$$

proving (i).

To prove (ii), suppose $u \in \operatorname{span}\{v_1, \ldots, v_n\}$. Then there exists $c_1, \ldots, c_n \in \mathbf{F}$ such that

$$u = c_1 v_1 + \cdots + c_n v_n.$$

□

Proposition 4.9 says that no linearly independent list in V is longer than a spanning list in V.

Proposition 4.9. The length of every linearly independent list of vectors is less than or equal to the length of every spanning list of vectors.

Proof. Suppose $\{u_1, \ldots, u_m\}$ linearly independent in V, $\{w_1, \ldots, w_n\}$ spans V. We want to show $m \leq n$. We do so through the following steps:

Step 1

□

§4.2 *Bases*

Definition 4.10 (Basis)**.** $\{v_1 \ldots, v_n\}$ is a ***basis*** of V if it is linearly independent and spans V.

Example 4.11
Let $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ where the i-th coordinate is 1. $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis of $\mathbf{F}^n$, called the ***standard basis*** of $\mathbf{F}^n$.

Example 4.12
$\{1, z, \ldots, z^n\}$ is a basis of $\mathcal{P}^n(\mathbf{F})$.

Lemma 4.13 (Criterion for basis)**.** The following are equivalent:

(i) $\{v_1, \ldots, v_n\}$ is a basis of V.

(ii) Every $v \in V$ is uniquely expressed as a linear combination of $v_1, \ldots, v_n$.

(iii) $v_i \neq 0$, $V = Fv_1 \oplus \cdots \oplus Fv_n$.

Proof. □

Proposition 4.14. Every spanning list in a vector space can be reduced to a basis of the vector space.

Proposition 4.15. Every finite-dimensional vector space has a basis.

Proof. By definition, a finite-dimensional vector space has a spanning list. The previous result tells us that each spanning list can be reduced to a basis. □

Proposition 4.16. Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.

Proposition 4.17. Suppose U is a subspace of V. Then there exists a subspace W of V such that $V = U \oplus W$.

Proof. □

§4.3 *Dimension*

Definition 4.18 (Dimension)**.** The ***dimension*** of V is the length of any basis of V, denoted by $\dim V$.

Proposition 4.19. Suppose U is a subspace of V, then $\dim U \leq \dim V$.

Proposition 4.20. Every linearly independent list of vectors in V with length $\dim V$ is a basis of V.

Proposition 4.21. Every spanning list of vectors in V with length $\dim V$ is a basis of V.

Lemma 4.22 (Dimension of a sum)**.** Suppose U_1 and U_2 are subspaces of V, then

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

5 Linear Maps

§5.1 *Vector Space of Linear Maps*

Definition 5.1 (Linear map)**.** A ***linear map*** from V to W is a function $T : V \to W$ with the following properties:

(i) Additivity: $T(v + w) = Tv + Tw$ for all $v, w \in V$

(ii) Homogeneity: $T(\lambda v) = \lambda T(v)$ for all $\lambda \in \mathbf{F}$, $v \in V$

Notation. The set of all linear maps from V to W is denoted $\mathcal{L}(V, W)$; the set of linear transformations on V is denoted $\mathcal{L}(V)$.

Proposition 5.2 (Linear map lemma)**.** Suppose $\{v_1, \ldots, v_n\}$ is a basis of V and $w_1, \ldots, w_n \in W$. Then there exists a unique linear map $T : V \to W$ such that

$$Tv_i = w_i \quad (i = 1, \ldots, n)$$

Proof. First we show the existence of a linear map T with the desired property. Define $T : V \to W$ by

$$T(c_1v_1 + \cdots + c_nv_n) = c_1w_1 + \cdots + c_nw_n,$$

for some $c_i \in \mathbf{F}$. Since $\{v_1, \ldots, v_n\}$ is a basis of V, by Lemma 4.13, each $v \in V$ can be uniquely expressed as a linear combination of $v_1, \ldots, v_n$, thus the equation above does indeed define a function $T : V \to W$. For $i = 1, \ldots, n$, take $c_i = 1$ and the other c's equal to 0 to show that $Tv_i = w_i$.

We now show that $T : V \to W$ is a linear map:

(i) If $u, v \in V$ with $u = a_1v_1 + \cdots + a_nv_n$ and $c_1v_1 + \cdots + c_nv_n$, then

$$\begin{aligned} T(u + v) &= T\left((a_1 + c_1)v_1 + \cdots + (a_n + c_n)v_n\right) \\ &= (a_1 + c_1)w_1 + \cdots + (a_n + c_n)w_n \\ &= (a_1w_1 + \cdots + a_nw_n) + (c_1w_1 + \cdots + c_nw_n) \\ &= Tu + Tv \end{aligned}$$

so T satisfies additivity.

(ii) If $\lambda \in \mathbf{F}$ and $v = c_1v_1 + \cdots + c_nv_n$, then

$$\begin{aligned} T(\lambda v) &= T(\lambda c_1v_1 + \cdots + \lambda c_nv_n) \\ &= \lambda c_1w_1 + \cdots + \lambda c_nw_n \\ &= \lambda(c_1w_1 + \cdots + c_nw_n) \\ &= \lambda Tv \end{aligned}$$

so T satisfies homogeneity.

To prove uniqueness, now suppose that $T \in \mathcal{L}(V,W)$ and $Tv_i = w_i$ for $i = 1,\dots,n$. Let $c_i \in \mathbf{F}$. The homogeneity of T implies that $T(c_i v_i) = c_i w_i$. The additivity of T now implies that

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n.$$

Thus T is uniquely determined on $\operatorname{span}\{v_1,\dots,v_n\}$. Since $\{v_1,\dots,v_n\}$ is a basis of V, this implies that T is uniquely determined on V. □

Proposition 5.3. $\mathcal{L}(V,W)$ is a vector space, with the operations addition and scalar multiplication defined as follows: suppose $S,T \in \mathcal{L}(V,W)$, $\lambda \in \mathbf{F}$,

(i) $(S+T)(v) = Sv + Tv$

(ii) $(\lambda T)(v) = \lambda(Tv)$

for all $v \in V$.

Definition 5.4 (Product of linear maps)**.** $T \in \mathcal{L}(U,V)$, $S \in \mathcal{L}(V,W)$, then the ***product*** $ST \in \mathcal{L}(U,W)$ is defined by

$$(ST)(u) = S(Tu) \quad (\forall u \in U)$$

Remark. In other words, ST is just the usual composition $S \circ T$ of two functions.

Remark. ST is defined only when T maps into the domain of S.

Proposition 5.5 (Algebraic properties of products of linear maps)**.**

(i) Associativity: $(T_1T_2)T_3 = T_1(T_2T_3)$ for all linear maps T_1, T_2, T_3 such that the products make sense (meaning that T_3 maps into the domain of T_2, T_2 maps into the domain of T_1)

(ii) Iidentity: $TI = IT = T$ for all $T \in \mathcal{L}(V,W)$ (the first I is the identity map on V, and the second I is the identity map on W)

(iii) Distributive: $(S_1+S_2)T = S_1T + S_2T$ and $S(T_1+T_2) = ST_1 + ST_2$ for all $T, T_1, T_2 \in \mathcal{L}(U,V)$ and $S, S_1, S_2 \in \mathcal{L}(V,W)$

Proposition 5.6. $T \in \mathcal{L}(V,W)$. Then $T(0) = 0$.

Proof. By additivity, we have

$$T(0) = T(0+0) = T(0) + T(0).$$

Add the additive inverse of $T(0)$ to each side of the equation above to conclude that $T(0) = 0$. □

§5.2 *Kernel and Image*

Definition 5.7 (Kernel)**.** For $T \in \mathcal{L}(V,W)$, the ***kernel*** of T is the subset of V consisting of those vectors that T maps to 0:

$$\ker T \coloneqq \{v \in V \mid Tv = 0\}.$$

Proposition 5.8. $T \in \mathcal{L}(V,W)$, $\ker T$ is a subspace of V.

Proof. By Lemma 3.12, we check the conditions of a subspace:

(i) $T(0) = 0$ by Proposition 5.6, so $0 \in \ker T$.

(ii) For all $v, w \in \ker T$,

$$T(v+w) = Tv + Tw = 0 \implies v + w \in \ker T$$

so $\ker T$ is closed under addition.

(iii) For all $v \in \ker T$, $\lambda \in \mathbf{F}$,

$$T(\lambda v) = \lambda Tv = 0 \implies \lambda v \in \ker T$$

so $\ker T$ is closed under scalar multiplication.

□

Definition 5.9 (Injectivity)**.** $T : V \to W$ is ***injective*** if

$$Tu = Tv \implies u = v.$$

Proposition 5.10. $T \in \mathcal{L}(V, W)$, T is injective if and only if $\ker T = 0$.

Proof. □

Definition 5.11 (Image)**.** For $T : V \to W$, the ***image*** of T is the subset of W consisting of those vectors that are of the form Tv for some $v \in V$:

$$\operatorname{im} T \coloneqq \{Tv \mid v \in V\}.$$

Proposition 5.12. $T \in \mathcal{L}(V, W)$, $\operatorname{im} T$ is a subspace of W.

Proof. □

Definition 5.13 (Surjectivity)**.** $T : V \to W$ is ***surjective*** if $\operatorname{im} T = W$.

Theorem 5.14 (Fundamental Theorem of Linear Maps)**.** $T \in \mathcal{L}(V, W)$, then $\operatorname{im} T$ is finite-dimensional and

$$\dim V = \dim \ker T + \dim \operatorname{im} T.$$

§5.3 *Matrices*

Definition 5.15 (Matrix)**.** A $m \times n$ ***matrix*** A is a rectangular array with m rows and n columns:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

where $a_{ij} \in \mathbf{F}$.

Definition 5.16 (Matrix of a linear map)**.** $T \in \mathcal{L}(V, W)$, $\{v_1, \ldots, v_n\}$ is a basis of V, $\{w_1, \ldots, w_m\}$ is a basis of W. The ***matrix*** of T with respect to these bases of the $m \times n$ matrix $M(T)$ whose entries a_{ij} are defined by

§5.4 *Invertibility and Isomorphism*

Definition 5.17 (Invertibility)**.** $T \in \mathcal{L}(V, W)$ is ***invertible*** if there exists $S \in \mathcal{L}(W, V)$ such that $ST = I_V$, $TS = I_W$, where I_V and I_W are the ***identity maps*** on V and W respectively; S is known as the ***inverse*** of T.

Proposition 5.18 (Uniqueness of inverse)**.** An invertible linear map has a unique inverse.

Proof. Suppose $T \in \mathcal{L}(V, W)$ is invertible. S_1 and S_2 are inverses of T. Then

$$S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2.$$

Thus $S_1 = S_2$. □

Now that we know that the inverse is unique, we can give it a notation.

Part III

Abstract Algebra

6 Group Theory

§6.1 *Modular Arithmetic*

Let n be a fixed positive integer. Define a relation on $\mathbf{Z}$ by

$$a \sim b \iff n \mid (b-a).$$

Proposition 6.1. $a \sim b$ is a equivalence relation.

Proof.

(1) $a \sim a$, thus the relation is reflexive.

(2) $a \sim b \implies b \sim a$ for any integers a and b, thus the relation is symmetric.

(3) If $a \sim b$ and $b \sim c$ then $n \mid (a-b)$ and $n \mid (b-c)$, so $n \mid (a-b)+(b-c) = (a-c)$, so $a \sim c$ and the relation is transitive.

□

We write $a \equiv b \pmod{n}$ (read: a is ***congruent*** to b mod n) if $a \sim b$.

For any $k \in \mathbf{Z}$ we denote the equivalence class of a by $\overline{a}$, called the ***congruence class*** (residue class) of a mod n, consisting of the integers which differ from a by an integral multiple of n; that is,

$$\overline{a} = \{a + kn \mid k \in \mathbf{Z}\}.$$

There are precisely n distinct equivalence classes mod n, namely

$$\overline{0}, \overline{1}, \ldots, \overline{n-1}$$

determined by the possible remainders after division by n and these residue classes partition the integers $\mathbf{Z}$. The set of equivalence classes under this equivalence relation is denoted by $\mathbf{Z}/n\mathbf{Z}$, and called the ***integers modulo*** n.

We can define addition and multiplication for the elements of $\mathbf{Z}/n\mathbf{Z}$, defining ***modular arithmetic*** as follows: for $\overline{a}, \overline{b} \in \mathbf{Z}/n\mathbf{Z}$,

1. (addition) $\overline{a} + \overline{b} = \overline{a+b}$
2. (multiplication) $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$

This means that to compute the sum (product) of two elements $\overline{a}, \overline{b} \in \mathbf{Z}/n\mathbf{Z}$, take any ***representative*** integer $a \in \overline{a}$ and any representative integer $b \in \overline{b}$, and add (multiply) integers a and b as usual in $\mathbf{Z}$, then take the equivalence class containing the result.

Theorem 6.2. Addition and mulltiplication on $\mathbf{Z}/n\mathbf{Z}$ are well-defined; that is, they do not depend on the choices of representatives for the classes involved. More precisely, if $a_1, a_2 \in \mathbf{Z}$ and $b_1, b_2 \in \mathbf{Z}$ with $\overline{a_1} = \overline{b_1}$ and $\overline{a_2} = \overline{b_2}$, then $\overline{a_1 + a_2} = \overline{b_1 + b_2}$ and $\overline{a_1 a_2} = \overline{b_1 b_2}$, i.e., If

$$a_1 \equiv b_1 \pmod{n}, \quad a_2 \equiv b_2 \pmod{n}$$

then

$$a_1 + a_2 \equiv b_1 + b_2 \pmod{n}, \quad a_1 a_2 \equiv b_1 b_2 \pmod{n}.$$

Proof. Suppose $a_1 \equiv b_1 \pmod{n}$, i.e., $n \mid (a_1 - b_1)$. Then $a_1 = b_1 + sn$ for some integer s. Similarly, $a_2 \equiv b_2 \pmod{n}$ means $a_2 = b_2 + tn$ for some integer t.

Then $a_1 + a_2 = (b_1 + b_2) + (s + t)n$ so that $a_1 + a_2 \equiv b_1 + b_2 \pmod{n}$, which shows that the sum of the residue classes is independent of the representatives chosen.

Similarly, $a_1 a_2 = (b_1 + sn)(b_2 + tn) = b_1 b_2 + (b_1 t + b_2 s + stn)n$ shows that $a_1 a_2 \equiv b_1 b_2 \pmod{n}$ and so the product of the residue classes is also independent of the representatives chosen, completing the proof. □

An important subset of $\mathbf{Z}/n\mathbf{Z}$ consists of the collection of residue classes which have a multiplicative inverse in $\mathbf{Z}/n\mathbf{Z}$:

$$(\mathbf{Z}/n\mathbf{Z})^\times := \{\overline{a} \in \mathbf{Z}/n\mathbf{Z} \mid \exists \overline{c} \in \mathbf{Z}/n\mathbf{Z}, \overline{a} \cdot \overline{c} = \overline{1}\}.$$

Proposition 6.3. $(\mathbf{Z}/n\mathbf{Z})^\times$ is also the collection of residue classes whose representatives are relatively prime to n:

$$(\mathbf{Z}/n\mathbf{Z})^\times = \{\overline{a} \in \mathbf{Z}/n\mathbf{Z} \mid (a, n) = 1\}.$$

§6.2 *Group Axioms*

Definition 6.4. A ***binary operation*** $*$ on a set G is a function $* : G \times G \to G$. For any $a, b \in G$, we write $a * b$ for the image of (a, b) under $*$.

A binary operation $*$ on G is ***associative*** if, for any $a, b, c \in G$, $(a * b) * c = a * (b * c)$.

A binary operation $*$ on G is ***commutative*** if, for any $a, b \in G$, $a * b = b * a$.

Example 6.5

The following are examples of binary operations.

- $+$ (usual addition) is a commutative binary operation on $\mathbf{Z}$ (or on $\mathbf{Q}$, $\mathbf{R}$, or $\mathbf{C}$ respectively).
- $\times$ (usual multiplication) is a commutative binary operation on $\mathbf{Z}$ (or on $\mathbf{Q}$, $\mathbf{R}$, or $\mathbf{C}$ respectively).
- $-$ (usual subtraction) is a non-commutative binary operation on $\mathbf{Z}$.
- $-$ is not a binary operation on $\mathbf{Z}^+$ (nor $\mathbf{Q}^+$, $\mathbf{R}^+$) because for $a, b \in \mathbf{Z}^+$, with $a < b$, $a - b \notin \mathbf{Z}^+$; that is, $-$ does not map $\mathbf{Z}^+ \times \mathbf{Z}^+ \to \mathbf{Z}^+$.
- Taking the vector cross-product of two vectors in $\mathbf{R}^3$ is a binary operation which is not associative and not commutative.

Suppose that $*$ is a binary operation on G and $H \subseteq G$. If the restriction of $*$ to H is a binary operation on H, i.e. for all $a, b \in H$, $a * b \in H$, then H is said to be ***closed*** under $*$.

Remark. Observe that if $*$ is an associative (respectively, commutative) binary operation on G and $*$ is restricted to some $H \subseteq G$ is a binary operation on H, then $*$ is automatically associative (respectively, commutative) on H as well.

Definition 6.6 (Group)**.** A ***group*** is a pair $(G, *)$, where G is a set and $*$ is a binary operation on G satisfying the following group axioms:

(i) **(associativity)** for all $a, b, c \in G$, $a * (b * c) = (a * b) * c$.

(ii) **(identity)** there exists an identity element $e \in G$ such that for all $a \in G$, $a * e = e * a = a$.

(iii) **(invertibility)** for all $a \in G$, there exists a unique inverse $a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = e$.

G is ***abelian*** if the operation is commutative; it is ***non-abelian*** if otherwise.

Remark. The **closure** axiom (for all $a, b, c \in G$, $a * b \in G$) is implicitly implied, as a binary operation has to be closed under the set.

Notation. A group $(G, *)$ is usually simply denoted by G.

Notation. We abbreviate $a * b$ to just ab. Also, since the operation $*$ is associative, we can omit unnecessary parentheses: $(ab)c = a(bc) = abc$.

Notation. For any $a \in G$ and $n \in \mathbf{Z}^+$ we abbreviate $a^n = \underbrace{a \cdots a}_{n \text{ times}}$.

Example 6.7

- $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ are groups under $+$ with $e = 0$ and $a^{-1} = -a$ for all a.
- $\mathbf{Q} \setminus \{0\}$, $\mathbf{R} \setminus \{0\}$, $\mathbf{C} \setminus \{0\}$, $\mathbf{Q}^+$, $\mathbf{R}^+$ are groups under $\times$ with $e = 1$ and $a^{-1} = \frac{1}{a}$ for all a. Note however that $\mathbf{Z} \setminus \{0\}$ is not a group under $\times$ because the element 2 (for instance) does not have an inverse in $\mathbf{Z} \setminus \{0\}$.
- For $n \in \mathbf{Z}^+$, $\mathbf{Z}/n\mathbf{Z}$ is an abelian group under $+$.
- For $n \in \mathbf{Z}^+$, $(\mathbf{Z}/n\mathbf{Z})^\times$ is an abelian group under multiplication.

Definition 6.8 (Product group)**.** Let $(G, *_G)$ and $(H, *_H)$ be groups. Then the operation $*$ defined on $G \times H$ by

$$(g_1, h_1) * (g_2, h_2) = (g_1 *_G g_2, h_1 *_H h_2)$$

is a group operation. $(G \times H, *)$ is called the ***product group*** or the product of G and H.

Proof. As $*_G$ and $*_H$ are both associative binary operations then it follows easily from the definition to see that $*$ is also an associative binary operation on $G \times H$. We also note

$$e_{G \times H} = (e_G, e_H) \quad \text{and} \quad (g, h)^{-1} = (g^{-1}, h^{-1})$$

as for any $g \in G$, $h \in H$,

$$(e_G, e_H) * (g, h) = (g, h) = (g, h) * (e_G, e_H);$$
$$(g^{-1}, h^{-1}) * (g, h) = (e_G, e_H) = (g, h) * (g^{-1}, h^{-1}).$$

□

Proposition 6.9. Let G be a group. Then

(1) the identity of G is unique,

(2) for each $a \in G$, a^{-1} is unique,

(3) $(a^{-1})^{-1} = a$ for all $a \in G$,

(4) $(ab)^{-1} = b^{-1}a^{-1}$,

(5) for any $a_1, \ldots, a_n \in G$, $a_1 \cdots a_n$ is independent of how we arrange the parantheses (generalised associative law).

Proof.

(1) Let e_0 and e_1 both be identites, so $e_0 e_1 = e_0 = e_1$.

(2) Let c and c both be inverses to a and $e \in G$ the identity. Then $ab = e = ca$. Thus $c = ce = c(ab) = (ca)b = eb = b$.

(3) Clear.

(4) Let $c = (ab)^{-1}$ so that $(ab)c = e$, which gives $bc = a^{-1}$ and thus $c = b^{-1}a^{-1}$ by multiplying on the left.

(5) The result is trivial for $n = 1, 2, 3$. For all $k < n$ assume that any $a_1 \cdots a_k$ is independent of parantheses. Then
$$(a_1 \cdots a_n) = (a_1 \cdots a_k)(a_{k+1} \cdots a_n).$$
Then by assumption both are independent of parentheses since $k, n-k < n$ so by induction we are done.

□

Proposition 6.10 (Cancellation law)**.** Let $a, b \in G$. Then the equations $ax = b$ and $ya = b$ have unique solutions for $x, y \in G$. In particular, we can cancel on the left and right.

Proof. That $x = a^{-1}b$ is unique follows from the uniqueness of a^{-1} and the same for $y = ba^{-1}$. □

Definition 6.11 (Order)**.** For a group G and $x \in G$, the order of x is the smallest positive integer n such that $x^n = 1$, and denote this integer by $|x|$; in this case x is said to be of order n.

If no positive power of x is the identity, the order of x is defined to be infinity, and x is said to be of infinite order.

Example 6.12

- An element of a group has order 1 if and only if it is the identity.
- In the additive groups $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, every non-zero (i.e. non-identity) element has infinite order.
- In the multiplicative groups $\mathbf{R} \setminus \{0\}$ or $\mathbf{Q} \setminus \{0\}$, the element -1 has order 2 and all other non-identity elements have infinite order.
- In $\mathbf{Z}/9\mathbf{Z}$, the element $\overline{6}$ has order 3. (Recall that in an additive group, the powers of an element are integer multiples of the element.)
- In $(\mathbf{Z}/7\mathbf{Z})^\times$, the powers of the element $\overline{2}$ are $\overline{2}, \overline{4}, \overline{8} = \overline{1}$, the identity in this group, so 2 has order 3. Similarly, the element $\overline{3}$ has order 6, since 3^6 is the smallest positive power of 3 that is congruent to 1 mod 7.

Definition 6.13 (Group table)**.** Let $G = \{g_1, \ldots, g_n\}$ be a finite group with $g_1 = 1$. The ***group table*** (multiplication table) of G is the $n \times n$ matrix whose (i, j)-entry is the group element $g_i g_j$.

For a finite group the multiplication table contains, in some sense, all the information about the group.

§6.3 *Examples of Groups*

An important family of groups is the dihedral groups.

Definition 6.14 (Dihedral group)**.** For $n \in \mathbf{Z}^+$, $n \geq 3$, let D_{2n} be the set of symmetries of a regular n-gon.

Remark. Here "D" stands for "dihedral", meaning two-sided.

Let r be a rotation clockwise about the origin by $2\pi/n$ radians, let s be a reflection about the line of symmetry through the first labelled vertex and the origin.

Proposition 6.15.

(1) $|r| = n$

(2) $|s| = 2$

(3) $s \neq r^i$ for all i

(4) $sr^i \neq sr^j$ for all $i \neq j$. Thus

$$D_{2n} = \{1, r, \ldots, r^{n-1}, s, sr, \ldots, sr^{n-1}\}$$

and we see that $|D_{2n}| = 2n$.

(5) $rs = sr^{-1}$

(6) $r^i s = sr^{-i}$

Proof.

(1) This is clear.

(2) So is this.

(3) And this.

(4) Just cancel on the left and use the fact that $|r| = n$. We assume that $i \not\equiv j \pmod{n}$.

(5) Omitted.

(6) By (5), this is true for $i = 1$. Assume it holds for $k < n$. Then $r^{k+1}s = r(r^k s) = rsr^{-k}$. Then $rs = sr^{-1}$ so $rsr^{-k} = sr^{-1}r^{-k} = sr^{-k-1}$ so we are done.

□

A presentation for the dihedral group with $2n$ elements is

$$D_{2n} = \{r, s \mid r^n = s^2 = 1, rs = sr^{-1}\}.$$

permutation group, subgroup, order of group, homomorphism and isomorphism

An important (if rather elementary) family of groups is the cyclic groups.

Definition 6.16 (Cyclic group)**.** A group G is called ***cyclic*** if there exists $g \in G$ such that

$$G = \{g^k \mid k \in \mathbf{Z}\}.$$

Such a g is called a ***generator***.

As $g^i g^j = g^{i+j} = g^j g^i$ then cyclic groups are abelian.

Example 6.17
$\mathbf{Z}$ is cyclic and has generators 1 and -1.

Example 6.18
Let $n \geq 1$. The n-th cyclic group C_n is the group with elements

$$e, g, g_2, \ldots, g^{n-1}$$

which satisfy $g^n = e$. So given two elements in C_n we define

$$g_i g_j = \begin{cases} g^{i+j} & \text{if } 0 \leq i + j < n, \\ g^{i+j-n} & \text{if } n \leq i + j \leq 2n - 2. \end{cases}$$

Definition 6.19 (Subgroup)**.** Let G be a group. We say that a subset $H \subseteq G$ is a ***subgroup*** of G if the group operation $*$ restricts to make a group of H. That is H is a subgroup of G if:

(i) $e \in H$;

(ii) whenever $g_1, g_2 \in H$ then $g_1 g_2 \in H$.

(iii) whenever $g \in H$ then $g^{-1} \in H$.

Remark. Note that there is no need to require that associativity holds for products of elements in H as this follows from the associativity of products in G.

Example 6.20

The set of even integers is a subgroup of $\mathbf{Z}$; the set of odd integers is not a subgroup of $\mathbf{Z}$ because it does not even form a group, since it does not satisfy the closure axiom.

Definition 6.21 (Isomorphism)**.** An ***isomorphism*** $\phi : G \to H$ between two groups $(G, *_G)$ and $(H, *_H)$ is a bijection such that for any $g_1, g_2 \in G$ we have

$$\phi(g_1 *_G g_2) = \phi(g_1) *_H \phi(g_2).$$

Two groups are said to be ***isomorphic*** if there is an isomorphism between them, denoted by $G \cong H$.

Example 6.22 ($\mathbf{Z} \cong 10\mathbf{Z}$)

Consider the two groups

$$\mathbf{Z} = (\{\dots, -2, -1, 0, 1, 2, \dots\}, +)$$

and

$$10\mathbf{Z} = (\{\dots, -20, -10, 0, 10, 20, \dots\}, +).$$

These groups are "different", but only superficially so — you might even say they only differ in the names of the elements.

Formally, the map

$$\phi : \mathbf{Z} \to 10\mathbf{Z} \text{ by } x \mapsto 10x$$

is a bijection of the underlying sets which respects the group operation. In symbols,

$$\phi(x + y) = \phi(x) + \phi(y).$$

In other words, ϕ is a way of re-assigning names of the elements without changing the structure of the group.

§6.4 *Permutation Groups*

§6.5 *More on Subgroups & Cyclic Groups*

§6.6 *Lagrange's Theorem*

Definition 6.23 (Coset)**.** Let H be a subgroup of G.

Then the ***left cosets*** of H (or left H-cosets) are the sets

$$gH = \{gh \mid h \in H\}.$$

The ***right cosets*** of H (or right H-cosets) are the sets

$$Hg = \{hg \mid h \in H\}.$$

Two (left) cosets aH and bH are either disjoint or equal.

Since multiplication is injective, the cosets of H are the same size as H, and thus H partitions G into equal-sized parts.

Notation. We write G/H for the set of (left) cosets of H in G. The cardinality of G/H is called the ***index*** of H in G.

An important result relating the order of a group with the orders of its subgroups is Lagrange's theorem.

Theorem 6.24 (Lagrange's theorem)**.** If G is a finite group and H is a subgroup of G, then $|H|$ divides $|G|$.

Groups of small order (up to order 8). Quaternions. Fermat–Euler theorem from the group-theoretic point of view.

Theorem 6.25 (Fermat's Little Theorem)**.** For every finite group G, for all $a \in G$, $a^{|G|} = e$.

Proof. Consider the subgroup H generated by a: $H = \{a^i \mid i \in \mathbf{Z}\}$. Since G is finite, the infinite sequence $a^0 = e, a^1, a^2, a^3, \ldots$ must repeat, say $a^i = a^j, i < j$. Let $k = j - i$. Multiplying both sides by $a^{-i} = (a^{-1})^i$, we get $a^{j-i} = a^k = e$. Suppose k is the least positive integer for which this holds. Then $H = \{a_0, a_1, a_2, \ldots, a^{k-1}\}$, and thus $|H| = k$. By Lagrange's Theorem, k divides $|G|$, so $a^{|G|} = (a^k)^{\frac{|G|}{k}} = e$. □

Part IV

Real Analysis

7 Number Systems

§7.1 *Natural Numbers*

In Peano's development, it is assumed that there is a set $\mathbf{N}$ (the natural numbers) of undefined objects with a distinguished element 1 such that

(i) 1 is a natural number; that is $1 \in \mathbf{N}$;

(ii) every $n \in \mathbf{N}$ has a successor $S(n) \in \mathbf{N}$;

(iii) for every n, $S(n) \neq 1$ (there is no number with 1 as successor)

(iv) if $S(n) = S(m)$, then $n = m$;

(v) if A is a set of natural numbers such that $1 \in A$ and $n \in A \implies S(n) \in A$, then A contains all natural numbers.

