Real analysis

Peter Burton

June 20, 2019

Contents

1 Review of discrete mathematics **4**
1.1 Sets . . . 4
1.1.1 Basic information . . . 4
1.1.2 Subsets . . . 4
1.1.3 Boolean operations . . . 5
1.2 Functions . . . 6
1.2.1 Basic information . . . 6
1.2.2 Three essential properties . . . 7
1.3 The natural numbers . . . 9
1.3.1 Arithmetic structure . . . 9
1.3.2 Order structure . . . 9
1.4 Finite sets . . . 10
1.5 Infinite sets . . . 10
1.6 Mathematical induction . . . 11
1.6.1 The method . . . 11
1.6.2 Justification from well-ordering . . . 11
1.6.3 First example . . . 12
1.6.4 Delayed induction . . . 13
1.7 Integers and rational numbers . . . 13
1.7.1 Integers . . . 13
1.7.2 Countably infinite sets . . . 14
1.7.3 Rational numbers . . . 14

2 The real number system **16**
2.1 The real numbers . . . 16
2.1.1 Motivation . . . 16
2.1.2 Fourteen axioms . . . 16
2.1.3 Embedding $\mathbb{Q}$ in $\mathbb{R}$ . . . 19
2.1.4 Examples of axiomatic proofs . . . 19
2.1.5 Existence of $\sqrt{2}$ . . . 21
2.2 Uncountable sets . . . 23
2.2.1 Uncountability of the set of binary sequences . . . 23
2.2.2 Uncountability of the real numbers . . . 24
2.3 Absolute value and distance in the real line . . . 27

3 Sequences and series **29**
3.1 Sequences of real numbers 29
3.1.1 Basic information 29
3.1.2 Examples 30
3.1.3 Uniqueness of limits of sequences 31
3.1.4 Arithmetic of sequences 31
3.1.5 Squeeze theorem 32
3.1.6 Monotone sequences 33
3.1.7 Subsequences and the Bolzano-Weierstrass theorem 33
3.1.8 Cauchy sequences 35
3.2 Infinite sums 36
3.2.1 Basic information 36
3.2.2 Zero test 36
3.2.3 Examples 36
3.2.4 Convergence criteria 39

4 Limits and continuity **42**
4.1 Functions of a real variable 42
4.1.1 Examples 42
4.1.2 Operations on functions 43
4.2 Limits 44
4.2.1 Basic information 44
4.2.2 Uniqueness of limits 45
4.2.3 Examples 45
4.2.4 Manipulation of limits 47
4.2.5 Squeeze theorem 48
4.2.6 Extended limits 48
4.2.7 One-sided limits 49
4.3 Continuity 50
4.3.1 Basic information 50
4.3.2 Examples 50
4.3.3 Continuity and composition 52
4.3.4 Continuity and sequences 52
4.3.5 Intermediate value theorem 53
4.3.6 Continuity and compactness 55
4.3.7 Uniform continuity 57

5 Differentiation **60**
5.1 Fundamentals 60
5.1.1 Basic information 60
5.1.2 Differentiation and continuity 61
5.1.3 Examples 61
5.1.4 Rules for taking derivatives 62
5.2 Essential theorems about derivatives 64
5.2.1 Interior extremum theorem 65
5.2.2 Rolle's theorem 65

5.2.3 Mean value theorem 66

6 Integration **67**
6.1 Fundamental definitions 67
6.2 Uniqueness of the integral 67
6.3 Criteria for integrability 68
6.4 Integrability of continuous functions 69
6.5 First fundamental theorem 70
6.6 Second fundamental theorem 70

Chapter 1

Review of discrete mathematics

1.1 Sets

1.1.1 Basic information

A **set** is a collection of elements. If A is a set and x is an element of A, we denote this by $x \in A$. For example, I am an element of the set of human beings. One can think of membership in a set as being a property of the elements. Sets are the foundation of mathematics. We will not investigate sets axiomatically in this text, as for our purposes an intuitive understanding of sets is sufficient.

We use braces $\{\cdots\}$ to delimit sets. For example, the standard set of primary colors is

$$\{\text{red}, \text{blue}, \text{yellow}\}$$

We stipulate that two sets are equal if they have the same elements.

There is a special set called the **empty set** which has no elements. We denote the empty set by $\emptyset$.