These are known as ***Peano's axioms***.

Theorem 7.1 (Archimedean property of $\mathbf{N}$)**.** $\mathbf{N}$ is not bounded above.

Proof. Suppose, for a contradiction, that $\mathbf{N}$ is bounded above. Then $\mathbf{N}$ is non-empty and bounded above, so by completeness (of $\mathbf{R}$) $\mathbf{N}$ has a supremum.

By the Approximation property with $\epsilon = \frac{1}{2}$, there is a natural number $n \in \mathbf{N}$ such that $\sup \mathbf{N} - \frac{1}{2} < n \leq \sup \mathbf{N}$.

Now $n + 1 \in \mathbf{N}$ and $n + 1 > \sup \mathbf{N}$. This is a contradiction. □

§7.2 *Integers*

Definition 7.2. For $(a, b), (c, d) \in \mathbf{N} \times \mathbf{N}$, we define a relation

$$(a, b) \sim (c, d) \iff a + d = b + c.$$

Proposition 7.3. $\sim$ is an equivalence relation on $\mathbf{N} \times \mathbf{N}$.

Proof. Suppose $(a, b), (c, d), (e, f) \in \mathbf{N} \times \mathbf{N}$.

(i) $\sim$ is reflexive: $(a, b) \sim (a, b)$ because $a + b = b + a$ in $\mathbf{N}$, by commutativity in $\mathbf{N}$.

(ii) $\sim$ is symmetric: If $(a, b) \sim (c, d)$, then $(c, d) \sim (a, b)$ because if $a + d = b + c$, then $c + b = d + a$ in $\mathbf{N}$.

(iii) $\sim$ is transitive:

□

§7.3 *Rational Numbers*

Notation. $\mathbf{Z}' = \mathbf{Z} \setminus \{0\}$.

Definition 7.4. Let $\sim$ be the binary relation defined on $\mathbf{Z} \times \mathbf{Z}'$ by

$$(a, b) \sim (c, d) \iff ad = bc.$$

Proposition 7.5. $\sim$ is an equivalence on $\mathbf{Z} \times \mathbf{Z}'$.

Proof. We just check that $\sim$ is transitive. So suppose that $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then

$$ad = bc \tag{1}$$

$$cf = de \tag{2}$$

Multiplying (1) by f and (2) by b, we obtain

$$adf = bcf \tag{3}$$

$$bcf = bde \tag{4}$$

Hence $adf = bde$. Since $d \neq 0$, the Cancellation Law implies that $af = bc$. Hence $(a, b) \sim (e, f)$. □

Definition 7.6. The set of ***rational numbers*** is defined by

$$\mathbf{Q} := \mathbf{Z} \times \mathbf{Z}' / \sim$$

i.e. $\mathbf{Q}$ is the set of $\sim$ equivalence classes.

Notation. For each $(a, b) \in \mathbf{Z} \times \mathbf{Z}'$, the corresponding equivalence class is denoted by $[(a, b)]$.

We define addition $+_{\mathbf{Q}}$ and multiplication $\cdot_{\mathbf{Q}}$ on $\mathbf{Q}$ as follows:

$$[(a, b)] +_{\mathbf{Q}} [(c, d)] = [(ad + bc, bd)].$$

$$[(a, b)] \cdot_{\mathbf{Q}} [(c, d)] = [(ac, bd)].$$

Proposition 7.7. $+_{\mathbf{Q}}$ and $\cdot_{\mathbf{Q}}$ are well-defined.

Lemma 7.8. $\mathbf{Q}$ is a field, with addition and multiplication as defined above.

Proof. We check the field axioms.

(i) commutativity of addition

(ii) associativity of addition

(iii) Let $0_{\mathbf{Q}} = [(0, 1)]$. We now show that $0_{\mathbf{Q}}$ is an additive identity.
Let $q = [(a, b)]$. Then

$$\begin{aligned} q +_{\mathbf{Q}} 0_{\mathbf{Q}} &= [(a, b)] +_{\mathbf{Q}} [(0, 1)] \\ &= [(a \cdot 1 + 0 \cdot b, b \cdot 1)] \\ &= [(a, b)] \\ &= q. \end{aligned}$$

Since for any $q \in \mathbf{Q}$, $q +_{\mathbf{Q}} 0_{\mathbf{Q}} = q$, thus $0_{\mathbf{Q}}$ is an additive identity. Hence an additive identity exists.

(iv) Consider $r = [(-a, b)]$. Then

$$\begin{aligned} q +_{\mathbf{Q}} r &= [(a, b)] +_{\mathbf{Q}} [(-a, b)] \\ &= [(ab + (-a)b, b^2)] \\ &= [(0, b^2)] \end{aligned}$$

Since $0 \cdot 1 = 0 \cdot b^2$, we have $(0, b^2) = (0, 1)$. Hence

$$\begin{aligned} q +_{\mathbf{Q}} r &= [(0, b^2)] \\ &= [(0, 1)] \\ &= 0_{\mathbf{Q}} \end{aligned}$$

Since for any $q \in \mathbf{Q}$, there exists a unique $r \in \mathbf{Q}$ such that $q +_{\mathbf{Q}} r = 0_{\mathbf{Q}}$, hence the additive inverse exists.

(v) commutativity of multiplication

We want to show that for all $q, r \in \mathbf{Q}$, $q \cdot_{\mathbf{Q}} r = r \cdot_{\mathbf{Q}} q$.

(vi) associativity of multiplication

We want to show that for all $q, r \in \mathbf{Q}$, $(q \cdot_{\mathbf{Q}} r) \cdot_{\mathbf{Q}} s = q \cdot_{\mathbf{Q}} (r \cdot_{\mathbf{Q}} s)$.

(vii) distributivity

We want to show that for all $q, r, s \in \mathbf{Q}$, $q \cdot_{\mathbf{Q}} (r +_{\mathbf{Q}} s) = (q \cdot_{\mathbf{Q}} r) +_{\mathbf{Q}} (q \cdot_{\mathbf{Q}} s)$.

(viii) Let $1_{\mathbf{Q}} = [(1, 1)]$. We now show that $1_{\mathbf{Q}}$ is a multiplicative identity.
Let $q = [(a, b)]$. Then

$$\begin{aligned} q \cdot_{\mathbf{Q}} 1_{\mathbf{Q}} &= [(a, b)] \cdot_{\mathbf{Q}} [(1, 1)] \\ &= [(a \cdot 1, b \cdot 1)] \\ &= [(a, b)] \\ &= 1_{\mathbf{Q}} \end{aligned}$$

Since for all $q \in \mathbf{Q}$, $q \cdot_{\mathbf{Q}} 1_{\mathbf{Q}} = q$, $1_{\mathbf{Q}}$ is a multiplicative identity. Hence a multiplicative identity exists.

(ix) Suppose that $q = [(a, b)] \neq [(0, 1)]$. Then $a \neq 0$ and so $(b, a) \in \mathbf{Z} \times \mathbf{Z}'$. Let $r = [(b, a)]$. Then

$$\begin{aligned} q \cdot_{\mathbf{Q}} r &= [(a, b)] \cdot_{\mathbf{Q}} [(b, a)] \\ &= [(ab, ba)] \\ &= [(1, 1)] \\ &= 1_{\mathbf{Q}}. \end{aligned}$$

Since for every $0_{\mathbf{Q}} \neq q \in \mathbf{Q}$, there exists a unique $r \in \mathbf{Q}$ such that $q \cdot_{\mathbf{Q}} r = 1_{\mathbf{Q}}$, thus r is a multiplicative inverse. Hence a multiplicative inverse exists. □

Since $\mathbf{Q}$ is a field, we have the following results:

(1) The additive identity in $\mathbf{Q}$ is unique.

(2) The additive inverse of an element of $\mathbf{Q}$ is unique.

(3) The multiplicative identity of $\mathbf{Q}$ is unique.

(4) The multiplicative inverse of a nonzero element of $\mathbf{Q}$ is unique.

Notation. Since the additive inverse is unique, we denote the additive inverse of $q \in \mathbf{Q}$ by $-q$; we define the binary operation $-_{\mathbf{Q}}$ on $\mathbf{Q}$ by

$$q -_{\mathbf{Q}} r = q +_{\mathbf{Q}} (-r).$$

Notation. Since the multiplicative inverse is unique, we denote the additive inverse of $q \in \mathbf{Q}$ by q^{-1}.

Finally we want to define an order relation on $\mathbf{Q}$.

Definition 7.9 (Order on $\mathbf{Q}$)**.** Suppose that $r, s \in \mathbf{Q}$ and that $r = [(a, b)]$ and $s = [(c, d)]$, where $b, d > 0$. Then

$$r \leq_{\mathbf{Q}} s \iff ad < bc.$$

Proposition 7.10. $<_{\mathbf{Q}}$ is well-defined.

Definition 7.11. If $q \in \mathbf{Q}$, then

- q is ***positive*** if and only if $0_{\mathbf{Q}} <_{\mathbf{Q}} q$,
- q is ***negative*** if and only if $q <_{\mathbf{Q}} 0_{\mathbf{Q}}$.

Definition 7.12. If $q \in \mathbf{Q}$, then the ***absolute value*** of q is

$$|q| = \begin{cases} -q & \text{if } q \text{ is negative,} \\ q & \text{if otherwise.} \end{cases}$$

§7.4 *Real Numbers*

One method to construct $\mathbf{R}$ from $\mathbf{Q}$ is Dedekind cuts.

Definition 7.13 (Dedekind cut)**.** A ***Dedekind cut*** $\alpha \subset \mathbf{Q}$ satisfies the following properties:

(i) $\alpha \neq \emptyset$, $\alpha \neq \mathbf{Q}$;

(ii) if $p \in \alpha$, $q \in \mathbf{Q}$ and $q < p$, then $q \in \alpha$;

(iii) if $p \in \alpha$, then $p < r$ for some $r \in \alpha$.

Note that (iii) simply says that α has no largest member; (ii) implies two facts which will be used freely:

- If $p \in \alpha$ and $q \notin \alpha$ then $p < q$.
- If $r \notin \alpha$ and $r < s$ then $s \notin \alpha$.

Example 7.14
Let $r \in \mathbf{Q}$ and define

$$\alpha_r := \{p \in \mathbf{Q} \mid p < r\}.$$

We now check that this is indeed a Dedekind cut.

(1) $p = 1 + r \notin \alpha_r$ thus $\alpha_r \neq \mathbf{Q}$. $p = r - 1 \in \alpha_r$ thus $\alpha_r \neq \emptyset$.

(2) Suppose that $q \in \alpha_r$ and $q' < q$. Then $q' < q < r$ which implies that $q' < r$ thus $q' \in \alpha_r$.

(3) Suppose that $q \in \alpha_r$. Consider $\dfrac{q+r}{2} \in \mathbf{Q}$ and $q < \dfrac{q+r}{2} < r$. Thus $\dfrac{q+r}{2} \in \alpha_r$.

This example shows that every rational r corresponds to a Dedekind cut α_r.

Example 7.15
$\sqrt[3]{2}$ is not rational, but it is real. $\sqrt[3]{2}$ corresponds to the cut

$$\alpha = \{p \in \mathbf{Q} \mid p^3 < 2\}.$$

(1) Trivial.

(2) If $q < p$, by the monotonicity of the cubic function, this implies that $q^3 < p^3 < 2$ thus $q \in \alpha$.

(3) If $p \in \alpha$, consider $\left(p + \frac{1}{n}\right)^3 < 2$.

Definition 7.16. The set of real numbers, denoted by $\mathbf{R}$, is the set of all Dedekind cuts.

$$\mathbf{R} := \{\alpha \mid \alpha \text{ is a Dedekind cut}\}$$

Proposition 7.17. $\mathbf{R}$ has an order.

Proof. We define $\alpha < \beta$ to mean that $\alpha \subset \beta$. Let us check if this is an order (check for transitivity and trichotomy).

(1) For $\alpha, \beta, \gamma \in \mathbf{R}$, if $\alpha < \beta$ and $\beta < \gamma$ it is clear that $\alpha < \gamma$. (A proper subset of a proper subset is a proper subset.)

(2) It is clear that at most one of the three relations

$$\alpha < \beta, \quad \alpha = \beta, \quad \beta < \alpha$$

can hold for any pair α, β.

To show that at least one holds, assume that the first two fail. Then α is not a subset of β. Hence there exists some $p \in \alpha$ with $p \in \beta$.

If $q \in \beta$, it follows that $q < p$ (since $p \notin \beta$), hence $q \in \alpha$, by (ii). Thus $\beta \subset \alpha$. Since $\beta \neq \alpha$, we conclude that $\beta < \alpha$.

Thus $\mathbf{R}$ is an ordered set. □

Proposition 7.18. The ordered set $\mathbf{R}$ has the least-upper-bound property.

Proof. Let $A \neq \emptyset$, $A \subset \mathbf{R}$. Assume that $\beta \in \mathbf{R}$ is an upper bound of A.

Define β to be the union of all $\alpha \in A$; in other words, $p \in \gamma$ if and only if $p \in \alpha$ for some $\alpha \in A$. We shall prove that $\gamma \in \mathbf{R}$ by checking the definition of Dedekind cuts:

(1) Since A is not empty, there exists an $\alpha_0 \in A$. This α_0 is not empty. Since $\alpha_0 \subset \gamma$, γ is not empty. Next, $\gamma \subset \beta$ (since $\alpha \subset \beta$ for every $\alpha \in A$), and therefore $\gamma \neq \mathbf{Q}$.

(2) Pick $p \in \gamma$. Then $p \in \alpha_1$ for some $\alpha_1 \in A$. If $q < p$, then $q \in \alpha_1$, hence $q \in \gamma$.

(3) If $r \in \alpha_1$ is so chosen that $r > p$, we see that $r \in \gamma$ (since $\alpha_1 \subset \gamma$).

Next we prove that $\gamma = \sup A$.

(1) It is clear that $\alpha \leq \gamma$ for every $\alpha \in A$.

(2) Suppose $\delta < \gamma$. Then there is an $s \in \gamma$ and that $s \notin \delta$. Since $s \in \gamma$, $s \in \alpha$ for some $\alpha \in A$. Hence $\delta < \alpha$, and δ is not an upper bound of A.

□

Definition 7.19. Given $\alpha, \beta \in \mathbf{R}$. Define

$$\alpha + \beta \coloneqq \{r \in \mathbf{Q} \mid r = a + b, a \in \alpha, b \in \beta\}.$$

Proposition 7.20 (Addition on $\mathbf{R}$ is closed)**.** $\alpha + \beta \in \mathbf{R}$.

Proof.

(1) $\alpha \neq \emptyset$ and $\beta \neq \emptyset$ implies there exists $a \in \alpha$ and $b \in \beta$. Hence $r = a + b \in \alpha + \beta$ so $\alpha + \beta \neq \emptyset$.

Since $\alpha \neq \mathbf{Q}$ and $\beta \neq \mathbf{Q}$, there is $c \neq \alpha$ and $d \neq \beta$. $r' = c + d > a + b$ for any $a \in \alpha, b \in \beta$, so $r' \notin \alpha + \beta$. Hence $\alpha + \beta \neq \mathbf{Q}$.

(2) Suppose that $r \in \alpha + \beta$ and $r' < r$. We want to show that $r' \in \alpha + \beta$.

$r = a + b$ for some $a \in \alpha, b \in \beta$. $r' - a < b$. Since $\beta \in \mathbf{R}$, $r' - a \in \beta$ so $r' - a = b_1$ for some $b_1 \in \beta$. Hence $r' = a + b_1 \in \alpha + \beta$.

(3) Suppose $r \in \alpha + \beta$, so $r = a + b$ for some $a \in \alpha, b \in \beta$. There exists $a' \in \alpha, b' \in \beta$ with $a < a'$ and $b < b'$. Then $r = a + b < a' + b' \in \alpha + \beta$. We define $r' = a' + b' \in \alpha + \beta$ with $r < r'$.

□

We now prove that the set of real numbers satisfies the commutative, associative, and identity field axioms with respect to addition.

Proposition 7.21 (Addition on $\mathbf{R}$ is commutative)**.** $\forall \alpha, \beta \in \mathbf{R}$, $\alpha + \beta = \beta + \alpha$.

Proof. We need to show that $\alpha + \beta \subseteq \beta + \alpha$ and $\beta + \alpha \subseteq \alpha + \beta$.

Let $r \in \alpha + \beta$. Then $r = a + b$ for $a \in \alpha$ and $b \in \beta$. Thus $r = b + a$ since $+$ is commutative on $\mathbf{Q}$. Hence $r \in \beta + \alpha$. Therefore $\alpha + \beta \subseteq \beta + \alpha$.

Similarly, $\beta + \alpha \subseteq \alpha + \beta$.

Therefore $\alpha + \beta = \beta + \alpha$. □

Proposition 7.22 (Addition on $\mathbf{R}$ is associative)**.** $\forall \alpha, \beta, \gamma \in \mathbf{R}$, $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.

Proof. Let $r \in \alpha + (\beta + \gamma)$. Then $r = a + (b + c)$ where $a \in \alpha, b \in \beta, c \in \gamma$. Thus $r = (a + b) + c$ by associativity of $+$ on $\mathbf{Q}$. Therefore $r \in (\alpha + \beta) + \gamma$, hence $\alpha + (\beta + \gamma) \subseteq (\alpha + \beta) + \gamma$.

Similarly, $(\alpha + \beta) + \gamma \subseteq \alpha + (\beta + \gamma)$. □

Proposition 7.23. Define $0^* \coloneqq \{p \in \mathbf{Q} \mid p < 0\}$. Then $\alpha + 0^* = \alpha$.

Proof. Let $r \in \alpha + 0^*$. Then $r = a + p$ for some $a \in \alpha, p \in 0^*$. Thus $r = a + p < a + 0 = a$ by ordering on $\mathbf{Q}$ and identity on $\mathbf{Q}$. Hence $\alpha + 0^* \subseteq \alpha$.

Let $r \in \alpha$. Then there exists $r' > p$ where $r' \in \alpha$. Thus $r - r' < 0$, so $r - r' \in 0^*$. We see that

$$r = \underbrace{r'}_{\in \alpha} + \underbrace{(r - r')}_{\in 0^*}.$$

Hence $\alpha \subseteq \alpha + 0*$. □

Exercise 22

Express $-\alpha$ in terms of α; show

$$\alpha + (-\alpha) = 0 = (-\alpha) + \alpha$$

Proof. We split this into two cases.

Case 1: α is a rational number, then $\alpha = (A, B)$ where $A = \{x \mid x < \alpha\}$, $B = \{x \mid x \geq \alpha\}$.

Let $-\alpha = (A', B')$, where $A' = \{x \mid x < -\alpha\}$, $B' = \{x \mid x \geq -\alpha\}$. We see that $\alpha + (-\alpha) \leq 0$ is obvious.

On the other hand, since $0 = (O, O')$, for any $\epsilon < 0$ we have

$$\epsilon = \left(\alpha + \frac{\epsilon}{2}\right) + \left(-\alpha + \frac{\epsilon}{2}\right) \in A + A'$$

Hence $\alpha + (-\alpha) = 0$.

Case 2: α is irrational, let $\alpha = (A, B)$ where B does not have a lowest value. Then $-B = \{-x \mid x \in B\}$ does not have a highest value.

We wish to define $-\alpha = (-B, -A)$, but first we need to show that this is well-defined by checking through all the conditions.

- Property 1: This is trivial.
- Property 2: Prove that $-A$ and B are disjoint.

  Note that $\forall x \in \mathbf{R}$, if $x = -y$, then exactly one out of $y \in A$ and $y \in B$ is true $\implies$ exactly one out of $x \in -B$ and $x \in -A$ is true.
- Property 3: Prove $-B$ is closed downwards.

  Suppose otherwise, that $x < y, y \in -B$ but $x \notin -B$. Then $-y \in B$, $-x \notin B$. Since A is the complement of B, $-y \notin A$, $-x \in A$. But $-y < -x$, which is a contradiction.
- Property 4 is already guaranteed by the irrationality of α.

All of these properties imply that the real numbers form a commutative group by addition. □

Negation

Given any set $X \subset \mathbf{R}$, let $-X$ denote the set of the negatives of those rational numbers. That is $x \in X$ if and only if $-x \in -X$.

If (A, B) is a Dedekind cut, then $-(A, B)$ is defined to be $(-B, -A)$.

This is pretty clearly a Dedekind cut. - proof

Signs

A Dedekind cut (A, B) is **positive** if $0 \in A$ and **negative** if $0 \in B$. If (A, B) is neither positive nor negative, then (A, B) is the cut representing 0.

If (A, B) is positive, then $-(A, B)$ is negative. Likewise, if (A, B) is negative, then $-(A, B)$ is positive. The cut (A, B) is non-negative if it is either positive or 0.

Definition 7.24 (Multiplication on $\mathbf{R}$)**.** Given $\alpha, \beta \in \mathbf{R}$, $\alpha, \beta > 0^*$. Define

$$\alpha \cdot \beta \coloneqq \{p \in \mathbf{Q} \mid p \leq ab, a \in \alpha, b \in \beta, a, b > 0\}.$$

Proposition 7.25 (Multiplication on $\mathbf{R}$ is closed)**.** $\alpha \cdot \beta \in \mathbf{R}$.

Proof.

(1) $\alpha \neq \emptyset$ means there exists $a \in \alpha, a > 0$. Similarly, $\beta \neq \emptyset$ means there exists $b \in \beta, b > 0$. Then $a \cdot b \in \mathbf{Q}$ and $ab \leq ab$, so $ab \in \alpha \cdot \beta \neq \emptyset$.

$\alpha \neq \mathbf{Q}$ means there exists $a' \notin \alpha, a' > a$ for all $a \in \alpha$. $\beta \neq \mathbf{Q}$ means there exists $b' \in \beta, b' > b$ for all $b \in \beta$. Then $a'b' > ab$ for all $a \in \alpha, b \in \beta$, so $a'b' \notin \alpha \cdot \beta$, thus $\alpha \cdot \beta \neq \mathbf{Q}$.

(2) $p < \alpha \cdot \beta$ means $p \leq a \cdot b$ for some $a \in \alpha, b \in \beta, a, b > 0$.

For $q < p$, $q < p \leq a \cdot b$ so $q \in \alpha \cdot \beta$.

(3) $p \in \alpha \cdot \beta$ means $p \leq a \cdot b$ for some $a \in \alpha, b \in \beta, a, b > 0$. Pick $a' \in \alpha$ and $b' \in \beta$ with $a' > a$ and $b' > b$. Form $a'b' > ab \geq p$, $a'b' \leq a'b'$ means $a'b' \in \alpha \cdot \beta$.

Hence $\alpha \cdot \beta$ is a Dedekind cut. □

§7.4.1 *Properties*

Theorem 7.26 (**R** is archimedian)**.** For any $x \in \mathbf{R}^+$ and $y \in \mathbf{R}^+$, there exists some $n \in \mathbf{Z}^+$ so that

$$n \cdot x > y.$$

Proof. In particular, if we take $x = 1$ from this theorem, we immediately get the following statement.

Proposition 7.27. For any $y \in \mathbf{R}$, there exists some positive integer n so that $n > y$.

We now give a proof of Proposition 7.27 directly without using Theorem 7.26, and then we prove Theorem 7.26 from Proposition 7.27. This shows that these two statements are in fact equivalent, though Proposition 7.27 looks much simpler.

Proof. Assume $n \in \mathbf{Z}^+$ does not exist; that is to say that the set of positive integers $\mathbf{Z}^+$ has an upper bound y. Then using the l.u.b. property of $\mathbf{R}$, $\sup \mathbf{Z}^+$ exists, which we denote by $x_0 \in \mathbf{R}$.

Now we look at $x_0 - 1$. This is not an upper bound by definition of x_0, which means there exists some $N \in \mathbf{Z}^+$ such that

$$x_0 - 1 < N.$$

Then it follows that $x_0 < N + 1$. Notice that $N + 1 \in \mathbf{Z}^+$. So this contradicts the assumption that x_0 is an upper bound.

Hence our original assumption cannot be true, and thus there exists $n \in \mathbf{Z}^+$ with $n > y$. □

For any $x \in \mathbf{R}^+$ and $y \in \mathbf{R}$, consider $y \cdot x^{-1} \in \mathbf{R}$. From Proposition 7.27, there exists some $n \in \mathbf{Z}^+$ such that

$$n > y \cdot x^{-1}.$$

Then this is equivalent to $n - yx^{-1} > 0$. Since $x > 0$, and $\mathbf{R}$ is an ordered field, we have

$$(n - y \cdot x^{-1}) \cdot x > 0.$$

This is equivalent to $n \cdot x > y$. □

Remark. The archimedian property guarantees that we can use decimals to represent real numbers.

Theorem 7.28 (**Q** is dense in **R**)**.** For any $a, b \in \mathbf{R}$ with $a < b$, there exists some $x \in \mathbf{Q}$ such that $a < x < b$.

Proof. This means one can find some $m \in \mathbf{Z}$ and $n \in \mathbf{Z}^+$ so that

$$a < \frac{m}{n} < b,$$

which is further equivalent to finding $m \in \mathbf{Z}$ and $n \in \mathbf{Z}^+$ so that

$$an < m < bn.$$

Notice that $b - a > 0$, so by the archimedian property, there exists $n \in \mathbf{Z}^+$ so that

$$bn - an = (b - a)n > 1.$$

We now argue that there exists some integer between two real numbers, whenever their difference is larger than 1.

Lemma 7.29. For any $\alpha, \beta \in \mathbf{R}$ with $\beta - \alpha > 1$, there exists some integer m so that $\alpha < m < \beta$.

Proof. We prove this lemma by finding such m. First, using archimedian property of $\mathbf{R}$, we can find some integer $N > 0$ so that

$$-N < \alpha < \beta < N.$$

Then consider the integers which are smaller than N and greater than α, i.e., the set

$$A \coloneqq \{k \in \mathbf{Z} \mid a < k \leq N\}.$$

It is not empty since $N \in A$. Since this a subset of $\{-N+1, -N+2, \ldots, N-2, N-1, N\}$ which is a finite set, it contains only finite elements. We can pick the smallest one from it and denote it by m, i.e., $m \coloneqq \min A$. We claim this m is just the one we are looking for.

First since $m \in A$, $m > \alpha$. Then we only need to check $m < \beta$. If this is not true, i.e., $m \geq \beta$, then we consider $m - 1$. It follows

$$m - 1 \geq \beta - 1 \geq \alpha.$$

This contradicts the fact that m is the smallest integer which is greater than α.

Above all, we are done with the lemma. □

At last, apply the lemma to $\alpha = an$ and $\beta = bn$, we are done. □

Theorem 7.30 ($\mathbf{R}$ is closed under taking roots)**.** For every $y \in \mathbf{R}^+$ and every $n \in \mathbf{Z}^+$, there exists a unique $x \in \mathbf{R}^+$ so that $x^n = y$.

Proof. We first claim that such $x \in \mathbf{R}^+$, if exists, must be unique. Otherwise, assume that both $x_1, x_2 \in \mathbf{R}^+$ are solutions of the equation

$$x^n = y, \quad y \in \mathbf{R}^+, n \in \mathbf{Z}^+.$$

Assume now $x_1 < x_2$, then from the fact that $\mathbf{R}$ is an ordered field, we have $x_1^n < x_2^n$ (why?), a contradiction. Similarly, $x_1 > x_2$ also leads to a contradiction, and so $x_1 = x_2$.

Now we look for a solution for the equation. Consider a subset of $\mathbf{R}$ as

$$S \coloneqq \{a \in \mathbf{R}^+ \mid a^n < y\}.$$

Try to check that

(1) $S \neq \emptyset$;

(2) S has an upper bound.

Then using the fact that $\mathbf{R}$ has the l.u.b. property, $\sup S$ exists. Define it as x, clearly $x \in \mathbf{R}^+$. We show that x solves the equation. (The idea of the proof is similar to the proof of $\sup_{\mathbf{Q}}\{x \in \mathbf{Q} \mid x^2 \leq 2\}$ does not exist.)

First, we show that if $x^n < y$, then we can construct some $x_0 \in S$ which is greater than x, which says x is not an upper bound of S. So $x^n \geq y$.

Second, we show that if $x^n > y$, then we can find an upper bound of S which is smaller than x, which says that x is not the least upper bound. So $x^n \leq y$.

Above all, we must have $x^n = y$.

From now on, we use $y^{\frac{1}{n}}$ to denote the unique solution for the equation

$$x^n = y, \quad y \in \mathbf{R}^+, n \in \mathbf{Z}^+,$$

and call it the n-th real root of y. The property

$$(ab)^{\frac{1}{n}} = a^{\frac{1}{n}} \cdot b^{\frac{1}{n}}$$

immediately follows from the uniqueness of n-th real root. □

Theorem 7.31 (Completeness axiom for $\mathbf{R}$)**.** If non-empty $E \subset \mathbf{R}$ is bounded above, then E has a supremum.