1.1.2 Subsets

If A and B are sets, we say that A is a **subset** of B if every element of A is also an element of B. We denote this by $A \subseteq B$. For example,

$$\{\text{red}, \text{yellow}\} \subseteq \{\text{red}, \text{blue}, \text{yellow}\}.$$

We can specify subsets of a set A by specifying a property P that they might have. We use the notation

$$\{x \in A : x \text{ has the property } P\}.$$

For example let

$$A = \{\text{red}, \text{blue}, \text{yellow}\}$$

Then we have

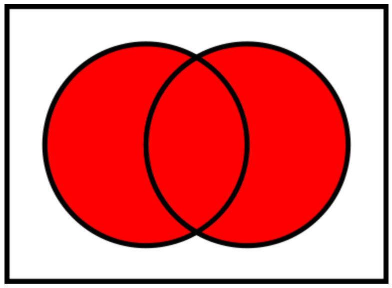
$$\{x \in A : x \text{ starts with the letter b}\} = \{\text{blue}\}$$

Two sets A and B are equal if and only if both of the conditions $A \subseteq B$ and $B \subseteq A$ hold.

1.1.3 Boolean operations

Union

The **union** of A and B is the collection of objects which are elements of A or elements of B. We write $A \cup B$ for the union of A and B. The red area in the graphic below is the union of the two discs.



As another example, we have

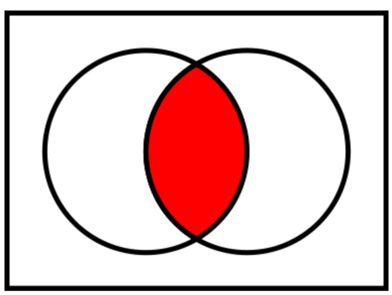
$$\{\text{red}, \text{blue}\} \cup \{\text{green}, \text{yellow}\} = \{\text{red}, \text{blue}, \text{green}, \text{yellow}\}$$

It is important to remember that elements of sets do not have multiplicity - an element is only counted once. Thus we have

$$\{\text{red}, \text{blue}\} \cup \{\text{blue}, \text{yellow}\} = \{\text{red}, \text{blue}, \text{yellow}\}$$

Intersection

The **intersection** of A and B is the collection of objects which are elements of A and elements of B. We write $A \cap B$ for the intersection of A and B. The red area in the graphic below is the intersection of the two discs.



As another example, we have

$$\{\text{red}, \text{blue}\} \cap \{\text{blue}, \text{yellow}\} = \{\text{blue}\}$$

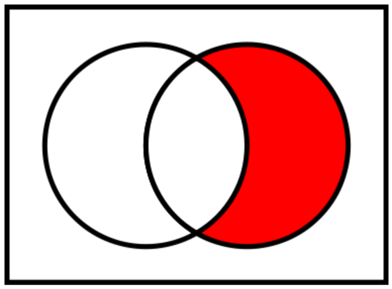
If the intersection of two sets is empty we say that the two sets are **disjoint**. For example,

$$\{\text{red}, \text{blue}\} \cap \{\text{green}, \text{yellow}\} = \emptyset \tag{1.1.1}$$

so that the two sets on the left side of (1.1.1) are disjoint.

Difference

The **difference** A minus B is the collection of objects which are elements of A but not elements of B. We write $A \setminus B$ for the difference A minus B. The red area in the graphic below is the right disc minus the left disc.



As another example, we have

$$\{\text{red}, \text{blue}\} \setminus \{\text{blue}, \text{yellow}\} = \{\text{red}\}$$

1.2 Functions

1.2.1 Basic information

The next level of mathematical structure on top of sets is the concept of a function. If A and B are sets, a **function** f from A to B is an object that takes any element of A as an input and returns a unique element of B as an output. We use the notation $f : A \to B$ to denote a function from A to B. We refer to the set A as the **domain** of f and we refer to the set B as the **codomain** of f. The set

$$\{y \in B : y = f(x) \text{ for some } x \in A\}$$

is called the **range** of f.

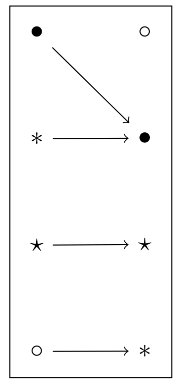
For example, let

$$A = \{\text{red}, \text{blue}, \text{yellow}\}$$

and let B be the English alphabet. We can define a function $f : A \to B$ by letting $f(x)$ be the second letter of x. We have

- $f(\text{red}) = \text{e}$
- $f(\text{blue}) = \text{l}$
- $f(\text{yellow}) = \text{e}$

Below is a graphical visualization of a function. The column of objects on the left is the domain and the column of objects on the right is the codomain. The arrows point from inputs to outputs. The essential property is that every object on the left has exactly one arrow pointing out of it.



1.2.2 Three essential properties

Injectivity

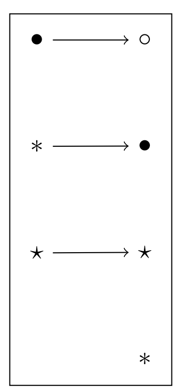
A function $f : A \to B$ is called **injective** or **one-to-one** if distinct inputs return distinct outputs. For example, let

$$A = \{\text{red}, \text{blue}, \text{yellow}\}$$

and let B be the English alphabet. If we define $f : A \to B$ by letting $f(x)$ be the first letter of x, then f is injective. The function returning the second letter is not injective, since

$$f(\text{red}) = \text{e} = f(\text{yellow})$$

Below is a graphical visualization of an injective function. The essential property is that each object on the right has at most one arrow is pointing at it.