Any set in the reals bounded from above/below must have a supremum/infimum.

Proof. We prove this using Dedekind cuts.

Let S be a real number set. We consider the rational number set $A = \{x \in \mathbf{Q} \mid \exists y \in S\}$. Set B is defined to be the complement of A in $\mathbf{Q}$.

We go through the definitions to check that $(A|B)$ is a Dedekind cut.

1. Since $S \neq \emptyset$, pick $y \in S$, then $[y] - 1$ is a real number smaller than some element in S, hence $[y] - 1 \in A$ and thus $A \neq \emptyset$.

   Since we're given that S is bounded, $\exists M > 0$ as the upper bound for S, thus $B \neq \emptyset$.

   (Note that an upper bound is simply a number that is bigger than anything from the set, and is not the supremum

2. We defined B to be the complement of A in $\mathbf{Q}$, so this condition is trivial.

3. For any $x, y \in A$, if $x < y$ and $y \in A$, then $\exists z \in S$ such that $y < z \implies x < z \implies x \in A$.

4. Suppose otherwise that $x \in A$ is the largest element in A, then $\exists y \in S$ such that $x < y$ We then pick a rational number z between x and y. Since we still have $z < y$, we have $z \in A$ but $z > x$, contradictory to z being the largest.

   Now there's actually an issue with the proof for property 4 here How exactly are we finding z?

   First $x \in \mathbf{Q}$. Then $y \in \mathbf{R}$ so we rewrite it as $y = (C|D)$ via definition.

   $x < y$ translates to the fact that $x \in C$.

   Since y is real, by definition we know that C must not have a largest element.

   In particular, x is not largest and we can pick $z \in C$ such that $z > x$. This is in fact the z that we need

Now that all the properties of a real number are validated, we may finally conclude that $\alpha = (A|B)$ is indeed a real number.

Now we need to show that $\alpha = \sup S$.

Let $x \in S$. If x is not the maximum value of S, i.e. $\exists y \in S, x < y$, then $x \in A$ and thus $x < \alpha$.

If x is the maximum value of S, then for any rational number $y < x$ we have $y \in A$, and for any rational number $y \geq x$ we have $y \in B$. Thus $x = (A|B) = \alpha$.

In conclusion, $x \leq \alpha$ for all $x \in S$.

For any upper bound x of S, since $\forall y \in S, x \geq y$ we have $x \in B$ and thus $x \geq \alpha$.

$\therefore \alpha$ is the smallest upper bound of S and thus $\sup S = \alpha$ exists. □

§7.4.2 *Extended real number system*

Definition 7.32. We add $\pm\infty$ to $\mathbf{R}$, and call the union $\mathbf{R} \cup \{\pm\infty\}$ the ***extended real number system***. Now any non-empty set $E \subset \mathbf{R}$ has a supremum and infimum, since we can define

$$\sup E = +\infty, \quad \text{if } E \text{ has no upper bound in } \mathbf{R}$$

and

$$\inf E = -\infty, \quad \text{if } E \text{ has no lower bound in } \mathbf{R}.$$

The extended real number system does not form a field, but it is customary to make the following conventions:

(1) If x is real then

$$x + \infty = +\infty, \quad x - \infty = -\infty, \quad \frac{x}{+\infty} = \frac{x}{-\infty} = 0.$$

(2) If $x > 0$ then $x \cdot (+\infty) = +\infty$, $x \cdot (-\infty) = -\infty$.

(3) If $x < 0$ then $x \cdot (+\infty) = -\infty$, $x \cdot (-\infty) = +\infty$.

When it is desired to make the distinction between real numbers on the one hand and the symbols $+\infty$ and $-\infty$ on the other quite explicit, the former are called ***finite***.

§7.5 *Complex Numbers*

We consider the Cartesian product of $\mathbf{R}$ with $\mathbf{R}$; that is,

$$\mathbf{R}^2 \coloneqq \mathbf{R} \times \mathbf{R} \coloneqq \{(x_1, x_2) \mid x_1, x_2 \in \mathbf{R}\}.$$

Over $\mathbf{R}^2$, we can define operations

- Addition $+$: $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$;
- Scalar multiplication $\mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}^2$: $c \cdot (x_1, x_2) = (c \cdot x_1, c \cdot x_2)$.

This two operations make $\mathbf{R}^2$ a 2-dimensional vector space (linear space) over the real field $\mathbf{R}$. We also say $\mathbf{R}^2$ is a $\mathbf{R}$-linear space of real dimension 2. For example, $\{(1, 0), (0, 1)\}$ form a basis of $\mathbf{R}^2$.

Moreover, over the linear space $\mathbf{R}^2$, one can define an inner product as

$$\langle (x_1, x_2), (y_1, y_2) \rangle = x_1 y_1 + x_2 y_2.$$

The inner product induces a norm

$$|(x_1, x_2)| = \sqrt{\langle (x_1, x_2), (x_1, x_2) \rangle} = \sqrt{x_1^2 + x_2^2}.$$

From now on, we use $\vec{x}$ to denote (x_1, x_2).

Proposition 7.33.

- $|\vec{x}| \geq 0$, where equality holds if and only if $\vec{x} = \vec{0}$.
- $|c \cdot \vec{x}| = |c||\vec{x}|$
- $|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|$
- $|\langle \vec{x}, \vec{y} \rangle| \leq |\vec{x}||\vec{y}|$

All constructions here can be easily generalised to any $\mathbf{R}^n$ with $n \in \mathbf{Z}^+$.

Over $\mathbf{R}^2$, we can define a multiplication $\cdot$ as

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

If we identity $\mathbf{R}^2$ with

$$\mathbf{C} \coloneqq \{x + yi \mid x, y \in \mathbf{R}\}$$

via $(x, y) \mapsto x + yi$, then all structures defined above are induced to $\mathbf{C}$. In particular, the multiplication is induced to $\mathbf{C}$ via requiring $i^2 = -1$. A nontrivial fact is that $(\mathbf{C}, +, \cdot)$ is a field. A element in $\mathbf{C}$ is called a complex number. Usually, people prefer to use $z = x + yi$, $x, y \in \mathbf{R}$, to denote a complex number. Here x is called the real part of z and y is called the imaginary part of z. We use $|z|$ to denote its norm.

§7.6 *Euclidean Spaces*

For each positive integer n, let $\mathbf{R}^n$ be the set of all ordered n-tuples

$$\mathbf{x} = (x_1, x_2, \ldots, x_n),$$

where $x_1, \ldots, x_n$ are real numbers, called the ***coordinates*** of $\mathbf{x}$. The elements of $\mathbf{R}^n$ are called points, or vectors, especially when $n > 1$. We shall denote vectors by boldfaced letters.

Since $\mathbf{R}^n$ is a vector space (over $\mathbf{R}$), $\mathbf{R}^n$ has the following extra properties

- For any two vectors $\mathbf{x}$ and $\mathbf{y}$ we may perform addition:

  $$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$$

  Properties of addition:

  1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
  2. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
  3. Zero vector $\mathbf{0} = (0, \ldots, 0)$ satisfies $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$
  4. For any vector $\mathbf{x}$, its negative $-\mathbf{x}$ satisfies $\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$

- For any vector $\mathbf{x}$ and scalar $k \in \mathbf{R}$ we may perform scalar multiplication:

  $$k\mathbf{x} = (kx_1, \ldots, kx_n)$$

  Properties of scalar multiplication:

  1. $0 \cdot \mathbf{x} = \mathbf{0}, 1 \cdot \mathbf{x} = \mathbf{x}$
  2. $(kl)\mathbf{x} = k(l\mathbf{x}) = l(k\mathbf{x})$
  3. $k(\mathbf{x} + \mathbf{y}) = k\mathbf{x} + k\mathbf{y}$
  4. $(k + l)\mathbf{x} = k\mathbf{x} + l\mathbf{x}$

We define the **inner product** (or scalar product) of $\mathbf{x}$ and $\mathbf{y}$ by

$$\mathbf{x} \cdot \mathbf{y} \coloneqq \sum_{i=1}^{n} x_i y_i.$$

The Euclidean space builds upon the vector space $\mathbf{R}^n$; specifically speaking, it is $\mathbf{R}^n$ endowed with two extra notions:

- The **norm** of the Euclidean space $\|\cdot\|$ is a real-valued function $\|\cdot\| : \mathbf{R}^n \to \mathbf{R}$. Given a vector $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathbf{R}^n$, the norm of $\mathbf{x}$ is defined as

  $$\|\mathbf{x}\| \coloneqq \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

- The **metric** d of the Euclidean space is a real-valued function $d : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$. Given two vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, the distance between $\mathbf{x}$ and $\mathbf{y}$ is defined as

  $$d(\mathbf{x}, \mathbf{y}) \coloneqq \|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

Remark. The norm is something like the length of the vector itself (distant to the origin); the metric refers to the distance function which measures the length between two points in $\mathbf{R}^n$ (determined by their positional vectors $\mathbf{x}$ and $\mathbf{y}$). Essentially, the metric is a much more general notion than the norm: the norm can only be defined on vector spaces; the metric can literally be defined on any set.

Norms are required to satisfy the following properties:

(1) (**positive definiteness**) for any vector $\mathbf{x}$, $\|\mathbf{x}\| \geq 0$, and equality holds if and only if $\mathbf{x} = \mathbf{0}$.

(2) (**absolute homogeneity**) for any vector $\mathbf{x}$ and scalar a, $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$.

(3) (**triangle inequality**) for any two vectors $\mathbf{x}$ and $\mathbf{y}$, $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Metrics are required to satisfy the following properties:

(1) (**positive definiteness**) for any two elements $\mathbf{x}$ and $\mathbf{y}$, $d(\mathbf{x},\mathbf{y}) \geq 0$, equality holds if and only if $\mathbf{x} = \mathbf{y}$.

(2) (**symmetry**) for any two elements $\mathbf{x}$ and $\mathbf{y}$, $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$.

(3) (**triangle inequality**) for any three elements $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$, $d(\mathbf{x},\mathbf{z}) \leq d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})$.

Generally, if there is a norm $\|\cdot\|$ on some vector space, then this norm naturally determines a metric $d(x,y) = \|x-y\|$, which is precisely the case for Euclidean spaces.

Definition 7.34. $E \subset \mathbf{R}^n$ is ***bounded*** if there exists $M > 0$ such that $\|x\| \leq M$ for all $x \in E$.

Exercise 23
Given E and F in $\mathbf{R}^n$ and real number k, define

$$kE = \{kx \mid x \in E\}$$

$$E + F = \{x + y \mid x \in E, y \in F\}$$

(a) Show that if E is bounded, then kE is bounded;

(b) Show that if E and F are bounded, then $E + F$ is bounded.

Definition 7.35. The ***diameter*** of $E \subset \mathbf{R}^n$ is defined as

$$\operatorname{diam} E \coloneqq \sup_{x,y \in E} d(x,y).$$

Exercise 24
Find the diameter of the open unit ball in $\mathbf{R}^n$ given by

$$B = \{x \in \mathbf{R}^n \mid \|x\| < 1\}.$$

Solution. First note that

$$d(x,y) = \|x-y\| \leq \|x\| + \|-y\| = \|x\| + \|y\| < 1 + 1 = 2.$$

On the other hand, for any $\epsilon > 0$, we pick

$$x = \left(1 - \frac{\epsilon}{4}, 0, \ldots, 0\right), \quad y = \left(-\left(1 - \frac{\epsilon}{4}\right), 0, \ldots, 0\right).$$

Then $d(x,y) = 2 - \dfrac{\epsilon}{2} > 2 - \epsilon$.

Therefore $\operatorname{diam} B = 2$. □

Exercise 25
Given a set E in $\mathbf{R}^n$, show that E is bounded if and only if $\operatorname{diam} E < +\infty$.

Proof.

($\Longrightarrow$) If E is bounded, then there exists $M > 0$ such that $\|x\| \leq M$ for all $x \in E$.

Thus for any $x, y \in E$,

$$d(x, y) = \|x - y\| \leq \|x\| + \|y\| \leq 2M.$$

Thus $\operatorname{diam} E = \sup d(x, y) \leq 2M < +\infty$.

($\Longleftarrow$) Suppose that $\operatorname{diam} E = r$. Pick a random point $x \in E$, suppose that $\|x\| = R$.

Then for any other $y \in E$,

$$\|y\| = \big\|x + (y - x)\big\| \leq \|x\| + \|y - x\| \leq R + r.$$

Thus, by picking $M = R + r$, we obtain $\|y\| \leq M$ for all $y \in E$, and we are done.

Remark. Basically you use x to confine E within a ball, which is then confined within an even bigger ball centered at the origin.

□

Definition 7.36. The ***distance between sets*** $E \subset \mathbf{R}^n$ and $F \subset \mathbf{R}^n$ is defined as

$$d(E, F) \coloneqq \inf_{x \in E, y \in F} \|x - y\| .$$

Obviously $d(E, F) > 0$ implies that E and F are disjoint, but E and F may still be disjoint even if $d(E, F) = 0$. For example, the closed intervals $E = (-1, 0)$ and $F = (0, 1)$.

Exercise 26

Suppose that E and F are sets in $\mathbf{R}^n$ where E and F is finite. Prove that E and F are disjoint if and only if $d(E, F) > 0$.

Exercise 27 (Rudin Q1)
If $r \neq 0$ is rational and x is irrational, prove that $r + x$ and rx are irrational.

Solution. We prove by contradiction. Suppose $r + x$ is rational, then $r + x = \frac{m}{n}, m, n \in \mathbf{Z}$, and m, n have no common factors. Then $m = n(r + x)$. Let $r = \frac{p}{q}, p, q \in \mathbf{Z}$, the former equation implies that $m = n\left(\frac{p}{q} + x\right)$, i.e., $qm = n(p + qx)$, giving

$$x = \frac{mq - np}{nq},$$

which says that x can be written as the quotient of two integers, so x is rational, a contradiction.
The proof for the case rx is similar. □

Exercise 28 (Rudin Q2)
Prove that there is no rational number whose square is 12.

8 Basic Topology

§8.1 *Metric Space*

Definition 8.1. A set X, whose elements we shall call ***points***, is a ***metric space*** if for any two points $p, q \in X$ there is associated a real value function (called distance function or ***metric***) $d : X \times X \to \mathbf{R}$ which satisfies the following properties:

(i) (**positive definitiveness**) $d(p,q) \geq 0$, where equality holds if and only if $x = y$;

(ii) (**symmetry**) $d(p,q) = d(q,p)$;

(iii) (**triangle inequality**) $d(p,q) \leq d(p,r) + d(r,q)$ for any $r \in X$.

Example 8.2

Take $X = \mathbf{R}^n$. Then each of the following functions define metrics on X.

$$d_1(x,y) = \sum_{i=1}^{n} |x_i - y_i|;$$

$$d_2(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)}$$

$$d_\infty(x,y) = \max_{i \in \{1,2,\ldots,n\}} |x_i - y_i|.$$

These are called the ℓ^1-("ell one"), ℓ^2- (or Euclidean) and ℓ^∞-distances respectively. Of course, the Euclidean distance is the most familiar one.

The proof that each of d_1, d_2, d_∞ is a metric is mostly very routine, with the exception of proving that d_2, the Euclidean distance, satisfies the triangle inequality. To establish this, recall that the Euclidean norm $\|x\|_2$ of a vector $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$ is

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

where the inner product is given by

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

Then $d_2(x,y) = \|x - y\|_2$, and so the triangle inequality is the statement that

$$\|w - y\|_2 \leq \|w - x\|_2 + \|x - y\|_2 .$$

This follows immediately by taking $u = w - x$ and $v = x - y$ in the following lemma.

Lemma 8.3. If $u, v \in \mathbf{R}^n$ then $\|u+v\|_2 \leq \|u\|_2 + \|v\|_2$.

Proof. Since $\|u\|_2 \geq 0$ for all $u \in \mathbf{R}^n$, the desired inequality is equivalent to

$$\|u+v\|_2^2 \leq \|u\|_2^2 + 2\|u\|_2\|v\|_2 + \|v\|_2^2 .$$

But since $\|u+v\|_2^2 = \langle u+v, u+v\rangle = \|u\|_2^2 + 2\langle u, v\rangle + \|v\|_2^2$, this inequality is immediate from the Cauchy–Schwarz inequality, that is to say the inequality $|\langle u, v\rangle| \leq \|u\|_2\|v\|_2$. □

Example 8.4 (Discrete metric)

Let X be an arbitrary set. The **discrete metric** on a set X is defined as follows:

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

The axioms for a metric are easy to check.

Now we turn to some metrics which come up very naturally in diverse areas of mathematics. Our first example is critical in number theory, and also serves to show that metrics need not conform to one's most naïve understand of "distance".

Example 8.5 (2-adic metric)

Let $X = \mathbf{Z}$, and define $d(x,y)$ to be 2^{-m}, where 2^m is the largest power of two dividing $x-y$. The triangle inequality holds in the following stronger form, known as the ultrametric property:

$$d(x,z) \leq \max\{d(x,y), d(y,z)\}.$$

Indeed, this is just a rephrasing of the statement that if 2^m divides both $x-y$ and $y-z$, then 2^m divides $x-z$.

This metric is very unlike the usual distance. For example, $d(999, 1000) = 1$, whilst $d(0, 1000) = \frac{1}{8}$.

The role of 2 can be replaced by any other prime p, and the metric may also be extended in a natural way to the rationals $\mathbf{Q}$.

Metrics are also ubiquitous in graph theory:

Example 8.6 (Path metric)

Let G be a graph, that is to say a finite set of vertices V joined by edges. Suppose that G is connected, that is to say that there is a path joining any pair of distinct vertices. Define a distance d as follows: $d(v,v) = 0$, and $d(v,w)$ is the length of the shortest path from v to w. Then d is a metric on V, as can be easily checked.

They also come up in group theory:

Example 8.7 (Word metric)

Let G be a group, and suppose that it is generated by elements a, b and their inverses. Define a distance on G as follows: $d(v,w)$ is the minimal k such that $v = wg_1 \cdots g_k$, where $g_i \in \{a, b, a^{-1}, b^{-1}\}$ for all i.

When G is finite, the word metric is a special case of the path metric – you may wish to think about why.

There are many metrics with a prominent position in computer science, for instance:

Example 8.8 (Hamming distance)

Let $X = \{0,1\}^n$ (the boolean cube), the set of all strings of n zeroes and ones. Define $d(x,y)$ to

be the number of coordinates in which x and y differ.

It hardly need be said that metrics are ubiquitous in geometry.

Example 8.9 (Projective space)

Consider the set $\mathbf{P}(\mathbf{R}^n)$ of one-dimensional subspaces of $\mathbf{R}^n$, that is to say lines through the origin. One way to define a distance on this set is to take, for lines L_1, L_2, the distance between L_1 and L_2 to be

$$d(L_1, L_2) = \sqrt{1 - \frac{|\langle v, w\rangle|^2}{\|v\|^2\|w\|^2}}$$

where v and w are any non-zero vectors in L_1 and L_2 respectively. It is easy to see this is independent of the choice of vectors v and w. The Cauchy–Schwarz inequality ensures that d is well-defined, and moreover the criterion for equality in that inequality ensures positivity. The symmetry property is evident, while the triangle inequality is left as an exercise.

It is useful to think of the case when $n = 2$ here, that is, the case of lines through the origin in the plane $\mathbf{R}^2$. The distance between two such lines given by the above formula is then $\sin\theta$ where θ is the angle between the two lines (another exercise).

§8.1.1 *Norms*

Definition 8.10 (Norms)**.** Let V be any vector space (over the reals). A function $\|\cdot\| : V \to [0, \infty)$ is called a norm if the following are all true:

(1) $\|x\| = 0$ if and only if $x = 0$;

(2) $\|\lambda x\| = |\lambda|\|x\|$ for all $\lambda \in \mathbf{R}$, $x \in V$;

(3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$.

Given a norm, it is very easy to check that $d(x, y) := \|x - y\|$ defines a metric on V. Indeed, we have already seen that when $V = \mathbf{R}^n$, $\|\cdot\|_2$ is a norm (and so the name "Euclidean norm" is appropriate) and we defined $d_2(x, y) = \|x - y\|_2$.

As we mentioned, the other metrics in Example 8.2 also come from norms. Indeed, d_1 comes from the ℓ^1-norm

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|,$$

whilst d_∞ comes from the ℓ^∞-norm

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

The principle of turning norms into metrics is important enough that we state it as a lemma in its own right

Lemma 8.11. Let V be a vector space over the reals, and let $\|\cdot\|$ be a norm on it. Define $d : V \times V \to [0, \infty)$ by $d(x, y) := \|x - y\|$. Then (V, d) is a metric space

Remark. The converse is very far from true. For instance, the discrete metric does not arise from a norm.

All metrics arising from a norm have the ***translation invariance property*** $d(x + z, y + z) = d(x, y)$, as well as the ***scalar invariance*** $d(\lambda x, \lambda y) = |\lambda| d(x, y)$, neither of which are properties of arbitrary metrics. Conversely one can show that a metric with these two additional properties does come from a norm, an exercise we leave to the reader (Hint: the norm must arise as $\|v\| = d(v, 0)$)

We call a vector space endowed with a norm $\|\cdot\|$ a ***normed space***. Whenever we talk about normed spaces it is understood that we are also thinking of them as metric spaces, with the metric being defined by $d(v,w) = \|v-w\|$.

Note that we do not assume that the underlying vector space V is finitedimensional. Here are some examples which are not finite-dimensional (whilst we do not prove that they are not finite-dimensional here, it is not hard to do so and we suggest this as an exercise).

Example 8.12 (ℓ^p spaces)
Let

$$\ell_1 = \left\{ (x_n)_{n=1}^{\infty} \mid \sum_{n\geq 1} |x_n| < \infty \right\},$$
$$\ell_2 = \left\{ (x_n)_{n=1}^{\infty} \mid \sum_{n\geq 1} x_n^2 < \infty \right\},$$
$$\ell_\infty = \left\{ (x_n)_{n=1}^{\infty} \mid \sup_{n\in\mathbf{N}} |x_n| < \infty \right\}.$$

The sets ℓ_1, ℓ_2, ℓ_∞ are all real vector spaces, and moreover $\left\|(x_n)\right\|_1 = \sum_{n\geq 1}|x_n|$, $\left\|(x_n)\right\|_2 = \left(\sum_{n\geq 1} x_n^2\right)^{\frac{1}{2}}$, $\left\|(x_n)\right\|_\infty = \sup_{n\in\mathbf{N}}|x_n|$ define norms on ℓ_1, ℓ_2 and ℓ_∞ respectively. Note that ℓ_2 is in fact an inner product space where

$$\langle (x_n), (y_n) \rangle = \sum_{n\geq 1} x_n y_n,$$

(the fact that the right-hand side converges if (x_n) and (y_n) are in ℓ_2 follows from the Cauchy–Schwarz inequality).

The space ℓ^2 is known as ***Hilbert space*** and it is of great importance in mathematics.

§8.1.2 *New metric spaces from old one*

A metric space (X,d) naturally induces a metric on any of its subsets.

Definition 8.13 (Subspace)**.** Suppose that (X,d) is a metric space and let Y be a subset of X. Then the restriction of d to $Y\times Y$ gives Y a metric so that $(Y, d_{Y\times Y})$ is a metric space. We call Y equipped with this metric a ***subspace***.

Example 8.14
If $X = \mathbf{R}$, we could take $Y = [0,1]$, for instance, or $Y = \mathbf{Q}$, or $Y = \mathbf{Z}$.

Definition 8.15 (Product space)**.** If (X, d_X) and (Y, d_Y) are metric spaces, then it is natural to try to make $X\times Y$ into a metric space. One method is as follows: if $x_1, x_2 \in X$ and $y_1, y_2 \in Y$ then we set

$$d_{X\times Y}\left((x_1,y_1),(x_2,y_2)\right) = \sqrt{d_X(x_1,x_2)^2 + d_Y(y_1,y_2)^2}.$$

The use of the square mean on the right, rather than the max or the sum, is appealing since then the product $\mathbf{R}\times\mathbf{R}$ becomes the space $\mathbf{R}^2$ with the Euclidean metric. However, either of those alternative definitions results in a metric which is equivalent.

Proposition 8.16. With notation as above, $d_{X\times Y}$ gives a metric on $X\times Y$.

Proof. Reflexivity and symmetry are obvious. Less clear is the triangle inequality. We need to prove

that

$$\begin{aligned}&\sqrt{d_X(x_1,x_3)^2+d_Y(y_1,y_3)^2}+\sqrt{d_X(x_3,x_2)^2+d_Y(y_3,y_2)^2}\\&\geq\sqrt{d_X(x_1,x_2)^2+d_Y(y_1,y_2)^2}\end{aligned}\tag{1}$$

Write $a_1=d_X(x_2,x_3)$, $a_2=d_X(x_1,x_3)$, $a_3=d_X(x_1,x_2)$ and similarly $b_1=d_Y(y_2,y_3)$, $b_2=d_Y(y_1,y_3)$ and $b_3=d_Y(y_1,y_2)$. Thus we want to show

$$\sqrt{a_2^2+b_2^2}+\sqrt{a_1^2+b_1^2}\geq\sqrt{a_3^2+b_3^2}.\tag{2}$$

To prove this, note that from the triangle inequality we have $a_1+a_2>a_3$, $b_1+b_2>b_3$. Squaring and adding gives

$$a_1^2+b_1^2+a_2^2+b_2^2+2(a_1a_2+b_1b_2)\geq a_3^2+b_3^2.$$

By Cauchy–Schwarz,

$$a_1a_2+b_1b_2\leq\sqrt{a_1^2+b_1^2}\sqrt{a_2^2+b_2^2}.$$

Substituting this into the previous line gives precisely the square of (2), and (1) follows. □

§8.1.3 *Balls and boundedness*

Definition 8.17 (Balls)**.** Let X be a metric space. IF $x\in X$ and $r>0$, we define the ***open ball*** centred at x with radius r to be the set

$$B_r(x)\coloneqq\{y\in X\mid d(x,y)<r\}.$$

Similarly we define the ***closed ball*** centred at x with radius r to be the set

$$\bar{B}_r(x)\coloneqq\{y\in X\mid d(x,y)\leq r\}.$$

The ***punctured ball*** is defined as

$$B_r(x)\setminus\{x\}=\{y\in X\mid 0<d(x,y)<r\}.$$

Definition 8.18 (Bounded)**.** Let X be a metric space, and let $Y\subseteq X$. Then we say that Y is ***bounded*** if Y is contained in some open ball.

Proposition 8.19. Let X be a metric space and let $Y\subseteq X$. Then the following are equivalent:

(i) Y is bounded;

(ii) Y is contained in some closed ball;

(iii) The set $\{d(y_1,y_2)\mid y_1,y_2\in Y\}$ is a bounded subset of $\mathbf{R}$.

Proof. That (i) implies (ii) is totally obvious. That (ii) implies (iii) follows immediately from the triangle inequality. Finally, suppose Y satisfies (iii). Then there is some K such that $d(y_1,y_2)\leq K$ whenever $y_1,y_2\in Y$. If Y is empty, it is certainly bounded. Otherwise, let $a\in Y$ be an arbitrary point. Then Y is contained in $B_r(a)$ where $r=K+1$. □

Example 8.20

An open (closed) ball in $\mathbf{R}$ is equivalent to a finite open (closed) interval, i.e. (a,b) ($[a,b]$), $a,b\in\mathbf{R}$.

Definition 8.21 (Neighbourhood)**.** A set $N\subset X$ is called a ***neighbourhood*** of $x\in X$ if $\exists r>0$ s.t. $B_r(x)\subset N$.

§8.1.4 *Open and closed sets*

Definition 8.22 (Open set)**.** If X is a metric space, we say $E \subseteq X$ is ***open*** (in X) if it is a neighbourhood of each of its elements, i.e., $\forall x \in E \exists r > 0$ s.t. $B_r(x) \subset E$.

Proposition 8.23. Any open ball is open.

Proof. Assume $B_r(x)$ is an open ball in a metric space (X, d). Then for any point $y \in B_r(x)$, there is

$$d(y, x) < r.$$

Now we define $r' \coloneqq r - d(y, x)$, which is positive.

Consider the ball $B_{r'}(y)$. We shall show it lives in $B_r(x)$. For this, take any point $z \in B_{r'}(y)$. Using the triangle inequality of a metric, we have

$$\begin{aligned} d(z, x) &\leq d(z, y) + d(y, x) \\ &< r' + d(y, x) \\ &= r. \end{aligned}$$

Hence $z \in B_r(x)$, and $B_{r'}(y) \subset B_r(x)$. □

Proposition 8.24. (1) Both $\emptyset$ and X are open.

(2) If E_1, E_2 are open, then $E_1 \cap E_2$ is open.

(3) If E_i is open for $i \in I$, then $\bigcup_{i \in I} E_i$ is open.

An arbitrary union of open sets is open; a finite intersection of open sets is open.

Proof.

(1) Obvious by definition.

(2) Take a point $x \in E_1 \cap E_2$, we need to find an open ball with radius $r > 0$ such that $x \in B_r(x) \subset E_1 \cap E_2$.

To find such $r > 0$, notice that since both E_1 and E_2 are open, there are open balls

$$\begin{aligned} x \in B_{r_1}(x) \subset E_1 \\ x \in B_{r_2}(x) \subset E_2 \end{aligned}$$

Take $r \coloneqq \min\{r_1, r_2\}$. Then $B_r(x) \subset B_{r_1}(x) \subset E_1$ and $B_r(x) \subset B_{r_2}(x) \subset E_2$, and hence $B_r(x) \subset E_1 \cap E_2$.

(3) Take a point $x \in \bigcup_{i \in I} E_\alpha$, then we can assume x lives in some E_k, $k \in I$. Since E_k is open, take an open ball

$$B_r(x) \subset E_k.$$

It follows

$$B_r(x) \subset E_k \subset \bigcup_{i \in I} E_i.$$

Hence $\bigcup_{i \in I} E_i$ is open.

□

Example 8.25

We know $I_n \coloneqq \left(-\frac{1}{n}, \frac{1}{n}\right) \subset \mathbf{R}$ is open for any $n \in \mathbf{Z}^+$. However, $\bigcap_{n \in \mathbf{Z}^+} I_n = \{0\}$ is not open.

The complement of an open set is a closed set.

Definition 8.26 (Closed set)**.** E is ***closed*** if its complement E^c is open.

Example 8.27
The closed interval $[a, b]$, $a \leq b$ is closed in $\mathbf{R}$.

Proposition 8.28. Any closed ball is closed.

Proof. To prove that $\bar{B}_r(x) = \{y \in X \mid d(x, y) \leq r\}$ is closed, we need to show that its complement $\bar{B}_r(x)^c = \{y \in X \mid d(x, y) > r\}$ is open.

Let $z \in \bar{B}_r(x)^c$. Choose $r' > 0$ such that $r + r' < d(x, z)$; that is, $r' < d(x, z) - r$.

We claim that $B_{r'} \subseteq \bar{B}_r(x)^c$. Pick $y \in B_{r'}(z)$. Then $d(y, z) < r'$. But $r + d(y, z) < d(x, z)$ so $r < d(x, z) - d(y, z) \leq d(x, y)$ by triangle inequality. Hence we have $r < d(x, y)$, thus $y \in \bar{B}_r(x)^c$. Therefore $\bar{B}_r(x)^c$ is open, so $\bar{B}_r(x)$ is closed. □

Proposition 8.29. (1) Both $\emptyset$ and X are closed.

(2) If E_1 and E_2 are closed, then $E_1 \cup E_2$ is closed.

(3) If E_i is closed for $i \in I$, then $\bigcap_{i \in I} E_i$ is closed.

An arbitrary intersection of closed sets is closed; a finite union of closed sets is closed.

Proof.

(1) It follows immediately from $\emptyset = X^c$ and $X = \emptyset^c$.

(2) It follows from above that
$$(E_1 \cup E_2)^c = {E_1}^c \cap {E_2}^c$$
is open (de Morgan's law applied), and hence $E_1 \cup E_2$ is closed.

(3) It follows from above that
$$\left(\bigcap_{i \in I} E_i\right)^c = \bigcup_{i \in I} {E_i}^c$$
is open (de Morgan's law applied), and hence $\bigcap_{i \in I} E_i$ is closed.

□

Example 8.30
Consider a sequence of closed sets $\left[-1 + \frac{1}{n}, 1 - \frac{1}{n}\right]$, $n \in \mathbf{Z}^+$, of $\mathbf{R}$. Take their union
$$\bigcup_{n \in \mathbf{Z}^+} \left[-1 + \frac{1}{n}, 1 - \frac{1}{n}\right] = (-1, 1)$$
which is open, not closed.

Definition 8.31 (Limit point)**.** p is a ***limit point*** of E if every neighborhood of p contains $q \neq p$ such that $q \in E$:
$$\forall r > 0, \exists q \in E, q \neq p \text{ s.t. } q \in B_r(p).$$
The ***induced set*** of E, denoted by E', is the set of all limit points of E in X.

The ***closure*** of E, denoted by $\bar{E}$, is the union set $E \cup E'$.

Example 8.32

- Consider the metric space $\mathbf{R}$, a and b are limit points $(a, b]$. The limit point set of $(a, b]$ is $[a, b]$, which is also the closure $(a, b]$.
- Consider the metric space $\mathbf{R}^2$. The limit point set of any open ball $B_r(x)$ is the closed ball

$\bar{B}_r(x)$, which is also the closure of $B_r(x)$.

- Consider $\mathbf{Q} \subset \mathbf{R}$. $\mathbf{Q}' = \bar{\mathbf{Q}} = \mathbf{R}$.

Proposition 8.33. If p is a limit point of E, then every neighbourhood of p contains infinitely many points of E.

Proof. Prove by contradiction. Suppose there is a neighborhood $B_r(p)$ which contains only a finite number of points of E: $q_1, \ldots, q_n$, which are distinct from p. Define

$$r = \min_{1 \le m \le n} d(p, q_m).$$

The minimum of a finite set of positive numbers is clearly positive, so that $r > 0$.

The neighborhood $B_r(p)$ contains no point $q \in E, q \neq p$ so that p is not a limit point of E, a contradiction. □

Corollary 8.34. A finite point set has no limit points.

Definition 8.35. E is called ***dense*** if $\bar{E} = X$.

Proposition 8.36. (1) A is a dense set in X if and only if A intersects with all open sets in X.

(2) If A is dense in X and B is dense in A, then B is dense in X.

(3) If A and B are dense in X where A is open, then $A \cap B$ is dense in X.

Definition 8.37. (1) p is a ***limit point*** of E if every neighborhood of p contains $q \neq p$ such that $q \in E$:

$$\forall r > 0, \exists q \in E, q \neq p \text{ s.t. } q \in B_r(p).$$

The ***induced set*** of E, denoted by E', is the set of all limit points of E in X.

(2) p is an ***isolated point*** of E if it not a limit point of E.

(3) E is ***closed*** if every limit point of E is a point of E, i.e $\bar{E} = E$.
The ***closure*** of E, denoted by $\bar{E}$, is the union set $E \cup E'$.

(4) p is an ***interior point*** of E if there is a neighborhood N of p such that $N \subset E$:

$$\exists r > 0 \text{ s.t. } B_r(p) \subset E.$$

The ***interior*** of E, denoted by E°, is the set of all interior points in E:

$$E^\circ \coloneqq \{p \in X \mid \exists r > 0 \text{ s.t. } B_r(p) \subset E\}$$

A point x is an ***exterior point*** of A if it is an interior point of A^c.

(5) E is ***open*** if every point of E is an interior point of E, i.e. $E^\circ = E$.

(6) E is ***perfect*** if E is closed and if every point of E is a limit point of E.

(7) The ***boundary*** of E, denoted by ∂E, is the set difference $\bar{E} \setminus E^\circ$.
p is a ***boundary point*** of E if $p \in \partial E$.
E is compact if it is a bounded closed set.

(8) E is ***dense*** in X if every point of X is a limit point of E, or a point of E (or both).
A subset $B \subset A$ is a dense subset of A if $\bar{B} = A$.
E is ***nowhere dense*** its closure has no interior, i.e. $(\bar{E})^\circ = \emptyset$.

Proposition 8.38. (1) $\bar{E}$ is closed;

(2) $E = \bar{E}$ if and only if E is closed;

(3) $\bar{E} \subset F$ for every closed set $F \subset X$ such that $E \subset F$.

By (1) and (3), $\bar{E}$ is the ***smallest*** closed subset of X that contains E.

Proof.

(1)

(2)

(3)

□

Proposition 8.39. (1) E° is open.

(2) E is open if and only if $E = E^\circ$.

(3) If $G \subset E$ and G is open, then $G \subset E^\circ$.

Proof.

(1) If $p \in E^\circ$ then $B_r(p) \subset E$ for some $r > 0$ and if $q \in B_r(p)$ then by triangle inequality, $B_{r-d(p,q)}(q) \subset E$ so $B_r(p) \subset E^\circ$ and hence E° is open.

(2) Certainly if E is open then $E = E^\circ$ since for each $p \in E$ there exists $r > 0$ such that $B_r(p) \subset E$. Conversely if $E^\circ = E$ then this holds for each $p \in E$ so E is open.

(3) If $G \subset E$ is open then for each $p \in G$ there exists $r > 0$ such that $B_r(p) \subset G$, hence $B_r(p) \subset E$ so $p \in E^\circ$ and it follows that $G \subset E^\circ$.

□

Proposition 8.40. The set of exterior points, $(A^c)^\circ$ is the same as $(\bar{A})^c$.

Proof.

$$\begin{aligned} x \in (A^c)^\circ &\iff \exists \epsilon > 0 \text{ such that } B(x,\epsilon) \subset A^c \\ &\iff B(x,\epsilon) \cap A = \emptyset \\ &\iff x \notin A \text{ and } B_0(x,\epsilon) \cap A = \emptyset \\ &\iff x \notin A \cup A' = \bar{A} \\ &\iff x \in (\bar{A}^c) \end{aligned}$$

□

Proposition 8.41. (1) A' is closed.

(2) $\bar{A}$ is closed, i.e. $\bar{\bar{A}} = \bar{A}$

Proof.

(1) In order to show that A' is closed, we need to show that if x is a limit point of A', then $x \in A'$, i.e. x is a limit point of A.

So we need to show that limit points of A' are always limit points of A: Let x be a limit point of A', then for all $\epsilon > 0$, $B_0(x, \epsilon/2)$ intersects with A' and we may pick $y \in B_0(x, \epsilon/2) \cap A'$

Now here's the tricky part Since $y \in A'$, y is a limit point of A, hence $B_0(y, |y - x|)$ intersects with A and thus we may pick $z \in B_0(y, |y - x|) \cap A$.

We show that $z \in B_0(x, \epsilon)$:

$$|z - x| \leq |z - y| + |y - x| < 2|y - x| < \epsilon,$$

hence $z \in B(x, \epsilon)$.

$$|z - y| < |x - y|,$$

hence $z \neq x$

$\therefore\ z \in B_0(x, \epsilon)$

(2)

□

Theorem 8.42 (Cantor's Intersection Theorem)**.** Given a decreasing sequence of compact sets $A_1 \supset A_2 \supset \cdots$, there exists a point $x \in \mathbf{R}^n$ such that x belongs to all A_i. In other words, $\bigcap_{i=1}^{\infty} A_i \neq \emptyset$. Moreover, if for all $i \in \mathbf{N}$ we have $\operatorname{diam} A_{i+1} \leq c \cdot \operatorname{diam} A_k$ for some constant $c < 1$, then such a point must be unique, i.e. $\bigcap_{i=1}^{\infty} A_k = \{x\}$ for some $x \in \mathbf{R}^n$.

Theorem 8.43 (Heine–Borel Theorem)**.** A set $A \subset \mathbf{R}^n$ is compact if and only if every open covering has a finite subcover, i.e. for any family of open sets $\mathscr{U} = \{U_i\}_{i \in I}$ satisfying $A \subset \bigcup_{i \in I} U_i$, there exists $\{U_1, \ldots, U_n\} \subset \mathscr{U}$ such that $A \subset \bigcup_{i=1}^{n} U_i$.

Theorem 8.44 (Bolzano–Weierstrass Theorem)**.** Infinite bounded sets in $\mathbf{R}^n$ must contain limit points.

We will follow a very specific sequence of steps to prove these three theorems:

(a) Cantor Intersection for $n = 1$

(b) Bolzano–Weierstrass for $n = 1$

(c) Bolzano–Weierstrass for general n

(d) Cantor Intersection for general n

(e) Heine–Borel for general n

Proof.

(a) Suppose that there is a decreasing sequence of compact sets $A_1, A_2, \ldots$ in the real numbers

Since A_k are bounded, we may let $a_k = \inf A_k$ Also since A_k are closed, $a_k \in A_k$

Note that since A_k is a decreasing sequence of sets we have $a_1 \leq a_2 \leq \ldots$

Also, whenever we have $n > k$, we have $a_n \in A_n$, but $A_n \subset A_k$ and thus $a_n \in A_k$.

Let $b_1 = \sup A_1$, then $a_k \in A_1$ and thus $a_k \leq b_1$ for all k.

This tells us that the sequence $\{a_k\}$ is bounded above, and thus we may let $a = \sup a_k$.

Our goal is to show that the number a appears in all A_k, thus showing that the entire intersection $\bigcap A_k$ contains a and thus must be non-empty.

Now we split this in two cases, which asks whether a is simply made from isolated points, or if it is actually some nontrivial point obtained from the boundaries of A_k

Case 1: $a_k = a$ for some k In this case we see that $a_k \leq a_n \leq a$ for all $n > k$ and thus $a_n = a$ in this case, therefore a is an element in A_n for all n

In this case you can imagine that there is a possibility where a is an isolated minimum point of A_n which stays there forever in the decreasing sequence of sets

Case 2: $a_k < a$ for all k; in this case we see that a is the limit point of the increasing sequence $\{a_k\}$

Exercise 1: Show that a is a limit point of each A_k.

Note that a_n is in A_k for each $n > k$, and since $a = \sup\{a_k\}$ where a_k is increasing, we can actually show that a is a limit point of $\{a_n \mid n \leq k\}$: For every $\epsilon > 0$, we pick n_0 such that $0 < a - a_{n_0} < \epsilon$ Pick $n\prime > \max\{k, n_0\}$, then $a_{n'} \geq a_{n_0}$ and so

$$0 < a - a_{n\prime} \leq a_{n_0} < \epsilon$$

This shows that there exists a'_n in $B_0(a, \epsilon) \cap \{a_n \mid n > k\}$ for all ϵ, and so a is a limit point of $\{a_n \mid n > k\}$.

Now since $\{a_n | n \geq k\}$ is a subset of A_k we also see that a is a limit point of A_k Finally, since A_k is closed, we conclude that a is in A_k for all k, and we are done

Wait hold on, I forgot about the second part

Now we consider a decreasing sequence of compact sets $A_1, A_2, \dots$ such that $\operatorname{diam} A_{k+1} \leq c \operatorname{diam} A_k$ for $c < 1$.

Suppose otherwise that there exists x, y in $\bigcap A_k$

You can imagine that this will form a fixed distance between two points, and thus there is a constant positive lower bound for the diameters:

$$\operatorname{diam} A_k \geq |x - y| > 0 \forall k$$

But this cannot be true because $\operatorname{diam} A_{k+1} \leq c \operatorname{diam} A_k$ and so the diameter is controlled by a decreasing geometric sequence:

$$\operatorname{diam} A_{k+1} \leq c^k \operatorname{diam} A_1$$

So we can simply pick a natural number k such that

$$k > \log_c \frac{|x - y|}{\operatorname{diam} A_1}$$

(b) We consider an infinite bounded set A in the real numbers. Since A is bounded, we can pick a closed interval $[a_1, b_1]$ containing A.

We then perform a series of binary cuts: Consider the two halves of $[a_1, b_1]$. We know that at least one of these two must contain infinitely many elements in A, otherwise A cannot be infinite. We pick this half of the interval and denote it by $[a_2, b_2]$. We continue this to pick a decreasing sequence of closed intervals $[a_n, b_n]$.

Now $\operatorname{diam}[a_{n+1}, b_{n+1}] = \frac{1}{2} \operatorname{diam}[a_n, b_n]$, so by the Cantor Intersection Theorem, there exists a unique real number c in the intersection $\bigcap [a_n, b_n]$.

We show that this c is in fact a limit point of A.

For any $\epsilon > 0$, we need to show that $B_0(c, \epsilon) \cap A \neq \emptyset$, i.e. we need to find an element $x \neq c$ in A that is less than ϵ apart from c.

We then realize that we can simply exploit the decreasing sequence $[a_n, b_n]$ Since $\operatorname{diam}[a_n, b_n]$ is controlled by a decreasing sequence:

$$\operatorname{diam}[a_{n+1}, b_{n+1}] \leq 1/2^n \operatorname{diam}[a_1, b_1]$$

We take a sufficiently large n so that $b_n - a_n < \epsilon$ Since c is in $[a_n, b_n]$, for all x in $[a_n, b_n]$ we have $|x - c| \leq b_n - a_n < \epsilon$ and therefore $[a_n, b_n]$ is within $B(c, \epsilon)$.

Here's the funny part: $[a_n, b_n]$ contains infinitely many elements of A, so it must contain at least one element in A that is not c.

Therefore this element $x \neq c$ is in $B_0(c, \epsilon)$.

(c) Now we have an infinte bounded set A in $\mathbf{R}^n$

The idea here is to consecutively come up with better and better sequences of points in A. We denote x_i to be the i-th coordinate in $\mathbf{R}^n$.

Our first wish is to pick some elements in A so that they sort of converge at x_1.

Because such considerations of 'restricting to a single coordinate' is important here, we define the projection map to the i-th coordinate by

$$f_i(x_1, \dots, x_n) = x_i$$

So, we look at $f_i(A)$ and try to apply BW for the case where $n = 1$.

However, the problem is that $f_i(A)$ need not be infinite. For example, the set $\{(0,0), (0,1), (0,2), \dots\}$ projected onto the first coordinate is simply $\{0\}$.

This forces us to consider two cases

Exercise 2: Show that $f_i(A)$ is bounded This is simple 1. $f_1(A)$ is infinite, then we can apply BW(n=1) to find a real number c_1 which is a limit point in $f_1(A)$

Here we can construct a sequence of points

$$\{x^{(1),1}, x^{(1),2}, ...\}$$

so that their first coordinates satisfy

$$|x_1^{(1),n} - c_1| < 1/n$$

for all natural number n (I know this notation is cumbersome but the problem is that we need multiple sequences for this proof)

2. $f_1(A)$ is finite, then by the Pigeonhole Principle there exists a real number c_1 such that its preimage $f_1^{-1}(c_1)$ in A is infinite

In this case we can randomly pick a sequence $\{x^{(1),1}, x^{(1),2}, \dots\}$ in A so that their first coordinate is equal to c_1

I forgot to mention something that is implied, but we actually do have the need to vocabasize that the sequence $\{x^{(1),1}, x^{(1),2}, \dots\}$ can be chosen to contain mutually distinct entries

Now that we have a sequence that behaves nice on the first coordinate, we may then move on to the second coordinate

Let $A_1 = \{x^{(1),1}, x^{(1),2}, \dots\}$ We again consider $f_2(A_1)$ in two cases, infinite or finite

In any case, we are able to find a subsequence $\{x^{(2),1}, x^{(2),2}, \dots\}$, where $x^{(2),k} = x^{(1),n_k}$ for some strictly increasing sequence of natural numbers n_k

So that, for the limit point/point with infinite preimage c_2, this sequence satisfies

$$|f_2(x^{(2),n}) - c_2| < \frac{1}{n}$$

Note that the property we have for the second case (we in fact have $f_2(x^{(2),n}) = c_2$) is just a better version of this.

Now, take note that picking this subsequence does no harm whatsoever towards the first coordinate (if anything it would turn out to be better) since

$$|f_1(x^{(2),k}) - c_1| = |f_1(x^{(1),n_k} - c_1| < \frac{1}{n_k} \leq \frac{1}{k}$$

($n_1 < \cdots < n_k$ is a strictly increasing sequence of natural numbers so $n_k \geq k$)

This continues on until we obtain a sequence of points $\{x^{(n),1}, x^{(n),2}, \dots\}$ in A so that

$$|f_i(x^{(n),k} - c_i| < \frac{1}{k} \quad \forall i, k$$

As we can see, the point $c = (c_1, \dots, c_n)$ is in fact a limit point of A as we can always choose a big enough k so that $x^{(n),k}$ is in $B(c, \epsilon) \cap A$.

Since $\{x^{(n),k}\}$ was always chosen to be a sequence of distinct entries, there is no danger for this sequence to always be c, and so c must be a limit point of A.

(d) We may now return to the general case of Cantor.

Suppose that there is a sequence of decreasing compact sets $A_1, A_2, \dots$ in $\mathbf{R}^n$. Note that every point is contained in A_1, so boundedness will never be an issue here.

Since A_k are all nonempty, we can simply pick any element a_k from A_k.

For the uncannily specific case that there are only finitely many $\{a_k\}$ chosen, we simply note that, again by Pigeonhole Principle, one of the a_k appears infinitely often; thus for each A_n we simply pick $n_k > n$ so that A_{n_k} contains a_k, then a_k is in A_{n_k} which is a subset of A_n.

Otherwise, we can then note that $\{a_k\}$ is an infinite bounded set of points, so there must exist a limit point a of $\{a_k\}$.

We can now see that a is always an element of A_k: Using the same technique as Exercise 1, we see that a is a limit point of $\{a_n \mid n > k\}$ and so is a limit point of A_k, therefore a is in A_k as A_k is closed.

This proves the first part of the statement The second part is completely identical to the second part of the $n = 1$ case so we don't need to waste our time there either

(e) We now consider a compact set A with some open covering $\mathscr{U}$.

This theorem is proved by contradiction: Suppose otherwise that set A cannot be covered by any finite collection of open sets in $\mathscr{U}$

Since A is compact, we may enclose it in a closed cube Q_1 (whose edges are parallel to the axes)

Now, for each step, we partition Q into 2^n cubes by cutting it in half from each direction.

Then, starting from Q_1, there must exist one of these smaller cubes, denoted by Q_2, such that $A \cap Q_2$ cannot be covered by a finite collection of open sets in $\mathscr{U}$. Otherwise, if each $A \cap Q$ has a finite cover, then we simply collect all of these open sets together to form a finite cover of A, which violates our assumption.

We continue on to partition Q_n and pick Q_{n+1} so that A_{n+1} has no finite cover (denote $A_n = A \cap Q_n$).

Note that A and Q_n are both compact, so A_n is compact Also we see that there is a decreasing sequence $A_1, A_2, \ldots$ (we can't exactly obtain a relation between $\operatorname{diam} A_n$ and $\operatorname{diam} A_{n+1}$ here)

By Cantor Intersection Theorem we can always find a point x in A located in the intersection $\bigcap A_k$.

Now, since $\mathscr{U}$ is an open covering of A, there exists an open set U in $\mathscr{U}$ such that $x \in U$.

The final key step is to exploit the sequence of decreasing cubes Q_n. So even though there isn't a clear cut way to control the sizes of $\operatorname{diam} A_n$, we do in fact have the property that $\operatorname{diam} Q_{n+1} = \frac{1}{2^n} \operatorname{diam} Q_1$.

Therefore, by picking a sufficiently large n, we can obtain Q_n that is contained in U.

But this is a contradiction. This is because we've specifically chosen the sequence A_n to be sets that do not possess any finite cover $\{U_1, ..., U_n\}$ in $\mathscr{U}$. But here A_n simply would have a one-element cover $\{U\}$.

This completes our proof.

□

Proposition 8.45. Suppose $Y \subset X$. A subset E of Y is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.

Proof.

($\Longrightarrow$) Suppose E is open relative to Y. Thus for each $p \in E$ there exists $r_p > 0$ such that $d(p, q) < r_p$, $q \in Y$ imply $q \in E$.

Let V_p be the set of all $q \in X$ such that $d(p, q) < r_p$, and define

$$G := \bigcup_{p \in E} V_p.$$

Then G is an open subset of X, by

($\Longleftarrow$) □

§8.1.5 *Interiors, closures, limit points*

Definition 8.46. Let X be a metric space, and let $E \subset X$. The interior $int(E)$ of E is defined to be the union of all open subsets of X contained in E.

The ***closure*** $\bar{E}$ is defined to be the intersection of all closed subsets of X containing E.

The set $\bar{E} \setminus int(E)$ is known as the boundary of E and denoted ∂E.

A set $E \subseteq X$ is said to be ***dense*** if $\bar{E} = X$.

Definition 8.47. If X is a metric space and $E \subseteq X$ is any subset, then we say a point $x \in X$ is a limit point of E if any open ball about a contains a point of E other than x itself.

Notation. We will write $L(E)$ for the set of limit points of E.

Definition 8.48. x is a ***isolated point*** of E if $\exists r > 0$ s.t. $B_r(x) \cap S = \{x\}$.

§8.2 *Compactness*

Definition 8.49. By an ***open cover*** of a set E in a metric space X we mean a collection $\{G_i \mid i \in I\}$ of open subsets of X such that

$$E \subset \bigcup_{i \in I} G_i.$$

For $I' \subset I$, if the subcollection $\{G_i \mid i \in I'\}$ is also an open cover of S; that is,

$$E \subset \bigcup_{i \in I'} G_\alpha,$$

then $\{G_i \mid i \in I'\}$ is called a ***subcover***. If moreover, I' is finite, then it is called a ***finite subcover***.

Definition 8.50 (Compactness)**.** A subset K of metric space X is said to be ***compact*** if every open cover of K contains a finite subcover.

Proposition 8.51. Suppose $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y.

Proof.

($\implies$) Suppose K is compact relative to X. Let $\{V_i \mid i \in I\}$ be a collection of sets open relative to Y, such that $K \subset \bigcup_{i \in I} V_i$. By □

sequential compactness A set K is compact if and only if every sequence of points in K has a subsequence that converges to a point in K.

Any continuous function defined on a compact set is bounded.

extreme value theorem

§8.3 *Perfect Sets*

§8.4 *Connectedness*

Definition 8.52. Two subsets A and B of a metric space X are said to be ***separated*** if both $A \cap \bar{B}$ and $\bar{A} \cap B$ are empty, i.e. no point of A lies in the closure of B and no point of B lies in the closure of A.

A set $E \subset X$ is said to be ***connected*** if E is not a union of two non-empty separated sets.

Remark. Separated sets are of course disjoint, but disjoint sets need not be separated. For example, the interval $[0, 1]$ and the segment $(1, 2)$ are not separated, since 1 is a limit point of $(1, 2)$. However, the segments $(0, 1)$ and $(1, 2)$ are separated.

The connected subsets of the line have a particularly simple structure:

Proposition 8.53. A subset $E \subset \mathbf{R}^1$ is connected if and only if it has the following property: if $x, y \in E$ and $x < z < y$, then $z \in E$.

Proof. ($\impliedby$) If there exists $x, y \in E$ and some $z \in (x, y)$ such that $z \notin E$, then $E = A_z \cup B_z$ where

$$A_z = E \cap (-\infty, z), \quad B_z = E \cap (z, \infty).$$

Since $x \in A_z$ and $y \in B_z$, A and B are non-empty. Since $A_z \subset (-\infty, z)$ and $B_z \subset (z, \infty)$, they are separated. Hence E is not connected.

($\implies$) Suppose E is not connecetd. Then there are non-empty separated sets A and B such that $A \cup B = E$. Pick $x \in A$, $y \in B$, and WLOG assume that $x < y$. Define

$$z := \sup(A \cap [x, y].)$$

By □

Definition 8.54. We say that a metric space is ***disconnected*** if we can write it as the disjoint union of two nonempty open sets. We say that a space is ***connected*** if it is not disconnected.

If X is written as a disjoint union of two nonempty open sets U and V then we say that these sets ***disconnect*** X.

Example 8.55
If $X = [0,1] \cup [2,3] \subset \mathbf{R}$ then we have seen that both $[0,1]$ and $[2,3]$ are open in X. Since X is their disjoint union, X is disconnected.

The following lemma gives some equivalent ways to formulate the concept of connected space.

Lemma 8.56. Let X be a metric space. Then the following are equivalent:

(1) X is connected.

(2) If $f : X \to \{0,1\}$ is a continuous function then f is constant.

(3) The only subsets of X which are both open and closed are X and $\emptyset$.

(Here the set $\{0,1\}$ is viewed as a metric space via its embedding in $\mathbf{R}$, or equivalently with the discrete metric.)

Proof. □

Frequently one has a metric space X and a subset Y of it whose connectedness or otherwise one wishes to ascertain. To this end, it is useful to record the following lemma.

Lemma 8.57. Let X be a metric space, and let $Y \subseteq X$ be a subset, considered as a metric space with the metric induced from X. Then Y is connected if and only if the following is true: if U, V are open subsets of X, and $U \cap V \cap Y = \emptyset$, then whenever $Y \subseteq U \cup V$, either $Y \subseteq U$ or $Y \subseteq V$.

Proof. □

We now turn to some basic properties of the notion of connectedness. These broadly conform with one's intuition about how connected sets should behave.

Lemma 8.58 (Sunflower lemma)**.** Let X be a metric space. Let $\{Ai \mid i \in I\}$ be a collection of connected subsets of X such that $\bigcap_{i \in I} A_i \neq \emptyset$. Then $\bigcup_{i \in I} A_i$ is connected.

Proof. □

9 Numerical Sequences and Series

§9.1 *Convergent Sequences*

Definition 9.1 (Convergence)**.** Suppose that $(x_n)_{n=1}^{\infty}$ is a sequence of elements of a metric space (X, d). Let $x \in X$. Then we say that $x_n \to x$, or that $\lim_{n\to\infty} x_n = x$, if

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall n \geq N, d(x_n, x) < \epsilon.$$

We call x the ***limit*** of (x_n).

If (x_n) does not converge, it is said to ***diverge***.

> **Exercise 29**
> Show that $\dfrac{1}{n} \to 0$ as $n \to \infty$.

Solution. $\forall \epsilon > 0$, pick $N = \frac{1}{\epsilon} + 1$. Then $\forall n > N$,

$$\frac{1}{n} < \frac{1}{N} = \frac{1}{\frac{1}{\epsilon} + 1} < \frac{1}{\frac{1}{\epsilon}} = \epsilon.$$

□

> **Exercise 30**
> Let (x_n) be a sequence in metric space X, and let $x \in X$. Define what it means for (x_n) to not converge to x.