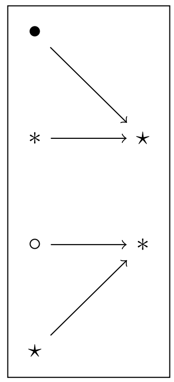
Surjectivity

A function $f : A \to B$ is called **surjective** or **onto** if every possible output has a corresponding input. For example, let

$$A = \{\text{red}, \text{blue}, \text{yellow}\}$$

and let B be the set $\{\text{e}, \text{l}\}$. If we define $f : A \to B$ by letting $f(x)$ be the second letter of x, then f is surjective. If we instead had $B = \{\text{e}, \text{l}, \text{k}\}$ then the second-letter function from A to B would not be surjective. A function is surjective if and only if its range is equal to its codomain.

Below is a graphical visualization of a surjective function. The essential property is that every object on the right has at least one arrow pointing at it.



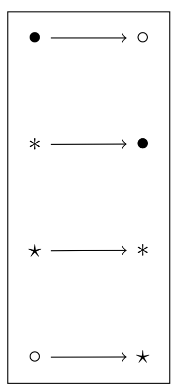
Bijectivity

A function $f : A \to B$ is called **bijective** if it is injective and surjective. Equivalently, every possible output is produced by exactly one input. A bijective function can be thought of as a one-to-one correspondence between the domain and the codomain. For example, let

$$A = \{\text{red}, \text{blue}, \text{yellow}\}$$

and let B be the set $\{\text{r}, \text{b}, \text{y}\}$. If we define $f : A \to B$ by letting $f(x)$ be the first letter of x, then f is bijective.

Below is a graphical visualization of a bijective function. The essential property is that every object on the right has exactly one arrow pointing at it.



1.3 The natural numbers

The **natural numbers** are the positive whole numbers $1, 2, 3, 4, \ldots$ We write $\mathbb{N}$ for the set of all natural numbers. The natural numbers carry two fundamental structures that we will use to start building our mathematical framework.

1.3.1 Arithmetic structure

We can add two natural numbers to get another natural number, and similarly we can multiply two natural numbers to get another natural number. Addition and multiplication are functions which take a pair of natural numbers as an input and return a single natural number as an output. Note that subtraction and division are not generally possible within the set of natural numbers.

1.3.2 Order structure

The natural numbers have an order relation $\leq$. For natural numbers n and k we define $n \leq k$ to mean that there exists an injective function from the set $\{1, \ldots, n\}$ to the set $\{1, \ldots, k\}$.

Essential properties

Here are the essential properties of the order relation.

(Reflexivity) For all $n \in \mathbb{N}$ we have $n \leq n$.

(Transitivity) Let $k, m, n \in \mathbb{N}$. If $k \leq m$ and $m \leq n$ then $k \leq n$.

(Antisymmetry) If $n \leq k$ and $k \leq n$ then $n = k$.

(Totality) For any pair $n, k \in \mathbb{N}$ we have $n \leq k$ or $k \leq n$.

These four properties will reappear several times as we develop more sophisticated forms of number.

Minimal elements

If $A \subseteq \mathbb{N}$, we will say that a natural number n is a **minimal element** of A if $n \in A$ and $n \leq k$ for all $k \in A$. The ordering of natural numbers has the following additional property.

Well-ordering principle. Every nonempty set of natural numbers has a minimal element.
The minimal element of a set of natural numbers represents the 'left endpoint'.

Proposition 1.3.1. *A nonempty set of natural numbers has a unique minimal element.*

Proof of Proposition 1.3.1. Suppose n and k are two minimal elements of a set A of natural numbers. We need to show that $n = k$.

The first clause in the definition of a minimal element implies that both n and k are elements of A. By applying the second clause directly we obtain $n \leq k$. By applying the second clause with the letters reversed we obtain $k \leq n$. Since $n \leq k$ and $k \leq n$, the antisymmetry property implies $n = k$. □

Proposition 1.3.1 makes it reasonable to speak of *the* minimal element of A.

1.4 Finite sets

We will say that a set A is **finite** if A is empty or if there exists a natural number n and a bijection between A and the set $\{1, \ldots, n\}$. For example, let

$$A = \{\text{red}, \text{blue}, \text{yellow}\}$$

This set is finite. We can define a bijection $f : A \to \{1, 2, 3\}$ by setting

- $f(\text{red}) = 1$
- $f(\text{blue}) = 2$
- $f(\text{yellow}) = 3.$

Finite sets are those which can enumerated with a terminating process.

Proposition 1.4.1. *For any nonempty finite set A, there is a unique natural number n such that there exists a bijection from A to $\{1, \ldots, n\}$.*

Proof of Proposition 1.4.1. Suppose that n and k are natural numbers such that there exist bijections $f : A \to \{1, \ldots, n\}$ and $g : A \to \{1, \ldots, k\}$. We need to show $n = k$.

The first idea is that a bijection can be inverted. In other words, if $h : B \to C$ is a bijection, then there exists a bijection $h^{-1} : C \to B$ such that $h^{-1}(h(x)) = x$ for all $x \in B$. The second idea is that the composition of two bijections is a bijection.

Therefore $g \circ f^{-1}$ is a bijection from $\{1, \ldots, n\}$ to $\{1, \ldots, k\}$. Here $g \circ f^{-1}$ is given by $(g \circ f^{-1})(m) = g(f^{-1}(m))$.

The existence of this bijection implies that $n \leq k$ and the fact that its inverse is a bijection implies $k \leq n$. Therefore antisymmetry implies $n = k$. □

The unique natural number n such that there exists a bijection between A and $\{1, \ldots, n\}$ is called the size or cardinality of A. We denote this by writing $|A| = n$. The empty set is the unique set with cardinality zero.

One can check that if A and B are finite sets then the following are equivalent.

- $|A| \leq |B|$
- There exists an injective function from A to B.
- There exists a surjective function from B to A.

1.5 Infinite sets

We say a set is **infinite** if it is not finite.

Proposition 1.5.1. *The whole set of natural numbers is infinite.*

Proof of Proposition 1.5.1. We use the technique of proof by contradiction. Suppose toward a contradiction that $\mathbb{N}$ were finite. Then there would exist a natural number n and a bijection $f : \mathbb{N} \to \{1, \ldots, n\}$.

Let $\overline{f} : \{1, \ldots, n+1\} \to \{1, \ldots, n\}$ be defined by $\overline{f}(k) = f(k)$ for all $k \in \{1, \ldots, n+1\}$. This process of restricting the domain of a bijective function always produces an injective function. Since $\overline{f}$ is an injective function from $\{1, \ldots, n+1\}$ to $\{1, \ldots, n\}$, we obtain $n + 1 \leq n$. This is obviously false. Thus, our initial hypothesis must also be false and so $\mathbb{N}$ is infinite. $\square$

A counterintuitive property of infinite sets is that there can exist an injective function from an infinite set to a proper subset of itself. For example, the function $f : \mathbb{N} \to \{2, 3, 4, \ldots, \}$ given by $f(n) = n + 1$ is injective, but its range misses the number 1.

Let E be the set of even natural numbers. The function $f : \mathbb{N} \to E$ given by $f(n) = 2n$ is also injective, but its range misses the whole infinite set of odd numbers.

The way to interpret this is that an infinite set can have the same size as a proper subset of itself.

1.6 Mathematical induction

1.6.1 The method

Mathematical induction is a method to prove statements about the natural numbers. Proofs by mathematical induction can be good practice for developing mathematical writing skills, as they are highly formalized and in some sense are intermediate between computations and more advanced proofs.

Let $P(n)$ be a property of a natural number n. For example, $P(n)$ could be the statement that $n \leq n^2$. The goal of a proof by induction is to prove that $P(n)$ is true for all $n \in \mathbb{N}$. The paradigm consists of two steps.

Initial step. In the initial step, one checks that $P(1)$ is true. This step is usually easy, but it is important to remember to perform it.

Inductive step. In the inductive step, one assumes that $P(n)$ holds for some natural number n. Then one uses this assumption to prove that $P(n + 1)$ holds.

If both steps are verified, one concludes that $P(n)$ is true for all $n \in \mathbb{N}$.

1.6.2 Justification from well-ordering

Recall the well-ordering principle for the natural numbers.

Well-ordering principle. Every nonempty set of natural numbers has a minimal element.

We can use the well-ordering principle to establish that the method of mathematical induction is valid.

Suppose that we have completed both steps for a given property P and let A be the set of natural numbers for which P fails. In order to justify the conclusion, we need to show that A is empty.

Suppose toward a contradiction that A were nonempty. Then by the well-ordering principle, A would have a minimal element a.

The initial step implies that $P(1)$ is true, so we must have $a > 1$. Therefore $a - 1$ is also a natural number. Since a is the least number for which P fails, it must be the case that P holds for $a - 1$.

The inductive step says that if P holds for some number n, then P also holds for $n + 1$. Therefore P holds for $a = (a - 1) + 1$. This contradicts the hypothesis that $a \in A$.

1.6.3 First example

We now give an example of a proof by mathematical induction.