Solution. Basically negate the definition for convergence:

$$\exists \epsilon > 0 \text{ s.t. } \forall N \in \mathbf{N}, \exists n \geq N \text{ s.t. } d(x_n, x) \geq \epsilon.$$

□

We now outline some important properties of convergent sequences in metric spaces.

Proposition 9.2. Let (x_n) be a sequence in metric space X.

(1) (x_n) converges to $x \in X$ if and only every neighbourhood of x contains x_n for all but finitely many n.

(2) (uniqueness of the limit) If $x \in X$, $x' \in X$, and if (x_n) converges to x and to x', then $x' = x$.

(3) (boundedness of convergent sequences) If (x_n) converges, then (x_n) is bounded.

(4) For $E \subset X$, x is a limit point of E, if and only if there exists a sequence (x_n) in $E \setminus \{x\}$ such that $x_n \to x$.

Proof.

(1) ($\implies$) Suppose $x_n \to x$. We want to prove that any neighbourhood U of x eventually contains all x_n.

Since U is a neighbourhood of x, pick a ball $B_\epsilon(x) \subset U$. Corresponding to this ϵ, there exists $N \in \mathbf{N}$ such that $n \geq N$ implies $d(x_n, x) < \epsilon$. Thus $n \geq N$ implies $x_n \in U$.

($\impliedby$) Suppose every neighbourhood of x contains all but finitely many of the x_n. Fix $\epsilon > 0$, pick a ball $B_\epsilon(x)$. Since $B_\epsilon(x)$ is a neighbourhood of x, it will also eventually contain all x_n. By assumption, there eists $N \in \mathbf{N}$ such that $x_n \in B_\epsilon(x)$ if $n \geq N$. Thus $d(x_n, x) < \epsilon$ if $n \geq N$, hence $x_n \to x$.

(2) Let $\epsilon > 0$ be given. There exists $N, N' \in \mathbf{N}$ such that

$$n \geq N \implies d(x_n, x) < \frac{\epsilon}{2}$$

and

$$n \geq N' \implies d(x_n, x') < \frac{\epsilon}{2}.$$

Take $N_1 \coloneqq \max\{N, N'\}$. Hence if $n \geq N_1$ we have $d(x_n, x) < \frac{\epsilon}{2}$ and $d(x_n, x') < \frac{\epsilon}{2}$ at the same time. By triangle inequality,

$$d(x, x') \leq d(x, x_n) + d(x_n, x') < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Since ϵ was arbitrary (i.e. holds for all $\epsilon > 0$), we must have $d(x, x') = 0$ and thus $x = x'$.

(3) Suppose $x_n \to x$. Then there exists $N \in \mathbf{N}$ such that $n > N$ implies $d(x_n, x) < 1$. Take

$$r \coloneqq \max\{1, d(x_1, x), \ldots, d(x_N, x)\}.$$

Then $d(x_n, x) \leq r$ for $n = 1, 2, \ldots, N$, so (x_n) is in $B_r(x)$.

(4) ($\implies$) If x is a limit point, then for all $\epsilon > 0$, $B_\epsilon \setminus \{x\}(x)$ contains points in E. We then construct such a sequence (x_n) in $E \setminus \{x\}$: pick any $x_n \in E$ so that x_n is contained in $B_{\frac{1}{n}} \setminus \{x\}(x)$. Then it is easy to show that (x_n) is a sequence in $E \setminus \{x\}$ which converges to x.

($\impliedby$) Suppose that there exists a sequence (x_n) in $E \setminus \{x\}$ such that $x_n \to x$. We wish to show that $B_\epsilon \setminus \{x\}(x)$ contains points in E for all $\epsilon > 0$.

Since (x_n) converges to x, for all $\epsilon > 0$ the sequence is eventually contained in $B_\epsilon(x)$. However because we have the precondition that (x_n) has to be in $E \setminus \{x\}$, the sequence is in fact eventually contained in $B_\epsilon \setminus \{x\}(x)$.

□

Lemma 9.3. If (a_n) and (b_n) are two convergent sequences, and $a_n \leq b_n$, then $\lim\limits_{n\to\infty} a_n \leq \lim\limits_{n\to\infty} b_n$.

Remark. Even if you have $a_n < b_n$, you cannot say that $\lim\limits_{n\to\infty} a_n < \lim\limits_{n\to\infty} b_n$. For example, $-\frac{1}{n} < \frac{1}{n}$ but their limits are both 0.

Proof. Let $A = \lim\limits_{n\to\infty} a_n$, $B = \lim\limits_{n\to\infty} b_n$. Suppose otherwise that $A > B$, take $\epsilon = A - B > 0$.

Since $\frac{\epsilon}{2} > 0$, then there exists N_1 such that for $n > N_1$ we have $|a_n - A| < \frac{\epsilon}{2}$; and there exists N_2 such that for $n > N_2$ we have $|b_n - B| < \frac{\epsilon}{2}$.

Let $N = \max\{N_1, N_2\}$, then for any $n > N$, the two inequalities above will hold simultaneously. But then we would have

$$a_n > A - \frac{\epsilon}{2}, \quad b_n < B + \frac{\epsilon}{2}$$

and thus

$$a_n - b_n > A - B - \epsilon = 0$$

so $a_n > b_n$, a contradiction. □

Theorem 9.4 (Sandwich Theorem)**.** Let $a_n \leq c_n \leq b_n$ where $(a_n), (b_n)$ are converging sequences such that $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = L$, then (c_n) is also a converging sequence and $\lim_{n\to\infty} c_n = L$.

Proof. □

Lemma 9.5 (Arithmetic properties)**.** Suppose (a_n) and (b_n) are convergent seqeunces of real numbers, $k \in \mathbf{R}$. Then

(1) Scalar multiplication: $\lim_{n\to\infty} ka_n = k \lim_{n\to\infty} a_n$

(2) Addition: $\lim_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$

(3) Multiplication: $\lim_{n\to\infty} (a_n b_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n$

(4) Division: $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n}$ $(b_n \neq 0,\ \lim_{n\to\infty} b_n \neq 0)$

Proof.

(1) The proof is left as an exercise. You will need to consider three cases, when k is positive, negative or 0.

(2) Let $A = \lim_{n\to\infty} a_n$, $B = \lim_{n\to\infty} b_n$.
$\forall \epsilon > 0$, $\exists N_1 \in \mathbf{N}$, $\forall n > N_1$
$$|a_n - A| < \frac{\epsilon}{2}.$$
$\forall \epsilon > 0$, $\exists N_2 \in \mathbf{N}$, $\forall n > N_2$,
$$|b_n - B| < \frac{\epsilon}{2}.$$
Let $N = \max\{N_1, N_2\}$, then for all $n > N$, by the triangle inequality we have
$$\big|(a_n + b_n) - (A + B)\big| \leq |a_n - A| + |b_n - B| < \epsilon.$$

(3) Let $A = \lim_{n\to\infty} a_n$, $B = \lim_{n\to\infty} b_n$.
Consider the limit $\lim_{n\to\infty} (a_n b_n - AB)$. We want to prove that this equals to 0. We write
$$\lim_{n\to\infty} (a_n b_n - AB) = \lim_{n\to\infty} (a_n b_n - Ab_n + Ab_n - AB)$$
The idea is to show that this is equal to
$$\lim_{n\to\infty} (a_n b_n - Ab_n) + \lim_{n\to\infty} (Ab_n - AB)$$
(Note that we cannot write this yet because we have not shown that these two sequences are convergent)
So let's examine these two sequences. The second one is easier since we have proved addition:
$$\lim_{n\to\infty} b_n = B \implies \lim_{n\to\infty} (b_n - B) = 0$$
Thus $\lim_{n\to\infty} (Ab_n - AB) = A \lim_{n\to\infty} (b_n - B) = 0$.
As for the first one, we want to show that $\lim_{n\to\infty} (a_n - A)b_n = 0$. Since b_n is convergent, b_n is bounded. Let $M > 0$ be a bound of b_n: $\forall n \in \mathbf{N}$,
$$|b_n| \leq M.$$
Since $\lim_{n\to\infty} a_n = a$, $\forall \epsilon > 0 \exists N \in \mathbf{N}$ s.t. $\forall n > N$,
$$|a_n - a| < \frac{\epsilon}{M}.$$

Combining the two above, we then conclude that $\forall \epsilon > 0 \exists N \in \mathbf{N}$ s.t. $\forall n > N$,

$$|a_n b_n - A b_n| = |(a_n - A) b_n| < \frac{\epsilon}{M} \cdot M = \epsilon.$$

Thus $\lim_{n\to\infty} (a_n b_n - A b_n) = 0$.

Since $\lim_{n\to\infty} (A b_n - AB) = 0$ and $\lim_{n\to\infty} (a_n b_n - A b_n) = 0$, we can conclude that $\lim_{n\to\infty} (a_n b_n - AB) = 0$.

(4) Since we have proven multiplication, it suffices to show that $\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{\lim_{n\to\infty} b_n}$.

Let $b = \lim_{n\to\infty} b_n$. Consider the limit

$$\lim_{n\to\infty} \left(\frac{1}{b_n} - \frac{1}{b} \right) = \lim_{n\to\infty} \left(\frac{b - b_n}{b_n b} \right).$$

Again, the important term here is $b - b_n$, but there is an extra term of $\frac{1}{b_n b}$, so we'll need to control this.

Since we need this to be bounded, we actually cannot have b_n to be close to 0. The good thing here is that $b \neq 0$, so we can restrict b_n to be close enough to b so that it stays away from 0.

Pick N_1 such that for all $n > N_1$,

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

Then

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

$$\frac{b^2}{2} < b_n b < \frac{3b^2}{2}$$

This show that if $n > N_1$, $b_n b$ would always be positive, and $\frac{1}{b_n b} < \frac{2}{b^2}$.

Let $M = \frac{2}{b^2}$, then we may refer back to the original statement

$$\left| \frac{b - b_n}{b_n b} \right| < M |b - b_n|$$

We pick N_2 such that for all $n > N_2$, $|b_n - b| < \frac{\epsilon}{M}$.

Let $N = \max\{N_1, N_2\}$, then for all $n > N$,

$$\left| \frac{b - b_n}{b_n b} \right| < M \cdot \frac{\epsilon}{M} = \epsilon.$$

□

Exercise 31

Let (x_n) be a sequence of real numbers and let $\alpha \geq 2$ be a constant. Define the sequence (y_n) as follows:

$$y_n = x_n + \alpha x_{n+1} \quad (n = 1, 2, \dots)$$

Show that if (y_n) is convergent, then (x_n) is also convergent.

Exercise 32

(1) $\lim_{n\to\infty} \frac{1}{n_p} = 0$ $(p > 0)$.

(2) $\lim_{n\to\infty} \sqrt[n]{p} = 1$ $(p > 0)$.

(3) $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

(4) $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$ $(p > 0,\ \alpha \in \mathbf{R})$.

(5) $\lim_{n\to\infty} x^n = 0$ $(|x| < 1)$.

§9.2 *Subsequences*

Definition 9.6 (Subsequence)**.** Given a sequence (x_n), consider a sequence (n_k) of positive integers such that $n_1 < n_2 < \cdots$. Then (x_{n_i}) is called a ***subsequence*** of (x_n). If (x_{n_i}) converges, its limit is called a ***subsequential limit*** of (x_n).

Proposition 9.7. (x_n) converges to x if and only if every subsequence of (x_n) converges to x.

Proof. Suppose (x_n) converges to x. Then $\forall \epsilon > 0$, $\exists N \in \mathbf{N}$, $\forall n > N$, $d(x_n, x) < \epsilon$. Every subsequence of (x_n) can be written in the form (x_{n_i}) where $n_1 < n_2 < \cdots$ is a strictly increasing sequence of positive integers. Pick M such that $n_M > N$, then $\forall i > M$, $d(x_{n_i}, x) < \epsilon$. Hence every subsequence of (x_n) converges to x.

Intuitively, if every neighbourhood of x eventually contains all x_n, then since (x_{n_i}) is a subset of (x_n) they should all be contained in the neighbourhood eventually as well. □

Proposition 9.8. Subsequential limits of a sequence are precisely the limit points of the sequence (viewed as a set)

Proof. This is just part (d) of the previous section.

Again, to make this work, we need to assume that nothing funny is going on at subsequential limits If the limits appear due to eventually constant subsequences, then they need not be limit points of the original sequence when viewed as a set

3.6, 3.7 are precisely the statements we've prepared for last week □

Proposition 9.9. If (x_n) is a sequence in a compact set (bounded closed set), then there exists a convergent subsequence of (x_n) (which converges to some number in the set).

Proof. This is Bolzano–Weierstrass together with part (b)

Essentially, compact sets satisfies the property akin to the statement in Heine-Borel: Given a topological space (X, τ), a compact set K in X is a set satisfying that, given any open covering $\{U_i\}$ of X, there exists a finite open cover $\{U_1, \ldots, U_n\}$ of X

This is difficult to process at this stage Since we're currently only working with Euclidean spaces it would be more beneficial if you consider the Heine-Borel Theorem as a property first It would be a lot easier to accept the definition after you're more accustomed to applying the theorem □

Proposition 9.10. The subsequential limits of a sequence (x_n) in metric space X form a closed subset of X.

Proof. Let E be the set of all subsequential limits of (x_n), and let q be a limit point of E. We want to show that $q \in E$.

Choose n_1 so that $x_{n_1} \neq q$. (If no such n_1 exists, then E has only one point, and there is nothing to prove.) Put $\delta = d(q, x_{n_1})$. Suppose $n_1, \ldots, n_{i-1}$ are chosen. Since q is a limit point of E, there is an $x \in E$ with $d(x, q) < 2^{-1}\delta$. Since $x \in E$, there is an $n_i > n_{i-1}$ such that $d(x, x_{n_i}) < 2^{-i}\delta$. Thus

$$d(q, x_{n_i}) < 2^{1-i}\delta$$

for $i = 1, 2, 3, \ldots$. This says that (x_{n_i}) converges to q. Hence $q \in E$. □

§9.3 *Cauchy Sequences*

This is a very helpful way to determine whether a sequence is convergent or divergent, as it does not require the limit to be known. In the future you will see many instances where the convergence of all sorts of limits are compared with similar counterparts; generally we describe such properties as ***Cauchy criteria***.

Definition 9.11 (Cauchy sequence)**.** A sequence (x_n) in a metric space X is said to be a ***Cauchy sequence*** if

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall n, m \geq N, d(x_n, x_m) < \epsilon.$$

Remark. This simply means that the distances between any two terms is sufficiently small after a certain point.

It is easy to prove that a converging sequence is Cauchy using the triangle inequality. The idea is that, if all the points are becoming arbitrarily close to a given point x, then they are also becoming close to each other. The converse is not always true, however.

Proposition 9.12. A sequence (x_n) in $\mathbf{R}^n$ is convergent if and only if it is Cauchy.

Proof.

($\implies$) Suppose that (x_n) converges to x, then there exists $N \in \mathbf{N}$ such that $\forall n > N$, $|x_n - x| < \dfrac{\epsilon}{2}$. Then for $n, m > N$, by triangle inequality,

$$|x_n - x_m| \leq |x_n - x| + |x_m - x| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence (x_n) is a Cauchy sequence.

($\impliedby$) First, we show that (x_n) must be bounded. Pick $N \in \mathbf{N}$ such that $\forall n, m > N$ we have $|x_n - x_m| < 1$. Centered at x_n, we show that (x_n) is bounded; to do this we pick

$$r = \max\{1, |x_n - x_1|, \ldots, |x_n - x_N|\}.$$

Then the sequence x_n is in $B_r(x_n)$ and thus is bounded.

Since (x_n) is bounded, by the corollary of Bolzano–Weierstrass we know that (x_n) contains a subsequence (x_{n_i}) that converges to x.

Then $\forall \epsilon > 0$, pick $N_1 \in \mathbf{N}$ such that for all $n, m > N$, $|x_n - x_m| < \dfrac{\epsilon}{2}$.

Simultaneously, since $\{x_{n_i}\}$ converges to x, pick M such that for $i > M$, $|x_{n_i} - x| < \dfrac{\epsilon}{2}$.

Now, since $n_1 < n_2 < \cdots$ is a sequence of strictly increasing natural numbers, we can pick $i > M$ such that $n_i > N$. Then $\forall n > N$, by setting $m = n_i$ we obtain

$$|x_n - x_{n_i}| < \frac{\epsilon}{2}, \quad |x_{n_i} - x| < \frac{\epsilon}{2}$$

and hence

$$|x_n - x| \leq |x_n - x_{n_i}| + |x_{n_i} - x| < \epsilon$$

by triangle inequality. Hence (x_n) is convergent. □

Definition 9.13. Let nonempty $E \subseteq X$. Let S be the set of all real numbers of the form $d(x, y)$, with $x, y \in E$. Then the ***diameter*** of E is

$$\operatorname{diam} E \coloneqq \sup S.$$

Definition 9.14. A sequence (x_n) of real number is said to be

(i) ***monotonically increasing*** if $x_n \leq x_{n+1}$ $(n = 1, 2, \ldots)$;

(ii) ***monotonically decreasing*** if $x_n \geq x_{n+1}$ $(n = 1, 2, \ldots)$.

The class of monotonic sequences consists of the increasing and decreasing sequences.

Lemma 9.15. Suppose (x_n) is monotonic. Then (x_n) converges if and only if it is bounded.

Proof. Suppose $x_n \leq x_{n+1}$ (the proof is analogous in the other case). Let E be the range of (x_n). Suppose (x_n) is bounded, then let $x = \sup E$. Then

$$x_n \leq x \quad (n = 1, 2, \dots)$$

... □

§9.4 *Upper and Lower Limits*

Definition 9.16. Let (x_n) be a real sequence with the following property: $\forall M \in \mathbf{R} \exists N \in \mathbf{N}$ s.t. $n \geq N \implies x_n \geq M$. We then write $x_n \to \infty$.

Similarly, if $\forall M \in \mathbf{R} \exists N \in \mathbf{N}$ s.t. $n \geq N \implies x_n \leq M$, we write $x_n \to \infty$.

Definition 9.17 (Upper and lower limits)**.** Let (x_n) be a real sequence. Let E be the set of real numbers x (in the extended real number system) such that $x_{n_k} \to x$ for some subsequence (x_{n_k}). This set E contains all subsequential limits, plus possibly the numbers $+\infty$ and $-\infty$.

Put $x^* = \sup E$, $x_* = \inf E$. The numbers x^* and x_* are called the ***upper limit*** and ***lower limit*** of (x_n), denoted by

$$\limsup_{n\to\infty} x_n = x^*, \quad \liminf_{n\to\infty} x_n = x_*.$$

Proposition 9.18. Let (x_n) be a real sequence. Let E and x^* have the same meaning as in Definition 9.17. Then x^* has the following two properties:

(1) $x^* \in E$.

(2) If $x > x^*$, there exists $N \in \mathbf{N}$ such that $n \geq N$ implies $x_n < x$.

Moreover, x^* is the only number with the properties (1) and (2).

Example 9.19

- Let (x_n) be a sequence containing all rationals. Then every real number is a subsequential limit, and $\limsup_{n\to\infty} x_n = +\infty$, $\liminf_{n\to\infty} = -\infty$.
- For a real-valued seqeunce (x_n), $\lim_{n\to\infty} x_n = x$ if and only if $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = x$.

Lemma 9.20. If $a_n \leq b_n$ for $n \geq N$ where N is fixed, then

$$\liminf_{n\to\infty} a_n \leq \liminf_{n\to\infty} b_n,$$
$$\limsup_{n\to\infty} a_n \leq \limsup_{n\to\infty} b_n.$$

Lemma 9.21 (Arithmetic properties)**.**

(1) If $k > 0$, $\limsup_{n\to\infty} ka_n = k \limsup_{n\to\infty} a_n$.
If $k < 0$, $\limsup_{n\to\infty} ka_n = k \liminf_{n\to\infty} a_n$.

(2) $\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$
Moreover, $\limsup(a_n + b_n)$ may be bounded from below as follows:

$$\limsup(a_n + b_n) \geq \limsup a_n + \liminf b_n$$

Your homework for today is to write down the analogous properties for liminf, and to prove (i) and (ii)

Now you should try to prove (i) for liminf as well; as for (ii), try to explain why properties (i),(ii) for limsup and property (i) for liminf would imply property (ii) for lim inf

§9.5 *Series*

Definition 9.22 (Series)**.** Given a sequence (a_n) in metric space X, we associate a sequence (S_n), where

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

which we call a ***series***. S_n is the ***n-th partial sum*** of the series.

We say that the (infinite) series ***converges*** if the sequence of partial sums (S_n) converges. We then define the ***sum*** of a convergent infinite series to be the limit of the convergent sequence (S_n); that is, given $S \in X$, $\sum_{n=1}^{\infty} a_n = S$ if

$$\forall \epsilon > 0, \exists N \in \mathbf{N}, \forall n > N, \left| \sum_{k=1}^{n} a_k - S \right| < \epsilon.$$

If (S_n) diverges, the series is said to ***diverge***.

3.22

Proposition 9.23 (Cauchy criterion)**.** $\sum_{n=1}^{\infty} a_n$ converges if and only if $\forall \epsilon > 0$, $\exists N \in \mathbf{N}$ such that $\forall m \geq n > N$,

$$\left| \sum_{k=m}^{n} a_k \right| < \epsilon.$$

Corollary 9.24. If $\sum_{n=1}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

Remark. The converse is not true; we have the very well known counterexample of the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$.

Now we talk about various methods to determine whether an infinite series converges or diverges.

Lemma 9.25 (Comparison test)**.** We consider two sequences (a_n) and (b_n).

(1) Suppose $|a_n| \leq b_n$ for all n (or for all sufficiently large n), if $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.

(2) Suppose $a_n \geq b_n \geq 0$ for all n (or for all sufficiently large n), if $\sum_{n=1}^{\infty} b_n$ diverges, then $\sum_{n=1}^{\infty} a_n$ diverges.

Refer to 3.25 for the proof

Lemma 9.26 (Root test)**.** Given $\sum a_n$, put $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Then

(i) if $\alpha < 1$, $\sum a_n$ converges;

(ii) if $\alpha > 1$, $\sum a_n$ diverges;

(iii) if $\alpha = 1$, the rest gives no information.

Lemma 9.27 (Ratio test)**.** The series $\sum a_n$

(i) converges if $\limsup_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| < 1$;

(ii) diverges if $\left| \frac{a_{n+1}}{a_n} \right| \geq 1$ for all $n \geq n_0$, where n_0 is some fixed integer.

The series $\sum a_n$ is said to ***converge absolutely*** if the series $\sum |a_n|$ converges.

Lemma 9.28. If $\sum a_n$ converges absolutely, then $\sum a_n$ converges.

Definition 9.29.

$$e := \sum_{n=0}^{\infty} \frac{1}{n!}$$

Lemma 9.30.

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Proof. Let

$$s_n = \sum_{k=0}^{n} \frac{1}{k!}, \quad t_n = \left(1 + \frac{1}{n}\right)^n.$$

Expanding t_n using the binomial theorem, and comparing s_n and t_n term by term, we have $t_n \leq s_n$, so

$$\limsup_{n\to\infty} t_n \leq e.$$

□

Proposition 9.31. e is irrational.

Proof. Suppose e is rational. Then $e = \frac{p}{q}$, where □

10 Continuity

§10.1 *Limit of Functions*

Assume (X, d_X) is metric space and $E \subset X$ is a subset of X. Then the metric d_X induces a metric on E. We now consider another metric space (Y, d_Y). A map $f : E \to Y$ is also called a function over E with values in Y. In particular, if $Y = \mathbf{R}$, then f is called a real-valued function; and if $Y = \mathbf{C}$, f is called a complex-valued function.

Definition 10.1 (Limit of function)**.** Consider a limit point $p \in E$ and a point $q \in Y$. We say the ***limit*** of the funcion $f(x)$ at p is q, denoted by $\lim_{x\to p} f(x) = q$, if

$$\forall \epsilon > 0, \exists \delta > 0 \text{ s.t. } \forall x \in E, 0 < d_X(x, p) < \delta, d_Y\left(f(x), q\right) < \epsilon.$$

We can recast this definition in terms of limits of sequences:

Proposition 10.2. Let X, Y, E, f, p be as in Definition 10.1. Then $\lim\limits_{x\to p} f(x) = q$ if and only if

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

for every sequence $\{p_n\}$ in E such that $p_n \neq p$ and $\lim\limits_{n\to\infty} p_n = p$.

Proof.

$(\implies)$ Suppose $\lim\limits_{x\to p} f(x) = q$. Choose $\{p_n\}$ in E satisfying $p_n \neq p$ and $\lim\limits_{n\to\infty} p_n = p$.

Let $\epsilon > 0$ be given. Then there exists $\delta > 0$ such that $d_Y\left(f(x), q\right) < \epsilon$ if $x \in E$ and $0 < d_X(x, p) < \delta$.

Also, there exists $N \in \mathbf{N}$ such that $n > N$ implies $0 < d_X(p_n, p) < \delta$. Thus for $n > N$, we have $d_Y\left(f(p_n), q\right) < \epsilon$, which shows that $\lim\limits_{n\to\infty} f(p_n) = q$.

$(\impliedby)$ □

By the same proofs as for sequences, limits are unique, and in $\mathbf{R}$ they add/multiply/divide as expected.

Definition 10.3. f is ***continuous*** at p if

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

In the case where p is not a limit point of the domain E, we say f is continuous at p. If f is continuous at all points of E, then we say f is continuous on E.

The sequential definition of continuity follows almost directly from the sequential definition of limits: f is continuous at p if for every sequence x_n converging to p, the sequence $f(x_n)$ converges to $f(p)$.

§10.2 *Continuous Functions*

Consider metric spaces (X, d_X) and (Y, d_Y), $U \subset X$.

Definition 10.4 (Continuity)**.** Let (X, d_X) and (Y, d_Y) be metric spaces, and $U \subseteq X$. We say that $f : U \to Y$ is ***continuous*** at $x_0 \in U$, if

$$\forall \epsilon > 0 \exists \delta > 0 \text{ s.t. } \forall x \in X, d_X(x, x_0) < \delta \implies d_Y\left(f(x), f(x_0)\right).$$

We say f is continuous in U if it is continuous at every $x_0 \in X$.

As for functions on the reals, one may also phrase the definition of continuity in terms of limits.

Lemma 10.5. Let $f : X \to Y$ be a function between metric spaces. Then f is continuous at a if and only if the following is true: for any sequence $(x_n)_{n=1}^{\infty}$ with $\lim_{n\to\infty} x_n = a$, we have $\lim_{n\to\infty} f(x_n) = f(a)$.

Proof. ($\implies$)

($\impliedby$) □

Definition 10.6 (Uniform continuity)**.** Let (X, d_X) and (Y, d_Y) be metric spaces, and $U \subseteq X$. We say that $f : U \to Y$ is ***uniformly continuous*** if

$$\forall \epsilon > 0 \exists \delta > 0 \text{ s.t. } \forall x, y \in U, d_X(x, y) < \delta \implies d_Y\left(f(x), f(y)\right) < \epsilon.$$

§10.2.1 *Continuity of linear functions in normed spaces*

A great deal of power comes from considering the set of all functions on a space satisfying some property, such as continuity, as a metric space in its own right. In this section we consider some important examples of such spaces.

We begin with the space of bounded real-valued functions on a set X. At this stage we assume nothing about X.

Definition 10.7 (Space of bounded real-valued functions)**.** If X is any set, we define $B(X)$ to be the space of functions $f : X \to \mathbf{R}$ for which $f(X) = \{f(x) \mid x \in X\}$ is bounded. If $f \in B(X)$, define $\|f\|_\infty = \sup_{x\in X} |f(x)|$.

Lemma 10.8. For any set X, $B(X)$ is a vector space, and $\|\cdot\|_\infty$ is a norm.

Proof. □

Now we turn to the space of continuous real-valued functions, $C(X)$. To make sense of what this means we now need X to be a metric space.

Definition 10.9. Let X be a metric space. We write $C(X)$ for the space of all continuous functions $f : X \to \mathbf{R}$.

§10.3 *Continuity and Compactness*

Assume (X, d_X) and (Y, d_Y) are metric spaces.

Theorem 10.10. Assume $f : X \to Y$ is a continuous map. Then for any compact subset $K \subset X$, the image set $f(K)$ is a compact subset of Y.

Proof. We prove it by definition. Assume $\{V_i \mid i \in I\}$ is an open cover of $f(K)$. By the continuity of f and □

§10.4 *Continuity and Connectedness*

Proposition 10.11. If f is a continous mapping of a metric space X into a metric space Y, and if E is a connected subset of X, then $f(E)$ is connected.

Proof. □

Theorem 10.12 (Intermediate value theorem)**.** Let $f : [a, b] \to \mathbf{R}$ be continuous. If $f(a) < f(b)$ and $f(a) < c < f(b)$, then $\exists x \in (a, b)$ s.t. $f(x) = c$.

Proof. □

§10.5 *Discontinuities*

Let $f : X \to Y$. If f is not continuous at $x \in X$, we say that f is discontinuous at x, or that f has a discontinuity at x.

If f is defined on an interval or a segment, it is customary to divide discontinuities into two types. Before giving this classification, we have to define the ***right-hand*** and the ***left-hand limits*** of f at x, denoted by $f(x+)$ and $f(x-)$ respectively.