Proposition 1.6.1. *Let $P(n)$ be following statement:*

$$1 + 2 + 3 + \cdots + (n-1) + n = \frac{n(n+1)}{2}$$

Then $P(n)$ holds for all $n \in \mathbb{N}$.

Proof of Proposition 1.6.1. We proceed to follow the method of induction.

Initial step. We check that $P(1)$ holds:

$$1 = \frac{1(1+1)}{2}.$$

Inductive step. Here we assume that $P(n)$ holds for a natural number n. Explicitly, this means we assume that

$$1 + 2 + 3 + \cdots + (n-1) + n = \frac{n(n+1)}{2}.$$

We need to show that $P(n+1)$ holds. Explicitly, this means we need to show that

$$1 + 2 + 3 + \cdots + (n-1) + n + (n+1) = \frac{(n+1)(n+2)}{2}.$$

We compute

$$\begin{aligned}
1 + 2 + 3 + \cdots + (n-1) + n + (n+1) &= (1 + 2 + 3 + \cdots + (n-1) + n) + n + 1 && (1.6.1)\\
&= \frac{n(n+1)}{2} + n + 1 && (1.6.2)\\
&= \frac{n^2+n}{2} + \frac{2n+2}{2}\\
&= \frac{n^2+3n+2}{2}\\
&= \frac{(n+1)(n+2)}{2}
\end{aligned}$$

Here we used the inductive assumption to pass from (1.6.1) to (1.6.2). □

1.6.4 Delayed induction

Mathematical induction also works if we start from some number m greater than 1. In this case, for our initial step we check that $P(m)$ is true. In the inductive step we assume that $P(n)$ holds for some $n \geq m$ and we need to show that $P(n+1)$ holds. After both of these steps have been verified, we conclude that P holds for all numbers greater than or equal to m.

We now consider an example of this 'delayed' induction. Recall the notion of a factorial:

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-2) \cdot (n-1) \cdot n.$$

Proposition 1.6.2. *Let $P(n)$ be the statement $n! \geq 2^n$. Then $P(n)$ holds for all $n \geq 4$*

The statement in Proposition 1.6.2 is not true for *every* natural number, since $1! = 1 < 2 = 2^1$.

Proof of Proposition 1.6.2. We proceed with induction starting from the number 4.

Initial step. We check $P(4)$:

$$4! = 2 \cdot 3 \cdot 4 = 24 \geq 16 = 2^4.$$

Inductive step. We assume that $P(n)$ is true, that is we assume $n! \geq 2^n$. We compute

$$\begin{aligned} (n+1)! &= n! \cdot (n+1) && (1.6.3) \\ &\geq 2^n \cdot (n+1) && (1.6.4) \\ &\geq 2^n \cdot 2 \\ &= 2^{n+1}. \end{aligned}$$

Here we used the inductive hypothesis to pass from (1.6.3) to (1.6.4). □

1.7 Integers and rational numbers

1.7.1 Integers

We observed that subtraction and division are not generally possible within the natural numbers. This is problematic, as in order to develop calculus we will certainly need these operations.

We can enable subtraction by adjoining zero and the negatives of natural numbers. This results in the integers

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

We use the symbol $\mathbb{Z}$ for the set of integers.

Like the natural numbers, the integers have an order structure. However, the well-ordering principle fails for the integers: the whole set of integers is nonempty and it has no minimal element. Nevertheless, something close to the well-ordering principle holds.

Given a set A of integers, a **lower bound** for A is an integer n such that $n \leq k$ for all $k \in A$. We say that n is a **greatest lower bound** for A if n is a lower bound for A and any other lower bound m for

A satisfies $m \leq n$. One can check that in the context of the natural numbers, a greatest lower bound is the same thing as a minimal element.

Greatest lower bound principle for $\mathbb{Z}$. If a nonempty set of integers has a lower bound, then it has a greatest lower bound.

We can establish that the greatest lower bound of a set of integers is unique if it exists using the same argument we used to show that the minimal element of a set of natural numbers is unique.

1.7.2 Countably infinite sets

If there exists a bijection between a set A and the natural numbers, we say that A is countably infinite. We say that A is countable if it is countably infinite or finite.

Intuitively, countable sets are no larger than the natural numbers.

The set of integers is countable. For example, we can define a bijection $f : \mathbb{N} \to \mathbb{Z}$ by

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

1.7.3 Rational numbers

Basic information

We can enable division by considering the rational numbers. A rational number is an expression of the form $\frac{n}{k}$ where n is an integer and k is a nonzero integer. We make the convention that two such expressions $\frac{n}{k}$ and $\frac{m}{l}$ denote the same rational number if $nl = mk$. We use the symbol $\mathbb{Q}$ to denote the set of rational numbers.

Avoiding division by zero

The qualification that k is nonzero in the characterization of rational numbers reflects the mathematical consensus that division by zero is not advisable. Let us try to understand why this consensus exists.

Suppose that we could assign a numerical value to the expression $\frac{1}{0}$. Then we should certainly have $0 \cdot \frac{1}{0} = 1$. On the other hand, basic considerations about rational numbers entail that $0 \cdot x = 0$ for all $x \in \mathbb{Q}$. Combining the last two expressions, we see that if we allow division by zero by then we are forced to conclude that $0 = 1$. Following the principle that $0 \cdot x = 0$ for all x, we obtain the conclusion that $0 = 0 \cdot x = 1 \cdot x = x$ for all $x \in \mathbb{Q}$.

Thus the only way to make sense of division by zero is to conclude that zero is the only rational number that exists. Since this conclusion is obviously inconsistent with the fundamental utility of numbers, we are compelled to avoid division by zero in all circumstances.

Artihmetic of rational numbers

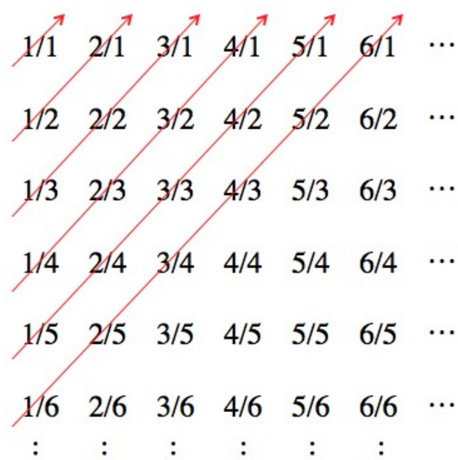
We can add and multiply rational numbers in the usual way:

$$\frac{n}{k} + \frac{m}{l} = \frac{nl + km}{kl}$$
$$\frac{n}{k} \cdot \frac{m}{l} = \frac{nm}{kl}$$

For the purposes of this text, we take the arithmetic of rational numbers as understood.

Countability of the rational numbers

The set of rational numbers is countable. Here is a picture of a surjective function from the natural numbers to the positive rational numbers.



If you want to, try to think about how to write down a formula that expresses the picture.

Chapter 2

The real number system

2.1 The real numbers

2.1.1 Motivation

The rational numbers are very nice arithmetically - one has all four basic operations. However, they are still not rich enough to develop calculus.

The reason is that the greatest lower bound property is essential for calculus, and it fails badly for the rational numbers. For example, consider the set

$$\{q \in \mathbb{Q} : q \geq 0 \text{ and } q^2 \geq 2\}.$$

This set has a lower bound of 0 by construction. However, it has no greatest lower bound in the rational numbers. The greatest lower bound 'should' be $\sqrt{2}$, but this is not a rational number.

In order to rectify this situation, we will need to extend our system to the real numbers. Intuitively, the real numbers are designed to fill in all the infinitesimal holes in the rational numbers. It is possible to construct the real numbers directly from the rational numbers, but instead we will take an axiomatic approach.

We write $\mathbb{R}$ for the set of real numbers. We will now proceed to list out fourteen axioms which characterize the real number system. These axioms will be the basis for all the work we do in the rest of the text.

2.1.2 Fourteen axioms

Addition is an abelian group

$\mathbb{R}$ is equipped with a binary operation called addition, which takes a pair of real numbers and returns a single real number. We denote addition with the infix notation $+$. The first four axioms describe the behavior of addition.

- Axiom 1 (associativity of addition) For all $x, y, z \in \mathbb{R}$ we have $(x + y) + z = x + (y + z)$.

- Axiom 2 (existence of additive identity) There exists a distinguished element $0 \in \mathbb{R}$ such that $0+x = x$ for all $x \in \mathbb{R}$.

- Axiom 3 (existence of additive inverses) For every $x \in \mathbb{R}$ there exists a corresponding element $-x$ such that $x + (-x) = 0$.

- Axiom 4 (commutativity of addition) For all $x, y \in \mathbb{R}$ we have $x + y = y + x$.

In algebraic terms, Axioms 1 - 3 assert that the real numbers with addition form an object called a group. When we add Axiom 4, the object is called an abelian group.

Axiom 1 belongs to the unusual class of mathematical statements which are so obvious to the intuition that their actual content is somewhat opaque. This axiom is necessary because we stipulated that addition was a binary operation, so *a priori* it does not make sense to add three real numbers. Axiom 1 asserts the equivalence of two plausible methods for decomposing the sum of three real numbers into iterated binary sums. Thus we can unambiguously speak of the sum of three real numbers. Repeated applications of Axiom 1 extend this to the sum of any finite list of real numbers. Even in our first axiomatic proofs below in Subsection 2.1.4, we will use Axiom 1 implicitly by writing expressions such as $x + y + z$.