Definition 10.13 (Right-hand and left-hand limits)**.** Let $f : (a, b) \to \mathbf{R}$. Consider any point x such that $a \leq x < b$.

Definition 10.14 (Discontinuities)**.** Let $f : [a, b] \to \mathbf{R}$. If f is discontinuous at x, and if $f(x+)$ and $f(x-)$ exist, then f is said to have a ***discontinuity of the first kind***, or a ***simple discontinuity***, at x. Otherwise the discontinuity is said to be of the ***second kind***.

There are two ways in which a function can have a simple discontinuity: either

§10.6 *Monotonic Functions*

Proposition 10.15. Let $f : [a, b] \to \mathbf{R}$ be monotonically increasing. Then $f(x+)$ and $f(x-)$ exist for all $x \in (a, b)$; more precisely,

$$\sup_{t \in (a,x)} f(t) = f(x-) \leq f(x) \leq f(x+) = \inf_{t \in (x,b)} f(t).$$

Furthermore, if $a < x < y < b$, then

$$f(x+) \leq f(y-).$$

Analogous results evidently hold for monotically decreasing functions.

§10.7 *Infinite Limits and Limits at Infinity*

Definition 10.16. For $c \in \mathbf{R}$, the set $\{x \in \mathbf{R} \mid x > c\}$ is called a neighbourhood of $+\infty$ and is written $(c, +\infty)$. Similarly, the set $(-\infty, c)$ is a neighbourhood of $-\infty$.

Definition 10.17. Let $f : E \subset \mathbf{R} \to \mathbf{R}$. We say that $\lim_{t \to x} f(t) = A$ where A and x are in the extended real number system, if for every neighbourhood of U of A there is a neighbourhood V of x such that $V \cap E$ is not empty, and such that $f(t) \in U$ for all $t \in V \cap E$, $t \neq x$.

11 Differentiation

§11.1 *The Derivative of A Real Function*

Definition 11.1. Suppose $f : [a, b] \to \mathbf{R}$. For any $x \in [a, b]$, we form the quotient

$$\phi(t) = \frac{f(t) - f(x)}{t - x} \quad (a < t < b, t \neq x)$$

and define

$$f'(x) = \lim_{t \to x} \phi(t),$$

provided this limit exists. f' is called the ***derivative*** of f.

If f' is defined at a point x, we say that f is ***differentiable*** at x; If f' is defined at every point of a set $E \subseteq [a, b]$, we say that f is differentiable on E.

Lemma 11.2. If $f : [a, b] \to \mathbf{R}$ is differentiable at $x \in [a, b]$, then f is continuous at x.

Proof.

$$\lim_{t \to x}[f(t) - f(x)] = \lim_{t \to x}\left[\frac{f(t) - f(x)}{t - x} \cdot (t - x)\right] = \lim_{t \to x} \frac{f(t) - f(x)}{t - x} \cdot \lim_{t \to x}(t - x) = f'(x) \cdot 0 = 0.$$

Since $\lim_{t \to x} f(t) = f(x)$, f is continuous at x. □

Remark. The converse of this theorem is not true. It is easy to construct continuous functions which fail to be differentiable at isolated points.

Notation. We use $C_1[a, b]$ to denote the set of differentiable functions over $[a, b]$ whose derivative is continuous. More generally, we use $C_k[a, b]$ to denote the set of functions whose k-th ordered derivative is continuous. In particular, $C_0[a, b]$ is the set of continuous functions over $[a, b]$.

Later on when we talk about properties of differentiation such as the intermediate value theorems, we usually have the following requirement on the function:

> f is a continuous function on $[a, b]$ which is differentiable in (a, b).

Lemma 11.3 (Differentiation rules)**.** Suppose $f, g : [a, b] \to \mathbf{R}$ are differentiable at $x \in [a, b]$. Then $f \pm g$, fg and f/g (when $g(x) \neq 0$) are differentiable at x. Moreover,

(1) $(f \pm g)'(x) = f'(x) \pm g'(x)$;

(2) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$;

(3) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$

Proof. We take (2) as an example.

We calculate

$$\begin{aligned}\frac{f(x)g(x)-f(x_0)g(x_0)}{x-x_0} &= \frac{\big(f(x)-f(x_0)\big)\,g(x)+f(x_0)\,\big(g(x)-g(x_0)\big)}{x-x_0}\\ &= \frac{f(x)-f(x_0)}{x-x_0}\cdot g(x)+f(x_0)\cdot\frac{g(x)-g(x_0)}{x-x_0}\\ &\to f'(x_0)g(x_0)+f(x_0)g'(x_0) \text{ as } x\to x_0\end{aligned}$$

where we use f and g are differentiable at x_0 and Lemma 11.2. □

Theorem 11.4 (Chain rule)**.** Let $f:[a,b]\to\mathbf{R}$ be a real-valued function that is differentiable at $x_0\in[a,b]$. Let g be a real-valued function defined on an interval that contains $f([a,b])$, and g is differentiable at $f(x_0)$. Then the composition

$$h(x)\coloneqq g\circ f(x)\coloneqq g\big(f(x)\big):[a,b]\to\mathbf{R}$$

is differentiable at x_0 and the derivative at x_0 can be calculated as

$$h'(x_0)=g'\big(f(x_0)\big)\,f'(x_0).$$

Proof. We know that

$$f'(x)=\lim_{t\to x}\frac{f(t)-f(x)}{t-x},$$

so under the assumption that t stays within the domain of f, $\frac{f(t)-f(x)}{t-x}$ should be a good approximation to $f'(x)$.

To actually quantify this, let $u(t)=\frac{f(t)-f(x)}{t-x}-f'(x)$.

Then the differentiability of f tells us that $\lim_{t\to x}u(t)=0$.

Similarly, let $v(s)=\frac{g(s)-g(y)}{s-y}-g'(y)$, then $\lim_{s\to y}v(s)=0$, as long as s stays in the domain of g

What's nice here is that we can let $s=f(t)$, then by our assumption s always stays in the domain of g, so nothing fishy will happen

Ah I forgot a small detail here Additionally we also need to define u(x)=0 and v(y)=0

Now let $h(t)=g(f(t))$, then h is defined on $[a,b]$, and we deduce that

$$h(t)-h(x)=(t-x)[f'(x)+u(t)][g'(y)+v(s)]$$

We then check that

$$\lim_{t\to x}\frac{h(t)-h(x)}{t-x}=\lim_{t\to x}[f'(x)+u(t)][g'(y)+v(s)]=f'(x)g'(f(x))$$

and we are done. □

Example 11.5

One of the best (worst?) family of pathological examples in calculus are functions of the form

$$f(x)=x^p\sin\frac{1}{x}.$$

- For $p=1$, the function is continuous and differentiable everywhere other than $x=0$.
- For $p=2$, the function is differentiable everywhere, but the derivative is discontinuous.

Other more advanced pathological results (just for fun):

- The graph for $y=\sin\frac{1}{x}$ on $(0,1]$, together with the interval $[-1,1]$ on the y-axis, is a connected closed set that is not path-connected.

- For $0 < p < 1$, we obtain functions that are continuous and bounded, but the graphs are of infinite length (ps. I think that this is also true for $p = 1$).

Regarding continuous but not differentiable functions, a more pathological example is the Weierstrass function, which is continuous everywhere over $\mathbf{R}$ but differentiable nowhere.

§11.2 *Mean Value Theorems*

Definition 11.6. Let f be a real valued function defined over a metric space X. We say f has a ***local maximum*** at $x_0 \in X$ if $\exists \delta > 0$ s.t. $\forall x \in B_\delta(x_0)$,

$$f(x_0) \geq f(x).$$

Similarly, we say f has ***local minimum*** at $x_0 \in X$ if $\exists \delta > 0$ s.t. $\forall x \in B_\delta(x_0)$,

$$f(x_0) \leq f(x).$$

Definition 11.7. For a function $f : (a, b) \to \mathbf{R}$, a point $x_0 \in [a, b]$ is called a ***critical point*** if f is not differentiable at x_0 or $f'(x_0) = 0$.

Theorem 11.8. Assume f is defined over $[a, b]$. If f has a local maximum or local minimum at some $x_0 \in (a, b)$, then x_0 is a critical point of f.

Proof. If f is not differentiable at x_0, we are done. Assume now f is differentiable at x_0 and x_0 is a local maximum.

Then $\exists \delta > 0$ s.t. $\forall x \in B_\delta(x_0)$,

$$f(x_0) \leq f(x).$$

It follows

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \geq 0 & x_0 - \delta < x < x + \delta \\ \leq 0 & x_0 < x < x_0 + \delta \end{cases}$$

Further since $f'(x_0)$ exists, there is

$$f'(x_0-) \geq 0, \quad f'(x_0+) \leq 0,$$

but $f'(x_0-) = f'(x_0+) = f'(x_0)$. Hence $f'(x_0) = 0$. □

Theorem 11.9 (Fermat's Theorem (Interior Extremum Theorem))**.** If the differential exists, then by comparing the left and right limits it is easy to see that the differential for a local maximum/maximum can only be 0.

To summarize in four words: Local extrema are stationary

There are three mean value theorems, from specific to general:

1. Rolle's Theorem
2. (Lagrange's) Mean Value Theorem
3. Generalised (Cauchy's) Mean Value Theorem

Theorem 11.10 (Rolle's Theorem)**.** If f is continuous on $[a, b]$, differentiable in (a, b) and $f(a) = f(b)$, then there exists $c \in (a, b)$ such that

$$f'(c) = 0.$$

Proof. Let $h(x)$ be a function defined on $[a, b]$ where $h(a) = h(b)$.

The idea is to show that h has a local maximum/minimum, then by Fermat's Theorem this will then be the stationary point that we're trying to find.

First note that h is continuous on $[a, b]$, so h must have a maximum M and a minimum m.

If M and m were both equal to $h(a) = h(b)$, then h is just a constant function and so $h'(x) = 0$ everywhere.

Otherwise, h has a maximum/minimum that is not $h(a) = h(b)$, so this extremal point lies in (a, b).

In particular, this extremal point is also a local extremum. Since h is differentiable on (a, b), by Fermat's theorem this extremum point is stationary, thus Rolle's Theorem is proven. □

Theorem 11.11 (Mean Value Theorem)**.** If f is continuous on $[a, b]$ and differentiable in (a, b), then there exists $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Exercise 2: Show that the Mean Value Theorem results directly from Rolle's Theorem (the other direction is trivial) : This isn't a very significant exercise because we're going to prove something more general

Theorem 11.12 (Generalised Mean Value Theorem)**.** If f and g are continuous on $[a, b]$ and differentiable in (a, b), then there exists $c \in (a, b)$ such that

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Now we return to the proof of the generalized MVT

We set the function $h(t) = [f(b) - f(a)]g(t) - [g(b) - g(a)]f(t)$, then h is continuous on $[a, b]$ and differentiable on (a, b)

Moreover, $h(a) = f(b)g(a) - f(a)g(b) = h(b)$, thus by Rolle's Theorem, there exists $c \in (a, b)$ such that $h'(c) = 0$, i.e. $[g(b) - g(a)]f'(c) = [f(b) - f(a)]g'(c)$

Corollary: If f and g are continuous on $[a, b]$ and differentiable in (a, b), and $g'(x) \neq 0$ for all $x \in (a, b)$, then there exists

$$c \in (a, b) \text{ s.t. } f'(c)/g'(c) = [f(b) - f(a)]/[g(b) - g(a)]$$

This form of the generalized MVT will be used to prove the most beloved rule of high school students

exercises for the Mean Value Theorem

Exercise 33

Let f and g be continuous on $[a, b]$ and differentiable on (a, b). If $f'(x) = g'(x)$, then $f(x) = g(x) + C$.

Exercise 34

Given that $f(x) = x^\alpha$ where $0 < \alpha < 1$. Prove that f is uniformly continuous on $[0, +\infty)$.

Exercise 35 (Olympiad level)

Let f be a function continuous on $[0, 1]$ and differentiable on $(0, 1)$ where $f(0) = f(1) = 0$. Prove that there exists $c \in (0, 1)$ such that

$$f(x) + f'(x) = 0.$$

§11.3 *Darboux's Theorem*

Darboux's Theorem implies some sort of a 'intermediate value' property of derivatives that is similar to continuous functions

This is Theorem 5.12 in the book

Now first and foremost, the requirement for this statement is that f must be differentiable on [a,b], not just in (a,b) Otherwise f'(a) and f'(b) may not make sense : One common theme in many of these problems is to construct auxiliary functions Suppose that $f'(a) < \lambda < f'(b)$, then we construct the

auxiliary function $g(x) = f(x) - \lambda x$: Then we only need to find a point $x \in (a, b)$ such that g'(x)=0 : This means that we only need to find a local maximum/minimum, which by Fermat's Theorem has to be a stationary point as well : Now we look at the values of g near a and b : Exercise 1: Using the fact that $g'(a) < 0$ and $g'(b) > 0$, show that a and b are local maxima of g

Here we regard g as simply a function on $[a, b]$, so we only need to show that a,b are maximum and corresponding semi-open neighbourhoods $[a, a + \epsilon)$ and $(b - \epsilon, b]$: Let m=g'(a)<0 be the slope of the tangent at a : Then lim(h→0+)[g(a+h)-g(a)]/h=m<0 : This means that there should exist $\delta > 0$ such that for $0 < h < \delta$, [g(a+h)-g(a)]/h<m/2<0 : Now we can rewrite the above as g(a+h)<g(a)+mh/2 : Since m<0 and h>0, we obtain $g(a + h) < g(a)$ for $0 < h < \delta$: Thus this proves that x=a is a local maximum of g A similar proof applies for x=b : Now since g is differentiable on [a,b], in particular it has to be continuous on [a,b] : Since [a,b] is compact, g([a,b]) is compact in R and thus g has both maximum and minimum values in [a,b] : Here we'll just focus on the minimum value : As we've shown, x=a is a 'strict' local maxima, in the sense that for any point $x \in (a, a + \epsilon)$, we actually have the strict inequality $g(x) < g(a)$: This means that x=a cannot be a local minimum : Similarly, x=b cannot be a local minimum, and therefore g achieves its minimum strictly inside (a,b) : Only then we can say that this local minimum is stationary (This will not work otherwise; note that a and b are both local maxima but are not stationary points of g) : An interesting implication of Darboux's Theorem is that if f is differentiable on [a,b], then f' cannot have simple discontinuities (removable or jump discontinuities), simply because these discontinuities do not allow this 'intermediate value' property : However, we should recall certain pathological examples like f(x)=x² sin 1/x (f(0)=0) Here f'(0)=lim(h→0)[x² sin 1/x-0]/x=0, but f'(x)=2x sin 1/x - cos 1/x, so f' is discontinuous at x=0

§11.4 *L'Hopital's Rule*

Theorem 11.13 (L'Hopital's Rule)**.** Assume f, g are differentiable over (a, b) with $g(x) \neq 0$. If either

(1) $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$; or

(2) $\lim_{x\to a} |g(x)| = +\infty$,

and

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = A \in [-\infty, +\infty]$$

assuming $g'(x) \neq 0$ over (a, b), then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = A.$$

Proof. Now the entire proof is quite tedious because there's actually eight main cases to think of 1. $\frac{0}{0}$ or $\frac{\infty}{\infty}$ 2. a is normal or $a = -\infty$ 3. A is normal or $A = \pm\infty$

We'll only prove the most basic one here: 0/0, a and A are normal This is the case which will be required for Taylor series

First we define f(a)=g(a)=0, so that f and g are continuous at $x = a$

Now let $x \in (a, b)$, then f and g are continuous on $[a, x]$ and differentiable in (a, x) : Thus by Cauchy's Mean Value Theorem, there exists $\xi \in (a, x)$ such that

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x)}{g(x)}$$

For each x, we pick ξ which satisfies the above, so that ξ may be seen as a function of x satisfying $a < \xi(x) < x$

Then by squeezing we have $\lim_{x\to a^+} \xi(x) = a$.

Since $\frac{f'}{g'}$ is continuous near a, the theorem regarding the limit of composite functions give

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(\xi)}{g'(\xi)} = \lim_{x\to a^+} \left(\frac{f'}{g'}\right)(\xi(x)) = A$$

Now the same reasoning can be used for b where we will use lim(x→b-) to replace all the $\lim_{x\to a^+}$, and ξ will be a function which maps to (x, b). □

Example 11.14

- $\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}$.
- $\lim_{x\to+\infty} \frac{x^2}{e^{3x}} = 0$.

§11.5 *Taylor Expansion*

Consider a function $f : [a, b] \to \mathbf{R}$. We first look at the mean value theorem from the viewpoint of approximations for $f(x)$ near a point $x = a$. We can regard the constant function

$$f_0(x) = f(a)$$

as the ***zero order approximation*** of $f(x)$. Then we ask if we can understand the remainder

$$R_1(x) \coloneqq f(x) - f(a), \quad x \in [a, b]$$

for this approximation. For this, if we assume $f \in C_0[a, b]$ and f' exists over (a, b), then the mean value theorem tells us that there exists some $a < \xi_x < x$ (here ξ_x vocabasises that ξ depends on x) so that we can write R_1 as

$$R_1(x) = f'(\xi_x)(x - a).$$

This is saying that the derivative of f can control the remainder $R_1(x)$ as an order 1 monomial.

The main expression is as follows:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots \tag{11.1}$$

So for example we have the following (we've used the ones for e^x and $\ln x$ for generating functions):

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \\ \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \\ \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \end{aligned}$$

There's a lot of things to say about these equations, for example the one for $\ln(1 + x)$ only works for $|x| < 1$

Also, if you want the RHS of the expression to be an infinite power series, $f(x)$ has to be smooth (infinitely differentiable)

Even then, the power series may never converge to $f(x)$ at any interval, no matter how small The most common example given here is $f(x) = e^{\frac{-1}{x^2}}$ (f(0)=0); the Taylor series for $f(x)$ is just 0

Now sometimes we don't actually that nice of a property for f, we're often given that fact that f is only finitely differentiable

Then we will have something along the lines of

$$f(x) \approx f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

where $f^{(n)}$ denotes the n-th differential.

There are two main forms of the statement regarding the error between the original function and the Taylor series estimate

The simpler form is what's known as the Peano form: Given that f is n times differentiable at a, then

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + o((x-a)^n)$$

To show this, we only need to show that we have the following limit:

$$\lim_{x\to a} \frac{f(x) - f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n}{(x-a)^n} = 0$$

The basic idea is to use the L'Hopital Rule n times. The numerator becomes $f^{(n)}(x) - f^{(n)}(a)$ which approaches 0, whereas the denominater is just $n!$, so the limit exists and is equal to 0.

However, we need to verify all the necessary conditions for L'Hopital : Here the main problem is that we don't know if we have the 0/0 indeterminate at each step, so we'll need to check this for the k-th step where k=1,...,n

Fortunately, the k-th derivative of the numerator is $f^{(k)}(x) - f^{(k)}(a) - (x-a)F_k(x)$ where F_k is just a bunch of random stuff, so the numerator approaches 0 as $x \to a$ The k-th derivative of the denominator is $n(n-1)\cdots(n-k+1)(x-a)^{n-k}$ so it also approaches 0, and we're done

The other form is actually a family of similar statements which gives more precise values for the error The Peano form has a fundamental obstacle when used in approximation, we don't have any control on the size of the final term other than its asymptotic behaviour : We'll be talking about the one given in the book, known as the Lagrange form: : Given that f is n times differentiable on (a,b) such that $f^{(n-1)}$ is continuous on $[a,b]$, then

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^(n-1) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n$$

Just like in L'Hopital, we intuitively think of (a,b) as just a very small interval at the right hand side of x=a : Here we are giving up on the second final term of Peano by combining it with the infinitesimal (small o) term to give an accurate description of the error

For the proof of this one we'll be using Cauchy's MVT

Fix any $x \in (a,b)$, then we construct the functions

$$F(t) = f(x) - \left(f(t) + \frac{f'(t)}{1!}(x-t) + \frac{f''(t)}{2!}(x-t)^2 + \cdots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} \right)$$

$$G(t) = (x-t)^n$$

We calculate $F'(t)$ as follows:

$$-[f'(t) + \frac{f''(t)}{1!} - f'(t) + \frac{f'''(t)}{2!} - \frac{f''(t)}{1!} + \cdots + \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} - \frac{f^{(n-1)}(t)}{(n-2)!}(x-t)^{n-2}] = -\frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1}$$

$G'(t) = -n(x-t)^{n-1}$, so we have

$$\frac{F'(t)}{G'(t)} = \frac{f^{(n)}(t)}{n!}$$

The main reason for why we come up with the strange-looking F and G is that we specifically swap out a for t so that $F(x) = G(x) = 0$, in hopes of getting rid of x:

We apply Cauchy's MVT to F and G on $[a,x]$, so that we obtain $\xi \in (a,x)$ satisfying

$$\frac{F'(\xi)}{G'(\xi)} = \frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F(a)}{G(a)}.$$

Thus the Lagrange form of the remainder is given by

$$F(a) = \frac{f^{(n)}(\xi)}{n!} G(a).$$

Theorem 5.19 is important, so do go through that proof as an exercise

12 Riemann–Stieltjes Integral

§12.1 *Definition of Riemann–Stieltjes Integral*

Assume $[a, b]$ is a closed interval in $\mathbf{R}$. By a ***partition*** P, we mean a finite set of points $x_0, x_1, \ldots, x_n$ where

$$a = x_0 \leq x_1 \leq \cdots \leq x_{n-1} \leq x_n = b.$$

Assume f is a bounded real-valued function over $[a, b]$ and α is an increasing function over $[a, b]$. Denote by

$$M_i = \sup_{[x_{i-1}, x_i]} f(x), \quad m_i = \inf_{[x_{i-1}, x_i]} f(x)$$

and by

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

Define the ***upper sum*** of f with respect to the partition P and α as

$$U(f, \alpha; P) = \sum_{i=1}^{n} M_i \Delta\alpha_i$$

and the ***lower sum*** of f with respect to the partition P and α as

$$L(f, \alpha; P) = \sum_{i=1}^{n} m_i \Delta\alpha_i.$$

Define the upper Riemann–Stieltjes integral as

$$\overline{\int_a^b} f(x)\, \mathrm{d}\alpha(x) \coloneqq \inf_P U(f, \alpha; P)$$

and the lower Riemann–Stieltjes integral as

$$\underline{\int_a^b} f(x)\, \mathrm{d}\alpha(x) \coloneqq \sup_P L(f, \alpha; P).$$

It is easy to see from definition that

$$\underline{\int_a^b} f(x)\, \mathrm{d}\alpha(x) \leq \overline{\int_a^b} f(x)\, \mathrm{d}\alpha(x)\,.$$

Definition 12.1. A function f is ***Riemann–Stieltjes integrable*** with respect to α over $[a, b]$, if

$$\underline{\int_a^b} f(x)\, \mathrm{d}\alpha(x) = \overline{\int_a^b} f(x)\, \mathrm{d}\alpha(x)\,.$$

Notation. We use $\displaystyle\int_a^b f(x)\, \mathrm{d}\alpha(x)$ to denote the common value, and call it the Riemann–Stieltjes of f with respect to α over $[a, b]$.

Notation. We use the notation $R_\alpha[a,b]$ to denote the set of Riemann–Stieljes integrable functions with respect to α over $[a,b]$.

In particular, when $\alpha(x) = x$, we call the corresponding Riemann–Stieljes integration the ***Riemann integration***, and use $R[a,b]$ to denote the set of Riemann integrable functions.

Definition 12.2. The partition P' is a ***refinement*** of P if $P' \supset P$. Given two partitions P_1 and P_2, we say that P' is their ***common refinement*** if $P' = P_1 \cup P_2$.

Intuitively, a refinement will give a better estimation than the original partition, so the upper and lower sums of a refinement should be more restrictive.

Proposition 12.3. If P' is a refinement of P, then

$$L(f,\alpha;P) \le L(f,\alpha;P')$$

and

$$U(f,\alpha;P') \le U(f,\alpha;P).$$

Proof. Suppose that

$$P : a \le x_0 \le x_1 \le ... \le x_n = b$$

and

$$P' : a \le y_0 \le y_1 \le ... \le y_m = b.$$

Then there exists a strictly increasing sequence of indices $j_0 = 0, j_1, \ldots, j_n = m$ such that $y_{j_k} = x_k$.

Now consider each closed interval $[x_{i-1}, x_i]$

Focusing on the upper sum, we have

$$\sup_{[x_{i-1},x_i]} f \ge \sup_{[y_{k-1},y_k]} f$$

for $k = j_{i-1}+1, \ldots, j_i$. This is because $[y_{k-1}, y_k]$ is contained in $[x_{i-1}, x_i]$

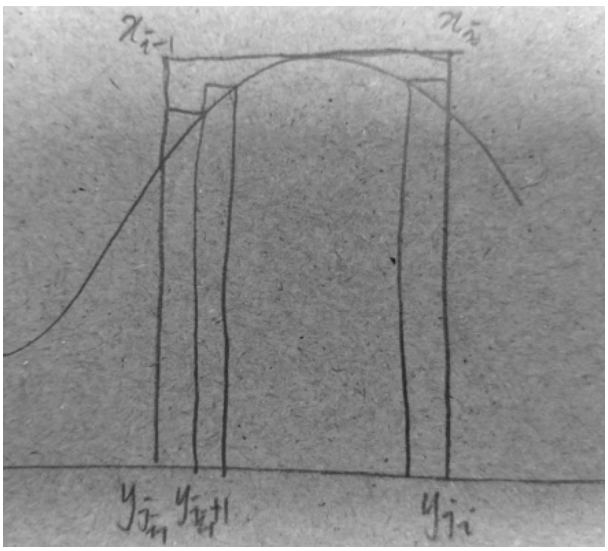

Figure 12.1: Partitions

Continuing from

$$\sup_{[x_{i-1},x_i]} f \ge \sup_{[y_{k-1},y_k]} f,$$

We then multiply by $\alpha(y_k) - \alpha(y_{k-1})$ on both sides and then take the sum from $k = j_{i-1}+1$ to $k = j_i$: The RHS corresponds to the (weighted) sum of the thin rectangles that you see in the above picture : The LHS is actually a telescoping sum, and the sum would be

$$(\sup_{[x_{i-1},x_i]} f) \cdot [\alpha(y_{j_i}) - \alpha(y_{j_{i-1}})] = (\sup_{[x_{i-1},x_i]} f) \cdot [\alpha(x_i) - \alpha(x_{i-1})]$$

Finally, we take the sum from $i = 1$ to $i = n$ of the above inequality LHS $\geq$ RHS (sorry I don't know of a better way to put it) We then obtain $U(P, f, \alpha) \geq U(P', f, \alpha)$

(On the LHS we're collecting all the rectangles for the upper sum wrt P, but on the RHS we're collecting up collections of upper rectangles to obtain the entire collective of upper rectangles for the upper sum wrt P') : Lower sum is similar : Now, a lemma used to prove 6.5 Given any two partitions P_1 and P_2, we have

$$L(P_1, f, \alpha) \leq U(P_2, f, \alpha)$$

So a lower sum will always be no larger than any other upper sum : So this includes the cases where we have the most refined of P_1's and P_2's, with no information regarding the partition points whatsoever To be honest, the result seems to be both intuitive and unclear at the same time

The key here is to use common refinements as a link for both sums The idea is stated in the proof of 6.5 and I don't think I need to elaborate further

What's nice here is that now we have two completely independent partitions P_1 and P_2, so by fixing one partition, say P_2, and taking the 'limit' over the other (here we take the supremum over all possible P_1) we then obtain an inequality between a Darboux integral and a Darboux sum (here it's the lower integral and an upper sum)

Since the Darboux integral is just a number, we can then safely take the 'limit' over the other partition to obtain the inequality in 6.5 □

Proposition 12.4.

$$\underline{\int_a^b} f \, \mathrm{d}\alpha = \overline{\int_a^b} f \, \mathrm{d}\alpha \, .$$

Proof. □

Now we move on to integrability conditions for f. The first one looks a lot like the $\epsilon - N$ or $\epsilon - \delta$ definition of limits:

Theorem 12.5. $f \in R_\alpha[a, b]$ if and only if for each $\epsilon > 0$, there exists some partition P such that

$$U(f, \alpha; P) - L(f, \alpha; P) < \epsilon.$$

Proof.

($\Longrightarrow$) Assume $f \in R_\alpha[a, b]$. By definition,

$$\inf_P U(f, \alpha; P) = \int_a^b f \, \mathrm{d}\alpha = \sup_P L(f, \alpha; P).$$

For every $\epsilon > 0$,

($\Longleftarrow$) □

Example 12.6 (Dirichlet function)

The Dirichlet function is given by

$$f(x) = \begin{cases} 1 & x \in \mathbf{Q} \\ 0 & x \notin \mathbf{Q} \end{cases}$$

We try to calculate the two on the interval $[0, 1]$.

The Dirichlet function is pathological because for each subinterval $[x_{i-1}, x_i]$, the supremum is always 1 and the infimum is always 0.

So no matter what partition we use, $U(f, P)$ is always 1 whereas $L(f, P)$ is always 0. This means that $U(f) = 1$ and $L(f) = 0$, so there are two different values for "the integral of f".

This is like the case where we try to find the limit of the Dirichlet function where x is approaching any given real number r, there exists two sequences approaching r whose image approaches two different values.