If x and y are real numbers, we will typically write $x - y$ for $x + (-y)$. If $x_1, \ldots, x_n$ is a finite list of real numbers we will typically write $\sum_{k=1}^{n} x_k$ for the sum $x_1 + x_2 + \cdots + x_{n-1} + x_n$. Observe that in this notation the variable k is bound by the sum. Thus we must avoid using the variable k for another purpose in an expression involving the notation $\sum_{k=1}^{n} x_k$.

Multiplication without zero is an abelian group

$\mathbb{R}$ is equipped with a second binary operation called multiplication, which we denote by $\cdot$ or simply by juxtaposition. The next four axioms describe the behavior of multiplication.

- Axiom 5 (associativity of multiplication) For all $x, y, z \in \mathbb{R}$ we have $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

- Axiom 6 (existence of multiplicative identity) There exists a distinguished element $1 \in \mathbb{R}$ such that $1 \cdot x = x$ for all $x \in \mathbb{R}$.

- Axiom 7 (existence of multiplicative inverses) For each $x \in \mathbb{R}$ such that $x \neq 0$ there exists a corresponding element x^{-1} such that $x \cdot x^{-1} = 1$.

- Axiom 8 (commutativity of multiplication) For all $x, y \in \mathbb{R}$ we have $x \cdot y = y \cdot x$.

These axioms assert that the nonzero real numbers form an abelian group under multiplication. As with Axiom 1 for addition, Axiom 5 allows us to multiply any finite list of real numbers.

Later in the text we will develop the theory of exponentiation, in which the association $x \mapsto x^{-1}$ described by Axiom 7 will be generalized to a family of operations where the superscript -1 can replaced by other real numbers. For the purposes of the axiomatization we regard the superscript -1 as purely a formal symbol which can be manipulated in accordance with Axiom 7.

If x and y are real numbers, we will typically write xy for $x \cdot y$. For $n \in \mathbb{N}$ we will typically write x^n for $x \cdot x \cdots x \cdot x$, where there are n terms in the product. If y is nonzero we will typically write $\frac{x}{y}$ for $x \cdot y^{-1}$.

In particular, in accordance with Axiom 6 we will write $\frac{1}{y}$ for $y^{-1} = 1 \cdot y^{-1}$. If $x_1, \ldots, x_n$ is a finite list of real numbers, we will typically write $\prod_{k=1}^{n} x_k$ for the product $x_1 \cdot x_2 \cdots x_{n-1} \cdot x_n$.

Inherent in Axiom 7 is that division by zero is not advisable.

Multiplication distributes over addition

The next axiom describes how addition and multiplication interact with each other.

- Axiom 9 (distributivity of multiplication over addition) For all $x, y, z \in \mathbb{R}$ we have $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

Nondegeneracy

- Axiom 10 (nondegeneracy) The numbers 0 and 1 are distinct.

Axiom 10 is necessary for the simple reason that explicitly stipulating $0 \neq 1$ is the only way to rule out the undesirable possibility that $0 = 1$. In accordance with the discussion in Subsection 1.7.3, we see that Axiom 10 is intimately connected with the avoidance of division by zero.

Taken together, Axioms 1 - 10 describe the properties of an algebraic object called a field.

Order structure

The real numbers have a special subset called the positive numbers. We denote the positive numbers by $\mathbb{P}$. The set $\mathbb{P}$ has the following properties.

- Axiom 11 (stability of $\mathbb{P}$ under addition) If $x, y \in \mathbb{P}$ then $x + y \in \mathbb{P}$.
- Axiom 12 (stability of $\mathbb{P}$ under multiplication) If $x, y \in \mathbb{P}$ then $x \cdot y \in \mathbb{P}$.
- Axiom 13 (trichotomy property) For any $x \in \mathbb{R}$ exactly one of the following conditions holds:

  (x is positive) We have $x \in \mathbb{P}$.

  (x is zero) We have $x = 0$.

  (x is negative) We have $-x \in \mathbb{P}$.

We can use $\mathbb{P}$ to define an ordering on the real numbers. We stipulate that $x < y$ if and only if $y - x \in \mathbb{P}$. We also use the notation $x \leq y$ to mean $x < y$ or $x = y$.

By applying Axiom 3 to $x = 0$ we see that $0 + (-0) = 0$. By applying Axiom 2 to the previous equality we see that $-0 = 0$. Therefore if $x \in \mathbb{P}$ we can apply Axiom 2 to see that $x - 0 = x + 0 = x \in \mathbb{P}$. This justifies the more common notation $x > 0$ for the assertion that $x \in \mathbb{P}$.

Completeness

Axioms 1 - 13 describe an object called an ordered field. All of the axioms listed so far hold for the rational numbers, so the rational numbers are also an ordered field. We now introduce the final axiom, which distinguishes the real numbers from the rational numbers and which makes calculus possible.