Now, a very important and fun case about the more general RS-integral, which we'll discuss next week (do try the exercise yourself first)

Exercise 36

The Heaviside step function H is a real-valued function defined by the following:

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

For the purpose of this question we assume the convention $\infty \cdot 0 = 0$.

(a) Let f be a real-valued function over $\mathbf{R}$. Show that $f \in \mathbf{R}_H[a, b]$ if and only if f is continuous at 0, and find the RS-integral $\int_{-\infty}^{\infty} f \, \mathrm{d}H$.

(b) Suppose that the definition for H is changed for $x = 0$, say $H(0) = \frac{1}{2}$. Show that the above result still holds.

(c) Examine the RS-integral of f over $\mathbf{R} \setminus \{0\}$ wrt H, where f is a real-valued function over $\mathbf{R} \setminus \{0\}$ such that $\lim_{x \to 0} f(x) = \infty$ or $-\infty$.

(You may read up on more information regarding the Heaviside function, and the (in)famous Dirac delta function)

Now we've been talking a lot about upper and lower sums because they're arguably the simplest way to define integrals, in the sense that there's not a whole lot of things that we could go wrong here By considering only upper and lower bound, we're essentially picking the most conservative route possible

It would be nice if we could just pick like one random point within each interval and consequently calculate the Riemann(-Stieltjes) sums

This method, of course, fails to be well defined for pathological functions like the Dirichlet function On the other hand, by using upper and lower sums, we could give a persuasive explanation as to why the Dirichlet function is not Riemann integrable

However, instead of throwing this idea away, there's actually a way for us to make this into a strict definition

When we were talking about the sequential definition for limits of functions, we noted that there are certain scenarios where the limit cannot exist because there may be two distinct sequences may give different limit Based on this observation, we then gave a reasonable condition as follows: "$\lim_{x \to a} f(x)$ exists and is equal to L if and only if for all sequences x_n converging but not containing a, $f(x_n)$ converges to L"

Well here, it's actually the same kind of scenario Given any partition P, we consider the Riemann sum $\sum f(\xi_i)\Delta x_i$ where ξ_i is any point where $x_{i-1} \leq \xi_i \leq x_i$

For the Dirichlet function over $[0, 1]$, given any partition P (here we may assume that the partition points are distinct), we will always be able to specifically pick $\xi_i, \eta_i \in [x_{i-1}, x_i]$ such that ξ_i is rational but η_i is irrational

Then $\sum f(\xi_i)\Delta x_i = 1$ but $\sum f(\eta_i)\Delta x_i = 0$

Now be very mindful that this alone cannot be evidence that f is non-integrable The key is that this somehow occured for all partitions P, no matter how refined they are; for every single partition P, there exists two sets of 'representing points' ξ_i, η_i such that the two Riemann sums are constantly far apart (1 and 0 in this case)

Let $\epsilon_0 = 1$, then this ultimately translates to the following: The Dirichlet function cannot be Riemann integrable because There exists some $\epsilon_0 > 0$, such that for any given partition P, there exists two sets of representing points ξ_i, η_i such that their corresponding Riemann sums satisfy that

$$|\sum f(\xi_i)\Delta x_i - \sum f(\eta_i)\Delta x_i| \geq \epsilon_0.$$

Now if we always pick the representatives such that $\xi_i > \eta_i$ then we can neglect the absolute value

So now, let's take the converse A function f is said to be RS-integrable if For every $\epsilon > 0$, There exists a partition P, such that For any two sets of representing points ξ_i, η_i, Their corresponding Riemann sums satisfy that

$$\sum [f(\xi_i) - f(\eta_i)]\Delta x_i < \epsilon$$

(The last one should be $\Delta\alpha_i$ for RS-integrals, not Δx_i)

Unfortunately this is still not quite the correct definition according to Apostol, but we're pretty close The problem with this definition is that it is too weak if we're considering general α of bounded variation; if we were only talking about monotonically increasing α then this will actually be an equivalent definition

The official definition for the RS-integral wrt α of bounded variation is as follows:

Definition 12.7. For every $\epsilon > 0$, there exists a partition P, such that [For any refinement P' of P, and] For any two sets of representing points ξ_i, η_i [of P'], their corresponding Riemann sums satisfy that

$$\sum [f(\xi_i) - f(\eta_i)]\Delta x_i < \epsilon.$$

Now this definition is what mathematicians would refer to as a 'Cauchy' definition, since it defines a notion by comparing a pair of arbitrary values that are similar to one another, and if they agree in some sense then we say that that something satisfies some property.

The integral is then obtained as follows: If f were to satisfy the above Cauchy definition, then we may pick an arbitrary sequence of refinements

$$P_1 \subset P_2 \subset P_3 \subset ...;$$

and for each partition we pick a set of representatives to obtain a sequence RS-sum $I_1, I_2, I_3, ...$: This sequence will be a Cauchy sequence of real numbers, and so will converge to a specific value I which we consider to be RS-integral of f : Now the reason why Apostol needed to strengthen the definition is that, otherwise this value I may not be unique : So if you look at the statement you see in 6.7(b)(c), then they correspond to the Cauchy definition and the 'value-based' definition respectively For monotonically increasing α, it is much easier to discuss them using upper and lower sums So your exercise today will be to read the statements and proofs in Theorem 6.7

Theorem 12.8. $f \in R_\alpha[a, b]$, $m \leq f \leq M$, and ϕ is uniformly continuous on $[m, M]$, then

$$\phi \circ f \in R_\alpha[a, b].$$

Proof. Choose $\epsilon > 0$. Since ϕ is uniformly continuous on $[m, M]$, there exists $\delta > 0$ such that $\delta < \epsilon$ and $|\phi(s) - \phi(t)|$ □

§12.2 *Properties of the Integral*

Theorem 12.9.

(1) If $f_1, f_2 \in R_\alpha[a, b]$, then

$$f_1 + f_2 \in R_\alpha[a, b];$$

$cf \in R_\alpha[a, b]$ for every $c \in \mathbf{R}$, and

$$\int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha\,,$$

$$\int_a^b (cf)\, d\alpha = c \int_a^b f\, d\alpha\,.$$

(2) If $f_1, f_2 \in R_\alpha[a, b]$ and $f_1 \leq f_2$, then

$$\int_a^b f_1\, d\alpha \leq \int_a^b f_2\, d\alpha\,.$$

(3) If $f \in R_\alpha[a,b]$ and $c \in [a,b]$, then $f \in R_\alpha[a,c]$ and $f \in R_\alpha[c,b]$, and

$$\int_a^b f \,\mathrm{d}\alpha = \int_a^c \mathrm{d}\alpha + \int_c^b \mathrm{d}\alpha\,.$$

(4) If $f \in R_\alpha[a,b]$ and $|f| \le M$, then

$$\left|\int_a^b f \,\mathrm{d}\alpha\right| \le M\left[\alpha(b) - \alpha(a)\right].$$

(5) If $f \in R_{\alpha_1}[a,b]$ and $f \in R_{\alpha_2}[a,b]$, then $f \in R_{\alpha_1+\alpha_2}[a,b]$ and

$$\int_a^b f \,\mathrm{d}(\alpha_1+\alpha_2) = \int_a^b f \,\mathrm{d}\alpha_1 + \int_a^b f \,\mathrm{d}\alpha_2\,;$$

if $f \in R_\alpha[a,b]$ and c is a positive constant, then $f \in R_{c\alpha}[a,b]$ and

$$\int_a^b f \,\mathrm{d}(c\alpha) = c\int_a^b f \,\mathrm{d}\alpha\,.$$

(6) If $f \in R_\alpha[a,b]$ and $g \in R_\alpha[a,b]$, then $fg \in R_\alpha[a,b]$.

Proof.

(1) If $f = f_1 + f_2$ and P is any partition of $[a,b]$, we have

$$\begin{aligned} L(f_1,\alpha;P) + L(f_2,\alpha;P) &\le L(f,\alpha;P) \\ &\le U(f,\alpha;P) \\ &\le U(f_1,\alpha;P) + U(f_2,\alpha;P). \end{aligned}$$

If $f_1 \in R_\alpha[a,b]$ and $f_2 \in R_\alpha[a,b]$, let $\epsilon > 0$ be given. There are partitions P_1 and P_2 such that

(2)

(3)

(4)

(5)

(6)

□

Theorem 12.10 (Triangle inequality)**.** $f \in R_\alpha[a,b]$, then $|f| \in R_\alpha[a,b]$,

$$\left|\int_a^b f \,\mathrm{d}\alpha\right| \le \int_a^b |f| \,\mathrm{d}\alpha\,.$$

Proof. □

6.14 6.15 Heaviside step function

6.16 corollary for intinite sum, need $\sum c_n$ to converge (23) comparison test

6.17 integration by substitution

Theorem 12.11. α increasing, $\alpha' \in R[a,b]$, f bounded on $[a,b]$, then

$$f \in R_\alpha[a,b] \iff f\alpha' \in R[a,b].$$

6.19 change of variables

§12.3 *Fundamental Theorem of Calculus*

6.20 6.21

Theorem 12.12.

6.22 integration by parts

13 Sequence and Series of Functions

§13.1 *Uniform Convergence*

Definition 13.1 (Pointwise convergence)**.** Suppose $\{f_n\}$, $n = 1, 2, 3, \ldots$ is a sequence of functions defined on a set E, and suppose that the sequence of numbers $\{f_n(x)\}$ converges for every $x \in E$. We can then define a function f by

$$f(x) = \lim_{n\to\infty} f_n(x).$$

We say that $\{f_n\}$ ***converges pointwise*** to f on E, denoted by $f_n \to f$, and f is the ***limit***, or the ***limit function***, of $\{f_n\}$.

$$\forall \epsilon > 0, \forall x \in E, \exists N \in \mathbf{N} \text{ s.t. } \forall n > N, |f_n(x) - f(x)| < \epsilon.$$

Similarly, if $\sum f_n(x)$ converges for every $x \in E$, and if we define

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

the function f is called the ***sum of the series*** $\sum f_n$.

Most properties are not preserved by pointwise continuity; that is, f does not inherit most properties of f_n.

Example 13.2 (f_n continuous, f discontinuous)
Let $f_n(x) = x^n$ for $x \in [0, 1]$. Then

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{if } x \in (0, 1] \\ 1 & \text{if } x = 1 \end{cases}$$

and so the limit function $f(x)$ is discontinuous.

Example 13.3 (f_n integrable, f not integrable)
Recall that the Dirichlet function

$$D(x) = \begin{cases} 1 & \text{if } x \in \mathbf{Q} \\ 0 & \text{if } x \in \mathbf{R} \setminus \mathbf{Q} \end{cases}$$

is not integrable.

Proof. Consider the interval $[0, 1]$. We partition $P : 0 = x_0 < x_1 < \cdots < x_n = 1$. The sum is given by $\sum_{i=1}^{n} D(t_i)\Delta x_i$. Then

$$M_i = \max_{t\in[x_{i-1},x_i]} D(t) = 1 \implies U(D; P) = 1 \quad \forall P$$

and

$$m_i = \min_{t\in[x_{i-1},x_i]} D(t) = 0 \implies L(D;P) = 0 \quad \forall P.$$

Hence

$$\bar{\int_0^1} D(x)\,\mathrm{d}x = 1, \quad \underline{\int_0^1} D(x)\,\mathrm{d}x = 0$$

so $\bar{\int_0^1} D(x)\,\mathrm{d}x \neq \underline{\int_0^1} D(x)\,\mathrm{d}x$, and thus $D(x)$ is not integrable. □

We define a sequence of functions as follows:

$$D_n(x) = \begin{cases} 1 & \text{if } x = \frac{p}{q}, p \in \mathbf{Z}, q \in \mathbf{Z}\setminus\{0\}, |q| \leq n \\ 0 & \text{if otherwise} \end{cases}$$

Definition 13.4 (Uniform convergence)**.** We say that $\{f_n\}$ ***uniformly converges*** to f over E, if

$$\forall \epsilon > 0, \exists N \in \mathbf{N} \text{ s.t. } \forall x \in E, \forall n > N, |f_n(x) - f(x)| < \epsilon.$$

We denote this by $f_n \rightrightarrows f$.

Uniform convergence is stronger than pointwise convergence, since N is uniform (or "fixed") for all $x \in E$; for pointwise convergence, the choice of N is determined by x.

Definition 13.5. If X is a metric space, we denote the set of all complex-valued, continuous, bounded functions with domain X by $C(X)$.

If $f \in C(X)$, we define

$$\|f\| \coloneqq \sup_{x\in X} |f(x)|,$$

known as the ***supremum norm*** of f.

Lemma 13.6. $\|f\|$ gives a norm on $C(X)$.

Proof. Check that $\|f\|$ satisfies the conditions for a norm:

(1)

□

Proposition 13.7. $\big(C(X), \|\cdot\|\big)$ is a metric space.

Lemma 13.8 (Cauchy criterion)**.** $\{f_n\} \rightrightarrows f$ if and only if

$$\forall \epsilon > 0, \exists N \in \mathbf{N} \text{ s.t. } \forall x \in E, \forall n, m > N, |f_n(x) - f_m(x)| < \epsilon.$$

Proof.

($\implies$) Suppose $f_n \rightrightarrows f$, then

$$\forall \epsilon > 0, \exists N \in \mathbf{N} \text{ s.t. } \forall x \in E, \forall n > N, |f_n(x) - f(x)| < \frac{\epsilon}{2}.$$

Then for all $n, m > N$,

$$\begin{aligned} |f_n(x) - f_m(x)| &= \Big|\big(f_n(x) - f(x)\big) - \big(f_m(x) - f(x)\big)\Big| \\ &\leq |f_n(x) - f(x)| + |f_m(x) - f(x)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

by triangle inequality.

($\impliedby$)

$$\forall \epsilon > 0, \exists N \in \mathbf{N} \text{ s.t. } \forall x \in E, \forall n, m > N, |f_n(x) - f_m(x)| < \epsilon.$$

□

The uniform convergence of series is defined similarly:

Lemma 13.9 (Cauchy criterion)**.**

Theorem 13.10 (Weierstrass M-test)**.** $\sum_{n=1}^{\infty} f_n(x)$ uniformly converges if

$$\exists \{M_n\} \in \mathbf{R}^+ \text{ s.t. } |f_n(x)| < M_n, \sum_{n=1}^{\infty} M_n \text{ convergent}$$

where $\sum_{n=1}^{\infty} M_n$ is convergent if

$$\forall \epsilon > 0, \exists N \in \mathbf{N} \text{ s.t. } \forall n, m > N, \sum_{k=m+1}^{n} M_k < \epsilon.$$

§13.2 *Uniform Convergence and Continuity*

We now consider properties preserved by uniform convergence.

§13.3 *Uniform Convergence and Integration*

Theorem 13.11. Assume $\{f_n\}$ is a sequence of functions defined over $[a, b]$ and each $f_n \in R_\alpha[a, b]$. If $f_n \to f$, then $f \in R_\alpha[a, b]$, and

$$\lim_{n\to\infty} \int_a^b f_n \, \mathrm{d}\alpha = \int_a^b f \, \mathrm{d}\alpha \,.$$

Proof. Define □

Corollary 13.12. Assume $a_n \in R_\alpha[a, b]$ and

$$f(x) \coloneqq \sum_{n=0}^{\infty} a_n(x)$$

converges uniformly. Then it follows

$$\int_a^b f \, \mathrm{d}\alpha = \sum_{n=0}^{\infty} a_n \, \mathrm{d}\alpha \,.$$

Proof. Consider the sequence of partial sums

$$f_n(x) \coloneqq \sum_{k=0}^{n} a_k(x), \quad n = 0, 1, \ldots$$

It follows $f_n \in R_\alpha[a, b]$ and $f_n \rightrightarrows f$. Apply above theorem to $\{f_n\}$ and the conclusion follows. □

§13.4 *Uniform Convergence and Differentiation*

Theorem 13.13. $\{f_n\}$ differentiable on $[a, b]$, $\exists x_0 \in [a, b]$ s.t. $f_n(x_0) \to y_0 = f(x_0)$ and $f_n' \rightrightarrows f'$. Then $f_n \rightrightarrows f$ on $[a, b]$, and f is differentiable, $f'(x) = \lim_{n\to\infty} f_n'(x)$ for any $x \in [a, b]$.

Proof. $f_n(x_0) \to y_0$ thus □

§13.5 *Stone–Weierstrass Approximation Theorem*

Theorem 13.14 (Weierstrass approximation theorem)**.** If f is a continuous complex function on $[a, b]$, there exists a sequence of polynomials P_n such that $P_n \rightrightarrows f$ on $[a, b]$.

If f is real, then P_n may be taken real.

14 Some Special Functions

§14.1 *Power Series*

Definition 14.1. ***Analytic functions*** are functions that can be represented by ***power series***, i.e. functions of the form

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

or, more generally,

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n.$$

The ***radius of convergence*** is the maximum R such that $f(x)$ converges in $(-R, R)$.

Theorem 14.2. Suppose the series

$$\sum_{n=0}^{\infty} c_n x^n$$

converges for $x \in (-R, R)$. Then

(1) $\sum_{n=0}^{\infty} c_n x^n$ converges uniformly on the closed interval $[-R+\epsilon, R-\epsilon]$ for all $\epsilon > 0$;

(2) $f(x)$ is continuous and differentiable on $(-R, R)$, and

$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1}.$$

Proof.

(1) Let $\epsilon > 0$ be given. For $|x| \le R - \epsilon$, we have

(2)

□

Corollary 14.3. f has derivatives of all orders in $(-R, R)$, which are given by

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) c_n x^{n-k}.$$

In particular,

$$f^{(k)}(0) = k! c_k, \quad k = 0, 1, 2, \dots$$

(Here $f^{(0)}$ means f, and $f^{(k)}$ is the k-th derivative of f, for $k = 1, 2, 3, \dots$)

Proof. Apply theorem successively to $f, f', f'', \ldots$. Put $x = 0$. □

Proposition 14.4. Suppose $\sum c_n$ converges. Put

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

for $x \in (-R, R)$

§14.2 *Exponential and Logarithmic Functions*

Definition 14.5 (Exponential function)**.** We define the exponential function as

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}. \tag{14.1}$$

Lemma 14.6. $\exp(z)$ converges for every $z \in \mathbf{C}$.

Proof. Ratio test. □

§14.3 *Trigonometric Functions*

§14.4 *Algebraic Completeness of the Complex Field*

We now prove that the complex field is ***algebraically complete***; that is, every non-constant polynomial with complex coefficients has a complex root.

Theorem 14.7 (Fundamental Theorem of Algebra)**.** Suppose $a_0, \ldots, a_n$ are complex numbers, $n \geq 1$, $a_n \neq 0$,

$$P(z) = \sum_{k=0}^{n} a_k z^k.$$

Then $P(z) = 0$ for some complex number z.

Proof. □

§14.5 *Fourier Series*

Definition 14.8. A ***trigonometric polynomial*** is a finite sum of the form

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

for $x \in \mathbf{R}$, where $a_0, \ldots, a_N, b_1, \ldots, b_N \in \mathbf{C}$.

On account of the identities (?), we can write the above in the form

$$f(x) = \sum_{n=-N}^{N} c_n e^{inx}.$$

It is clear that every trigonometric polynomial is periodic, with period 2π.

§14.6 *Gamma Function*

Definition 14.9 (Gamma function)**.** For $0 < x < \infty$,

$$\Gamma(x) \coloneqq \int_0^\infty t^{x-1} e^{-t} \, \mathrm{d}t \, .$$

The integral converges for these x. (When $x < 1$, both 0 and ∞ have to be looked at.)

Lemma 14.10.

(1) The functional equation

$$\Gamma(x+1) = x\Gamma(x)$$

holds for $0 < x < \infty$.

(2) $\Gamma(n+1) = n!$ for $n = 1, 2, 3, \ldots$

(3) $\log \Gamma$ is convex on $(0, \infty)$.

Proof.

(1) Integrate by parts.

(2) Since $\Gamma(1) = 1$, (1) implies (2) by induction.

(3)

□

In fact, these three properties characterise Γ completely.

Lemma 14.11 (Characteristic properties of Γ)**.** If f is a positive function on $(0, \infty)$ such that

(1) $f(x+1) = xf(x)$,

(2) $f(1) = 1$,

(3) $\log f$ is convex,

then $f(x) = \Gamma(x)$.

Proof. □

Definition 14.12 (Beta function)**.** For $x > 0$ and $y > 0$, the beta function is defined as

$$B(x, y) \coloneqq \int_0^1 t^{x-1} (1-t)^{y-1} \, \mathrm{d}t \, .$$

Lemma 14.13.

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Proof. Let $f(x) = \dfrac{\Gamma(x+y)}{\Gamma(y)} B(x, y)$. We want to prove that $f(x) = \Gamma(x)$, using Lemma 14.11.

(1)

$$B(x+1, y) = \int_0^1 t^x (1-t)^{y-1} \, \mathrm{d}t \, .$$

Integrating by parts gives

$$\begin{aligned} B(x+1,y) &= \underbrace{\left[t^x \cdot \frac{(1-t)^y}{y}(-1)\right]_0^1}_{0} + \int_0^1 x t^{x-1}\frac{(1-t)^y}{y}\,\mathrm{d}t \\ &= \frac{x}{y}\int_0^1 t^{x-1}(1-t)^{y-1}(1-t)\,\mathrm{d}t \\ &= \frac{x}{y}\left(\int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t - \int_0^1 t^x(1-t)^{y-1}\,\mathrm{d}t\right) \\ &= \frac{x}{y}\big(B(x,y) - B(x+1,y)\big) \end{aligned}$$

which gives $B(x+1,y) = \dfrac{x}{x+y}B(x,y)$. Thus

$$\begin{aligned} f(x+1) &= \frac{\Gamma(x+1+y)}{\Gamma(y)}B(x+1,y) \\ &= \frac{(x+y)B(x+y)}{\Gamma(y)}\cdot\frac{x}{x+y}B(x,y) \\ &= xf(x). \end{aligned}$$

(2)

$$B(1,y) = \int_0^1 (1-t)^{y-1}\,\mathrm{d}t = \left[-\frac{(1-t)^y}{y}\right]_0^1 = \frac{1}{y}$$

and thus

$$f(1) = \frac{\Gamma(1+y)}{\Gamma(y)}B(1,y) = \frac{y\Gamma(y)}{\Gamma(y)}\frac{1}{y} = 1.$$

(3) We now show that $\log B(x,y)$ is convex, so that

$$\log f(x) = \underbrace{\log\Gamma(x+y)}_{\text{convex}} + \log B(x,y) - \underbrace{\log\Gamma(y)}_{\text{constant}}$$

is convex with respect to x.

$$B(x_1,y)^{\frac{1}{p}}B(x_2,y)^{\frac{1}{q}} = \left(\int_0^1 t^{x_1-1}(1-t)^{y-1}\,\mathrm{d}t\right)^{\frac{1}{p}}\left(\int_0^1 t^{x_2-1}(1-t)^{y-1}\,\mathrm{d}t\right)^{\frac{1}{q}}$$

By Hölder's inequality,

$$\begin{aligned} B(x_1,y)^{\frac{1}{p}}B(x_2,y)^{\frac{1}{q}} &= \int_0^1 \left[t^{x_1-1}(1-t)^{y-1}\right]^{\frac{1}{p}}\left[t^{x_2-1}(1-t)^{y-1}\right]^{\frac{1}{q}}\,\mathrm{d}t \\ &= \int_0^1 t^{\frac{x_1}{p}+\frac{x_2}{q}-1}(1-t)^{y-1}\,\mathrm{d}t \\ &= B\left(\frac{x_1}{p}+\frac{x_2}{q},y\right). \end{aligned}$$

Taking log on both sides gives

$$\log B(x,y)^{\frac{1}{p}}B(x_2,y)^{\frac{1}{q}} \geq \log B\left(\frac{x_1}{p}+\frac{x_2}{q},y\right)$$

or

$$\frac{1}{p}\log B(x,y) + \frac{1}{q}\log B(x_2,y) \geq \log B\left(\frac{x_1}{p}+\frac{x_2}{q},y\right).$$

Hence $\log B(x,y)$ is convex, so $\log f(x)$ is convex.

Therefore $f(x) = \Gamma(x)$ which implies $B(x,y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$. □

An alternative form of Γ is as follows:

$$\Gamma(x) = 2\int_0^{+\infty} t^{2x-1}e^{-t^2}\,\mathrm{d}t\,.$$

Using this form of Γ, we present an alternative proof.

Proof.

$$\begin{aligned}\Gamma(x)\Gamma(y) &= \left(2\int_0^{+\infty} t^{2x-1}e^{-t^2}\,\mathrm{d}t\right)\left(2\int_0^{+\infty} s^{2y-1}e^{-s^2}\,\mathrm{d}s\right)\\ &= 4\iint_{[0,+\infty)\times[0,+\infty)} t^{2x-1}s^{2y-1}e^{-\left(t^2+s^2\right)}\,\mathrm{d}t\,\mathrm{d}s\end{aligned}$$

Using polar coordinates transformation, let $t = r\cos\theta$, $s = r\sin\theta$. Then $\mathrm{d}t\,\mathrm{d}s = r\,\mathrm{d}r\,\mathrm{d}\theta$. Thus

$$\begin{aligned}\Gamma(x)\Gamma(y) &= 4\int_0^{\frac{\pi}{2}}\left[\int_0^{+\infty} r^{2x-1}\cos^{2x-1}\theta\cdot r^{2y-1}\sin^{2y-1}\theta\cdot e^{-r^2}\cdot r\,\mathrm{d}r\right]\mathrm{d}\theta\\ &= \underbrace{2\int_0^{\frac{\pi}{2}}\cos^{2x-1}\theta\sin^{2y-1}\theta\,\mathrm{d}\theta}_{B(x,y)}\cdot\underbrace{2\int_0^{+\infty} r^{2(x+y)-1}e^{-r^2}\,\mathrm{d}r}_{\Gamma(x+y)}\end{aligned}$$

since

$$\begin{aligned}B(x,y) &= \int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t \quad t=\cos^2\theta\\ &= \int_{\frac{\pi}{2}}^0 \cos^{2(x-1)}\theta\sin^{2(y-1)}\theta\cdot 2\cos\theta(-\sin\theta)\,\mathrm{d}\theta\\ &= 2\int_0^{\frac{\pi}{2}}\cos^{2x-1}\theta\sin^{2y-1}\theta\,\mathrm{d}\theta\,.\end{aligned}$$

Hence $B(x,y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$. □

More on polar coordinates:

$$I = \int_{-\infty}^{+\infty} e^{-x^2}\,\mathrm{d}x \tag{14.2}$$

Proof.

$$\begin{aligned}I^2 &= \int_{-\infty}^{+\infty} e^{-x^2}\,\mathrm{d}x\int_{-\infty}^{+\infty} e^{-y^2}\,\mathrm{d}y\\ &= \iint_{\mathbf{R}^2} e^{-\left(x^2+y^2\right)}\,\mathrm{d}x\,\mathrm{d}y \quad x=r\cos\theta, y=r\sin\theta\\ &= \int_0^{2\pi}\underbrace{\int_0^{+\infty} e^{-r^2}r\,\mathrm{d}r}_{\text{constant w.r.t. }\theta}\mathrm{d}\theta \quad s=r^2, \mathrm{d}s = 2r\,\mathrm{d}r\\ &= 2\pi\int_0^{+\infty} e^{-s}\cdot\frac{1}{2}\,\mathrm{d}s\\ &= 2\pi\left[\frac{1}{2}e^{-s}(-1)\right]_0^{\infty} = \pi\end{aligned}$$

and thus

$$I = \int_{-\infty}^{+\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

□

From this, we have

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^\infty e^{-t^2}\,\mathrm{d}t = \sqrt{\pi}.$$

Lemma 14.14.

$$\Gamma(x) = \frac{2^{x-1}}{\sqrt{\pi}}\Gamma\left(\frac{x}{2}\right)\Gamma\left(\frac{x+1}{2}\right).$$

Proof. Let $f(x) = \frac{2^{x-1}}{\sqrt{\pi}}\Gamma\left(\frac{x}{2}\right)\Gamma\left(\frac{x+1}{2}\right)$. We want to prove that $f(x) = \Gamma(x)$.

(1)

$$\begin{aligned} f(x+1) &= \frac{2^x}{\sqrt{\pi}}\Gamma\left(\frac{x+1}{2}\right)\Gamma\left(\frac{x}{2}+1\right) \\ &= \frac{2^x}{\sqrt{\pi}}\Gamma\left(\frac{x+1}{2}\right)\frac{x}{2}\Gamma\left(\frac{x}{2}\right) \\ &= xf(x) \end{aligned}$$

(2) $f(1) = \frac{1}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}\right)\Gamma(1) = 1$ since $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

(3)

$$\log f(x) = \underbrace{(x-1)\log 2}_{\text{linear}} + \underbrace{\log\Gamma\left(\frac{x}{2}\right)}_{\text{convex}} + \underbrace{\log\Gamma\left(\frac{x+1}{2}\right)}_{\text{convex}} - \underbrace{\log\sqrt{\pi}}_{\text{constant}}$$

and hence $\log f(x)$ is convex.

Therefore $f(x) = \Gamma(x)$. □

Theorem 14.15 (Stirling's formula)**.** This provides a simple approximate expression for $\Gamma(x+1)$ when x is large (hence for $n!$ when n is large). The formula is

$$\lim_{x\to\infty}\frac{\Gamma(x+1)}{(x/e)^x\sqrt{2\pi x}} = 1. \tag{14.3}$$

Proof. □

Lemma 14.16.

$$B(p, 1-p) = \Gamma(p)\Gamma(1-p) = \frac{\pi}{\sin p\pi}.$$

Proof. □

Part V

Complex Analysis

15 Complex Plane

The aim of this part of the course is to study functions $f : \mathbf{C} \to \mathbf{C}$, asking what it means for them to be differentiable, how to integrate them, and looking at the applications of all this. Before we begin, we discuss some basic properties of the complex numbers $\mathbf{C}$.

§15.1 $\mathbf{C}$ *as a metric space*

We can identify $\mathbf{C}$ with the plane $\mathbf{R}^2$ by taking real and imaginary parts. Thus we have mutually inverse bijections

$$z \mapsto (\operatorname{Re} z, \operatorname{Im} z)$$

from $\mathbf{C}$ to $\mathbf{R}^2$, and

$$(x, y) \mapsto x + iy$$

from $\mathbf{R}^2$ to $\mathbf{C}$. As we have seen, $\mathbf{R}^2$ is a metric space with the metric induced from the Euclidean norm

$$\|(x, y)\|_2 = \sqrt{x^2 + y^2}.$$

This gives a metric on $\mathbf{C}$ by the identification $\mathbf{C} \cong \mathbf{R}^2$ described above.

If $z = \operatorname{Re} z + i \operatorname{Im} z$ is a complex number we write $|z|$ (called the ***modulus***) for this Euclidean norm; that is,

$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

The distance between the two points $z, w \in \mathbf{C}$ is then $|z - w|$.

Let us write down some basic properties of the modulus $|z|$. Recall that $e^{i\theta} = \cos\theta + i \sin\theta$ when $\theta \in \mathbf{R}$. For now, we will take this as the definition of $e^{i\theta}$. Later on we will define the complex exponential function e^z and link the two concepts.

Lemma 15.1. Let $z, w \in \mathbf{C}$. Then

(1) $|z|^2 = z\bar{z}$, where $\bar{z}$ is the complex conjugate of z;

(2) If $z = re^{i\theta}$, where $r \in [0, \infty)$ and $\theta] in \mathbf{R}$, then $|z| = r$;

(3) $|zw| = |z||w|$.

Proof.

(1) If $z = a + ib$ then $z\bar{z} = (a + ib)(a - ib) = a^2 + b^2$.

(2) We have $z = r\cos\theta + ir\sin\theta$ and so

$$|z| = \sqrt{r^2\cos^2\theta + r^2\sin^2\theta} = r.$$

(3) One can calculate directly, writing $z = a + ib$ and $w = c + id$. Alternatively, write $z = re^{i\theta}$, $w = r'e^{i\alpha}$, and then observe that $zw = rr'e^{i(\theta+\alpha)}$ and use (2).

□

§15.2 *Topological properties of* C

One can make sense of the notion of open set, closure, interior and so on by identifying $\mathbf{C}$ with $\mathbf{R}^2$.

Definition 15.2. $U \subset \mathbf{C}$ being ***open*** means that if $z \in U$ then some ball $B_r(z)$, $r > 0$, also lies in U, where

$$B_r(z) \coloneqq \{w \in \mathbf{C} \mid |z - w| < r\}.$$

In complex analysis it is often convenient to work with connected open sets, and these are called domains.

Definition 15.3. A connected open subset $D \subseteq \mathbf{C}$ of the complex plane will be called a domain.

We have seen that for open subsets of normed spaces (such as $\mathbf{R}^2$ with the Euclidean metric), the notions of connectedness and path-connectedness are the same thing. Therefore domains are always path-connected.

§15.3 *Geometry of* C

Let us take a closer look at the geometry of the complex plane in terms of the distance $|z - w|$. When we talk about lines and circles in $\mathbf{C}$, we mean sets that are lines and circles in $\mathbf{R}^2$ (under the identification of $\mathbf{R}^2$ with $\mathbf{C}$).

Lemma 15.4 (Lines)**.** Let $a, b \in \mathbf{C}$ be distinct complex numbers. Then the set $\{z \in \mathbf{C} \mid |z - a| = |z - b|\}$ is a line. Conversely, every line can be written in this form.

Proof. Given a and b, the set of z such that $|z - a| = |z - b|$ is the set of points equidistant from a and b, which is the perpendicular bisector of the line segment $\overline{ab}$. Conversely, every line is the perpendicular bisector of some line segment. □

Remark. Sometimes, the set of all complex numbers satisfying some given equation is called a ***locus***. Thus the locus of complex numbers satisfying $|z - a| = |z - b|$ is a line. The representation of lines in the above form is very much non-unique: for example, the x-axis (the set of z with zero imaginary part) can be described as $\{z \mid |z - a| = |z - \bar{a}\}$ for any complex number a.

Now we turn to circles. Evidently, the set $\{z \in \mathbf{C} \mid |z - c| = r\}$, where $c \in \mathbf{C}$ and $r \in (0, \infty)$, is a circle centred on c and with radius r. Conversely, every circle can be written in this form. Less obvious is the following.

Lemma 15.5 (Circles)**.** Let $a, b \in \mathbf{C}$ be distinct complex numbers, and let $\lambda \in (0, \infty)$, $\lambda \neq 1$. Then the locus of complex numbers satisfying $|z - a| = \lambda |z - b|$ is a circle. Conversely, every circle can be written in this form.

Proof. Without loss of generality, $b = 0$ (a translate of a circle is a circle). Now observe the identity

$$|tz + a|^2 = t(t + 1)|z|^2 - t|z - a|^2 + (t + 1)|a|^2,$$

valid for all $a, z \in \mathbf{C}$ and all $t \in \mathbf{R}$. This can be checked by a slightly tedious calculation. Applying it with $t = \lambda^2 - 1$ gives

$$|(\lambda^2 - 1)z + a| = \lambda |a|,$$

which is clearly the equation of a circle. Taking $a = -c(\lambda^2 - 1)$ and $\lambda = \frac{r}{|c|}$, this gives $|z - c| = r$, and so every circle can be written in this form. □

Remark. This lemma is an interesting and non-obvious fact in classical Eu- clidean geometry. Phrased in that language, if A, B are points in the plane, and if $\lambda \in (0, \infty)$, $\lambda \neq 1$, then the set of all points P such that $\frac{|PA|}{|PB|} = \lambda$ is a circle. We have just proven that this is true using complex numbers.

§15.4 *Extended Complex Plane* $\mathbf{C}_\infty$

§15.4.1 *Stereographic projection*

Let

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

be the unit sphere of radius 1 centred at the origin in $\mathbf{R}^3$. View the complex plane $\mathbf{C}$ as the copy of $\mathbf{R}^2$ inside $\mathbf{R}^3$ given by the plane $\{(x, y, 0) \in \mathbf{R} \mid x, y \in \mathbf{R}\}$. Thus $z = x + iy$ corresponds to the point $(x, y, 0)$. Let N be the "north pole" $N = (0, 0, 1)$ of $\mathbf{S}$.

We can define a bijective map $S : \mathbf{C} \to \mathbf{S} \setminus \{N\}$ as follows. To determine $S(z)$, join z to N by a straight line, and let $S(z)$ be the point where this line meets the sphere $\mathbf{S}$. This map (or more accurately its inverse) is called ***stereographic projection***.

It is not too hard to give an explicit formula for $S(z)$.

Lemma 15.6. Suppose that $z = x + iy$. Then

$$S(z) = \left(\frac{2x}{x^2 + y^2 + 1}, \frac{2y}{x^2 + y^2 + 1}, \frac{x^2 + y^2 - 1}{x^2 + y^2 + 1} \right).$$

Proof. The general point on the line joining z and N is $t(0, 0, 1) + (1 - t)(x, y, 0)$. There is a unique value of t for which this point lies on the sphere, namely $t = \frac{x^2+y^2-1}{x^2+y^2+1}$, as can be easily checked. □

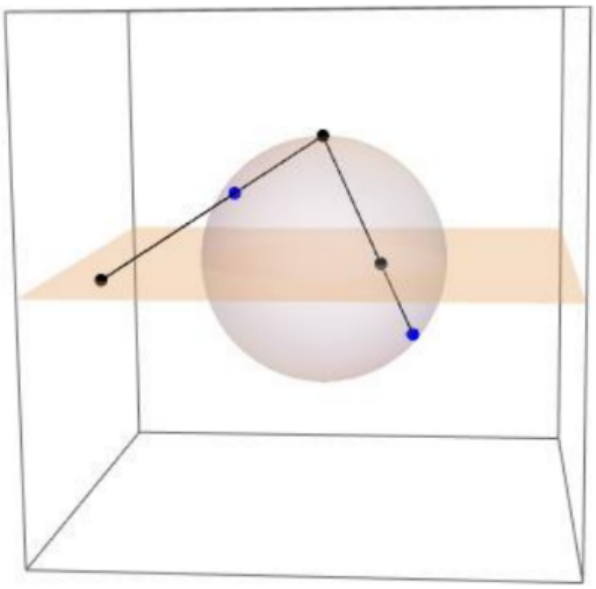

Figure 15.1: The stereographic projection map

We remark that the same formula can be written in the alternative form

$$S(z) = \frac{1}{1 + |z|^2} \left(2\,\mathrm{Re}(z), 2\,\mathrm{Im}(z), |z|^2 - 1 \right).$$

As we have seen, $\mathbf{C}$ may be identified with $\mathbf{S} \setminus \{N\}$ by stereographic projection. The set $\mathbf{S} \setminus \{N\}$ has a natural metric, namely the one induced from the Euclidean metric on $\mathbf{R}^3$. This induces a metric d on $\mathbf{C}$, the unique metric on $\mathbf{C}$ such that $\mathbf{S}$ is an isometry. To spell it out,

$$d(z, w) := \left\| S(z) - S(w) \right\|.$$

Here is a formula for this metric.

Lemma 15.7. For any $z, w \in \mathbf{C}$ we have

$$d(z,w) = \frac{2|z-w|}{\sqrt{1+|z|^2}\sqrt{1+|w|^2}}.$$

Proof. Since $\|S(z)\| = \|S(w)\| = 1$ we have $\|S(z) - S(w)\|^2 = 2 - 2\langle S(z), S(w)\rangle$, where $\langle , \rangle$ is the usual Euclidean inner product on $\mathbf{R}^3$. Using the formulae (and after a little computation),

$$\langle S(z), S(w)\rangle = 1 - \frac{2|z-w|^2}{(1+|z|^2)(1+|w|^2)}.$$

Therefore

$$\|S(z) - S(w)\|^2 = \frac{4|z-w|^2}{(1+|z|^2)(1+|w|^2)}$$

as required. □

§15.4.2 *Adding in* ∞

Now it is time to add in the point at infinity, which we will call ∞ (note this is just a symbol).

Now (exercise) as $|z| \to \infty$, $S(z) \to N$. Therefore, once we have identified $\mathbf{C}$ with $\mathbf{S} \setminus \{N\}$, it is natural to identify ∞ with N, and hence $\mathbf{C}_\infty = \mathbf{C} \cup \{\infty\}$ with the whole sphere $\mathbf{S}$. We extend the map S to a map $S : \mathbf{C}_\infty \to \mathbf{S}$ by defining $S(\infty) = N$.

Using, once again, the Euclidean metric on S, we can extend d to a metric on $\mathbf{C}_\infty$, the unique metric for which the map S is an isometry.

Lemma 15.8. For any $z \in \mathbf{C}$ we have

$$d(z,\infty) = \frac{2}{\sqrt{1+|z|^2}}.$$

Proof. By definition, $d(z,\infty) = \|S(z) - S(\infty)\| = \|S(z) - N\|$, where N is the north pole on the sphere. We may now proceed in much the same way as before, except the calculation is easier this time. The details are left as an exercise. □

We turn now to a few examples, which show that adding ∞ to $\mathbf{C}$ in this way leads to a space with nice analytic properties.

Example 15.9 (Translations)
Let $a \in \mathbf{C}$. Define $f : \mathbf{C}_\infty \to \mathbf{C}_\infty$ by $f(z) = z + a$ for $z \in \mathbf{C}$ and $f(\infty) = \infty$. Then f is a continuous bijection.

Proof. Clearly f is continuous with respect to the usual metric on $\mathbf{C}$. Therefore, restricted to $\mathbf{C}$, it is also continuous with respect to d, since d is equivalent to the usual metric.

It remains to check continuity at ∞. Let $\epsilon > 0$. Now if $\delta > 0$ and if $d(z,\infty) < \delta$ then $|z| > \sqrt{\frac{4}{\delta^2} - 1}$ and so $|f(z)| > \sqrt{\frac{4}{\delta^2} - 1} - |a|$. This tends to ∞ as $\delta \to 0$, so by choosing δ small enough in terms of ϵ it will follows that

$$d\left(f(z),\infty\right) = \frac{2}{\sqrt{1+|f(z)|^2}} < \epsilon.$$

□

Example 15.10 (Dilations)
Let $b \in \mathbf{C}$. Define $f : \mathbf{C}_\infty \to \mathbf{C}_\infty$ by $f(z) = bz$ for $z \in \mathbf{C}$ and $f(\infty) = \infty$. Then f is a continuous bijection.

Proof. This is very similar to the argument for translations and we leave the details as an exercise. □

The final example is the most interesting one.

Example 15.11 (Inversion)
Define $f : \mathbf{C}_\infty \to \mathbf{C}_\infty$ by $f(z) = \frac{1}{z}$ for $z \in \mathbf{C} \setminus \{0\}$, $f(0) = \infty$ and $f(\infty) = 0$. Then f is a continuous bijection.

Proof. As before, the equivalence of d and the usual metric on $\mathbf{C}$ means that f is continuous except possibly at 0 and ∞.

We prove that f is continuous at 0, leaving the continuity at ∞ as an exercise (similar to the example on translations).

Let $\epsilon > 0$ be small. Then there is δ such that $\frac{2t}{\sqrt{1+t^2}} \leq \epsilon$ for all $t \in [0, \delta]$. If $|z| < \delta$ then

$$d\left(f(z), f(0)\right) = d\left(\frac{1}{z}, \infty\right) = \frac{2}{\sqrt{1 + \frac{1}{|z|^2}}} = \frac{2|z|}{\sqrt{1+|z|^2}} \leq \epsilon.$$

This indeed shows that f is continuous at 0. □

§15.4.3 *Möbius maps*

In this subsection and subsequent ones we look at an important class of maps from $\mathbf{C}_\infty$ to itself, the Möbius maps.

§15.4.4 *The complex projective line* $\mathbf{P}^1(\mathbf{C})$

§15.4.5 *Decomposing Möbius maps*

§15.4.6 *Basic geometry of Möbius maps*

16 Complex Functions

§16.1 *Complex Differentiability*

Definition 16.1. Suppose $a \in \mathbf{C}$, $U \subseteq \mathbf{C}$.

(i) The ***open ball*** of radius $r > 0$ centred at a is defined as

$$B_r(a) \coloneqq \{z \in \mathbf{C} \mid |z - a| < r\}.$$

For the ***closed ball***, we use the condition $|z - a| \leq r$ instead.

(ii) U is ***open*** if for every $z \in U$ there exists an open ball $B_r(z) \subset U$.

(iii) a is a ***boundary point*** of U if every $B_r(a)$ contains both points of U and points not in U.

(iv) a is an ***interior point*** of U if there exists $B_r(a) \subset U$.

(v) U is ***closed*** if it contains all its boundary points. The complement of a closed set is then open.

(vi) The closure of U is defined to be the union of U and all its boundary points.

(vii) U is ***bounded*** if there exists $C > 0$ such that $|z| \leq C$ for all $z \in U$.

Definition 16.2 (Limit)**.** For $f : U \setminus \{a\} \to \mathbf{C}$, we say that $\lim\limits_{z \to a} f(z) = w$ if $\forall \epsilon > 0$, $\exists \delta > 0$ such that $\forall z \in U$,

$$0 < |z - a| < \delta \implies |f(z) - w| < \epsilon.$$

Definition 16.3 (Continuity)**.** Let $a \in U$. We say that f is ***continuous*** at a if

$$\lim_{z \to a} f(z) = f(a).$$

Definition 16.4 (Convergence)**.** Given a sequence of complex numbers $(z_n)_{n \in \mathbf{N}}$, we say that $w = \lim\limits_{n \to \infty} z_n$ if $\forall \epsilon > 0$, $\exists N \in \mathbf{N}$ s.t. $\forall n \geq N$, $|z_n - w| < \epsilon$.

Definition 16.5 (Cauchy sequence)**.** A sequence (z_n) is a ***Cauchy sequence*** if $\forall \epsilon > 0$, $\exists N \in \mathbf{N}$ s.t. $\forall m, n \geq N$, $|z_n - z_m| < \epsilon$.

Definition 16.6 (Complex differentiability)**.** Let $a \in \mathbf{C}$, and suppose that $f : U \to \mathbf{C}$ is a function, where U is a neighbourhood of a. In particular, f is defined on some ball $B_r(a)$. Then we say that f is (complex) differentiable at a if

$$\lim_{z \to a} \frac{f(z) - f(a)}{z - a}$$

exists. If the limit exists, we write $f'(a)$ for it and call this the derivative of f at a.

Since we will be talking exclusively about functions on $\mathbf{C}$, we just use the terms differentiable/derivative and omit the word "complex". The following lemma collects the basic facts about derivatives. We omit the proof, which is essentially identical to the real case.

Lemma 16.7. Let $a \in \mathbf{C}$, let U be a neighbourhood of a and let $f, g : U \to \mathbf{C}$.

(1) (Sums, products) If f, g are differentiable at a, then $f + g$ and fg are differeitnable at a, and

$$(f+g)'(a) = f'(a) + g'(a)$$

and

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

(2) (Quotients) If f, g are differentiable at a and $g(a) \neq 0$ then f/g is differentiable at a and

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

(3) (Chain rule) If U and V are open subsets of $\mathbf{C}$ and $f : V \to U$, $g : U \to \mathbf{C}$, where f is differentiable at $a \in V$ and g is differeitnable at $f(a) \in U$, then $g \circ f$ is differentiable at a, with

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Example 16.8

$f(z) = 1$ and $f(z) = z$ are analytic functions from $\mathbf{C}$ to $\mathbf{C}$, with derivatives $f'(z) = 0$ and $f'(z) = 1$ respectively.

Therefore, all polynomials $f(z) = a_n z^n + \cdots + a_1 z + a_0$ are analytic, with $f'(z) = na_n z^{n-1} + \cdots + a_1$.

Just as in the real-variable case one can formulate complex differentiability in the following form, which is in fact the better form to use in most instances.

Lemma 16.9. Let $a \in \mathbf{C}$, let U be a neighbourhood of a and let $f : U \to \mathbf{C}$. Then f is differentiable at a, with derivative $f'(a)$, if and only if we have

$$f(z) = f(a) + f'(a)(z-a) + \epsilon(z)(z-a) \tag{16.1}$$

where $\epsilon(z) \to 0$ as $z \to a$.

It is an easy exercise to check that this definition is indeed equivalent to (really just a reformulation of) the previous one.

Finally, we give an important definition.

Definition 16.10 (Holomorphic function)**.** Let $U \subseteq \mathbf{C}$ be an open set (for example, a domain). Let $f : U \to \mathbf{C}$ be a function. If f is complex differentiable at every $a \in U$, we say that f is holomorphic on U.

§16.1.1 *Cauchy–Riemann Equations*

A function from $\mathbf{C}$ to $\mathbf{C}$ may also be thought of as a function from $\mathbf{R}^2$ to $\mathbf{R}^2$, and it is useful to study what differentiability means in this language.

Let $a \in \mathbf{C}$, and let U be a neighbourhood of a. Let $f : U \to \mathbf{C}$ be a function. We abuse notation and identify $\mathbf{C} \cong \mathbf{R}^2$ in the usual way, and identify a with (a_1, a_2) (thus $a = a_1 + ia_2$). Then (again with some abuse of notation) we may think of U as an open subset of $\mathbf{R}^2$ and write $f = (u, v)$, where $u, v : \mathbf{R}^2 \to \mathbf{R}$ (the letters u, v are quite traditional in this context, and sometimes we call these the components of f). Another way to think of this is that $f(x + iy) = u(x, y) + iv(x, y)$.

Example 16.11

Consider the function $f(z) = z^2$ (which is holomorphic on all of $\mathbf{C}$). Since $(x + iy)^2 = (x^2 - y^2) + 2ixy$, we see that the components of f are given by $u(x, y) = x^2 - y^2$, $v(x, y) = 2xy$.

We have the partial derivatives

$$\frac{\partial u(a)}{\partial x} := \lim_{h\to 0} \frac{u(a_1+h, a_2) - u(a_1, a_2)}{h}$$

(if the limit exists) and

$$\frac{\partial u(a)}{\partial y} := \lim_{k\to 0} \frac{u(a_1, a_2+k) - u(a_1, a_2)}{k},$$

and similarly for v. It is important to note that h, k in these limits are real.

An important fact is that if f is differentiable then these partial derivatives do exist, and moreover they are subject to a constraint.

Theorem 16.12 (Cauchy–Riemann equations)**.** Let $a \in \mathbf{C}$, let U be a neighbourhood of a, and let $f : U \to \mathbf{C}$ be a function which is complex differentiable at a. Let $u, v : \mathbf{R}^2 \to \mathbf{R}$ be the components of f. Then the four partial derivatives $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}$ exist at a. Moreover, we have the Cauchy–Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \tag{16.2}$$

and $f'(a) = \dfrac{\partial u(a)}{\partial x} + i\dfrac{\partial v(a)}{\partial x}$.

Proof. Write $f(z) = u(z) + iv(z)$, where $u, v : \Omega \to \mathbf{R}$ are real-valued functions. Suppose f is analytic. We compare two ways of taking the limit $f'(z)$:

First take h to be a real number approaching 0. Then

$$f'(z) = \frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}.$$

Next, take h to be purely imaginary, i.e., let $h = ik$ for some $k \in \mathbf{R}$. Then

$$f'(z) = \lim_{k\to 0} \frac{f(z+ik) - f(z)}{ik} = -i\frac{\partial f}{\partial y} = -i\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}.$$

Comparing real and imaginary parts, we obtain

$$\frac{\partial f}{\partial x} = -i\frac{\partial f}{\partial y},$$

or, equivalently,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

□

Assuming for the time being that u, v have continuous partial derivatives of all orders (and in particular the mixed partials are equal), we can show that

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \Delta v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0.$$

Such an equation $\Delta u = 0$ is called Laplace's equation and its solution is said to be a harmonic function.

Let us pause to give a simple example using the Cauchy-Riemann equations, which shows that complex differentiation is a much more rigid property than one might think at first sight.

Example 16.13

The function $f(z) = \bar{z}$ is not (complex) differentiable anywhere.

Proof. Let $u, v : \mathbf{R}^2 \to \mathbf{R}$ be the components of f. Then clearly $u(x, y) = x$, $v(x, y) = -y$ and so $\partial_x u = 1, \partial_y u = 0, \partial_x v = 0, \partial_y v = -1$. Thus $\partial_x u$ is never equal to $\partial_y v$, so the Cauchy–Riemann equations are never satisfied. □

Part VI

Topology

17 Topological Spaces and Continuous Functions

§17.1 *Topological Spaces*

Definition 17.1 (Topological space)**.** A ***topological space*** $(X, \mathcal{T})$ consists of a non-mepty set X together with a family $\mathcal{T}$ of subsets of X satisfying:

(1) $X, \emptyset \in \mathcal{T}$;

(2) if $U, V \in \mathcal{T}$, then $U \cap V \in \mathcal{T}$;

(3) If $U_i \in \mathcal{T}$ for all $i \in I$, then $\bigcup_{i \in I} U_i \in \mathcal{T}$.

The family $\mathcal{T}$ is called a ***topology*** for X. The sets in $\mathcal{T}$ are called the ***open sets*** of X. When $\mathcal{T}$ is understood we talk about the topological space X.

Remark. A consequence of (2) is that if $U_1, \ldots, U_n$ is a collection of open sets, then $U_1 \cap \cdots \cap U_n$ is open. But the intersection of infinitely many open sets need not be open!

On the other hand, in (3), the indexing set I is allowed to be infinite. It may even be uncountable.

Proposition 17.2. Let (X, d) be a metric space. Then the open subsets of X form a topology, denoted by $\mathcal{T}_d$.

Proof. Check through the conditions in the definition for a topological space:

(1) Trivial.

(2) Let U and V be open subsets of X. Consider an arbitrary point $x \in U \cap V$.

As U is open, there exists $r_1 > 0$ such that $B_{r_1}(x) \subseteq U$. Likewise, as $x \in V$ and V is open, there exists $r_2 > 0$ such that $B_{r_2}(x) \subseteq V$.

Take $r \coloneqq \min\{r_1, r_2\}$. Then $B_r(x) \subseteq B_{r_1}(x) \subseteq U$ and $B_r(x) \subseteq B_{r_2}(x) \subseteq V$. Hence $B_r(x) \subseteq U \cap V$.

(3) For every $x \in \bigcup_{i \in I} U_i$ there exists $k \in I$ such that $x \in U_k$. Since U_k is open, there exists $r > 0$ such that $B_r(x) \subseteq U_k \subseteq \bigcup_{i \in I} U_i$.

□

Example 17.3

The following are some other examples of topological spaces.

- (Discrete spaces) Let X be any non-empty set. The **discrete topology** on X is the set of all subsets of X.

- (Indiscrete spaces) Let X be any non-empty set. The **indiscrete topology** on X is the family of subsets $\{X, \emptyset\}$.
- Let X be any non-empty set. The **co-finite topology** on X consists of the empty set together with every subset U of X such that $X \setminus U$ is finite.

Definition 17.4. A topological space $(X, \mathcal{T})$ is ***metrisable*** if it arises from (at least oe) metric space (X, d), i.e. there is at least one metric d on X such that $\mathcal{T} = \mathcal{T}_d$.

Definition 17.5. Two metrics on a set are ***topologically equivalent*** if they give rise to the same topology.

Example 17.6

- The metrics d_1, d_2, d_∞ on $\mathbf{R}^n$ are all topologically equivalent. (Recall that d_1, d_2, d_∞ are the metrics arising from the norms $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$, respectively.) We shall call the topology defined by the above metrics the **standard** (or canonical) topology on $\mathbf{R}^n$.
- The discrete topology on a non-empty set X is metrisable, using the metric

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

  It is easy to check that this is a metric. To see that is gives the discrete topology, consider any subset $U \subseteq X$. Then for every $x \in U$, $B_{\frac{1}{2}}(x) \subseteq U$.

Definition 17.7. Given two topologies $\mathcal{T}_1$ and $\mathcal{T}_2$ on the same set, we say $\mathcal{T}_1$ is ***coarser*** than $\mathcal{T}_2$ if $\mathcal{T}_1 \subseteq \mathcal{T}2$.

Remark. For any space $(X, \mathcal{T})$, the indiscrete topology on X is coarser than $\mathcal{T}$ which in turn is coarser than the discrete topology on X.

Definition 17.8. Let $(X, \mathcal{T})$ be a topological space. A subset V of X is ***closed*** in X if $X \setminus V$ is open in X (i.e. $X \setminus V \in \mathcal{T}$).

Example 17.9

- In the space $[0, 1)$ with the usual topology coming from the Euclidean metric, $[1/2, 1)$ is closed.
- In a discrete space, all subsets are closed since their complements are open.
- In the co-finite topology on a set X, a subset is closed if and only if it is finite or all of X.

Proposition 17.10. Let X be a topological space. Then

(1) X, $\emptyset$ are closed in X;

(2) if V_1, V_2 are closed in X then $V_1 \cup V_2$ is closed in X;

(3) if V_i is closed in X for all $i \in I$ then $\bigcap_{i \in I} Vi$ is closed in X.

Proof. These properties follow from (1), (2), (3) of definition of topological space, and from the De Morgan laws. □

Definition 17.11 (Convergent sequence)**.** A sequence $\{x_n\}_{n \in \mathbf{N}}$ in a topological space X converges to a point $x \in X$ if given any open set U containing x there exists $N \in \mathbf{N}$ such that $x_n \in U$ for all $n > N$.

Example 17.12

- In a metric space this is equivalent to the metric definition of convergence.
- In an indiscrete topological space X any sequence converges to any point $x \in X$.
- In an infinite space X with the co-finite topology any sequence $\{x_n\}$ of pairwise distinct elements (i.e. such that $x_n \neq x_m$ when $n \neq m$) converges to any point $x \in X$.

Bibliography

[Ahl79] L. V. Ahlfors. *Complex Analysis*. McGraw-Hill, 1979.

[Apo57] T. M. Apostol. *Mathematical Analysis*. Addison-Wesley, 1957.

[Axl15] S. Axler. *Linear Algebra Done Right*. Springer International Publishing, 2015.

[DF04] D. S. Dummit and R. M. Foote. *Abstract Algebra*. John Wiley & Sons, 2004.

[Lan99] Serge Lang. *Complex Analysis*. Springer International Publishing, 1999.

[Mun18] J. R. Munkres. *Topology*. Pearson Education Limited, 2018.

[Pól45] G. Pólya. *How to Solve It*. Princeton University Press, 1945.

[Rud53] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 1953.

[Sch92] A. H. Schoenfeld. "Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics". In: *Handbook for Research on Mathematics Teaching and Learning*. Macmillan, 1992, pp. 334–370.