If A is a nonempty set of real numbers, we say that $x \in \mathbb{R}$ is a **lower bound** for A if $x \leq y$ for all $y \in A$. We say that x is a **greatest lower bound** for A if x is a lower bound for A and if every lower bound z for A satisfies $z \leq x$. This is the same definition we made for our previous sets of numbers.

- Axiom 14 (completeness of $\mathbb{R}$) If a nonempty set of real numbers has a lower bound, then it has a greatest lower bound.

If A is a nonempty set of real numbers with an upper bound, we can apply Axiom 14 to the set $\{-x : x \in A\}$ to see that the set A has a least upper bound. We will use the word **infimum** to refer to the greatest lower bound and the word **supremum** to refer to the least upper bound. We will typically write $\inf A$ and $\sup A$ for the infimum and supremum of A respectively.

Axioms 1 - 14 assert that the real numbers form a complete ordered field. It is possible to prove that there is exactly one complete ordered field, in the sense that any two structures satisfying Axioms 1 - 14 are essentially the same. This uniqueness allows us to think of the set of real numbers as a single canonical object, in the same way that we think of the natural numbers.

2.1.3 Embedding $\mathbb{Q}$ in $\mathbb{R}$

There are canonical embeddings of all our previous sets of numbers into $\mathbb{R}$. Axiom 6 gave us the existence of the real number 1. We can identify the real number 1 with the natural number 1. More generally, we can identify the natural number n with the real number $1+1+\cdots+1+1$, where there are n terms in the sum.

Given the identification of $\mathbb{N}$ with a subset of $\mathbb{R}$, Axiom 3 provides us with a way to embed $\mathbb{Z}$ in $\mathbb{R}$. Specifically, for a natural number n we identify the integer $-n$ with the additive inverse of the real number n.

Finally, Axiom 7 provides us with a way to embed $\mathbb{Q}$ in $\mathbb{R}$. Specifically, we for an integer n and a nonzero integer k we identify the rational number $\frac{n}{k}$ with the real number $n \cdot k^{-1}$.

2.1.4 Examples of axiomatic proofs

In Subsection 2.1.4 we will give several examples of how to use the axioms to prove statements about the real numbers. These proofs will justify each step by explicitly citing an axiom. They will use only Axioms 1 - 13. In Subsection 2.1.5 below we will give an example of a proof using the more subtle Axiom 14.

Proposition 2.1.1 (cancellative law for addition)**.** *Suppose that x, y and z are real numbers such that $x + y = x + z$. Then $y = z$.*

Proof of Proposition 2.1.1. Let x, y and z be as in the statement of Proposition 2.1.1.

- Axiom 3 provides us with the real number $-x$. From the hypothesis $x + y = x + z$ we obtain

$$-x + x + y = -x + x + z \tag{2.1.1}$$

- By Axiom 4, we have $-x + x = x + (-x)$ so that from (2.1.1) we obtain

$$x + (-x) + y = x + (-x) + z \tag{2.1.2}$$

- By applying Axiom 3 to both sides of (2.1.2) we obtain

$$0 + y = 0 + z \tag{2.1.3}$$

- By applying Axiom 2 to both sides of (2.1.3) we obtain $y = z$ as required.

□

Proposition 2.1.2 (translation preserves inequalities)**.** *Let x, y and z be real numbers. If $x < y$ then $x + z < y + z$.*

Proof of Proposition 2.1.2. Let x, y and z be as in the statement of Proposition 2.1.2.

- By definition, the hypothesis $x < y$ means $y - x \in \mathbb{P}$.
- By Axiom 2, we have $y - x + 0 = y - x$, so that $y - x + 0 \in \mathbb{P}$.
- By Axiom 3, we have $y - x + 0 = y - x + z - z$, so that $y - x + z - z \in \mathbb{P}$.
- By Axiom 4, we have $y - x + z - z = y + z - x - z$, so that

$$y + z - x - z \in \mathbb{P} \tag{2.1.4}$$

- We claim that

$$-x - z = -(x + z) \tag{2.1.5}$$

  Indeed, by combining Axioms 2, 3 and 4 we have

$$x + z - x - z = x - x + z - z = 0 + 0 = 0 \tag{2.1.6}$$

  By Axiom 3 we have

$$x + z - (x + z) = 0 \tag{2.1.7}$$

  The claim (2.1.5) follows by applying Lemma 2.1.1 to (2.1.6) and (2.1.7).

- By combining (2.1.4) and (2.1.5) we obtain $y + z - (x + z) \in \mathbb{P}$. By definition, this means $x + z < y + z$ as required.

□

Corollary 2.1.1. *Let x be a real number and let z be a positive number. Then $x < x + z$.*

Proof of Corollary 2.1.1. Let x and z be as in the statement of Corollary 2.1.1. Since $z > 0$, Lemma 2.1.2 implies $x + z > x + 0$. By Axiom 2 we have $x + 0 = x$ so that $x + z > x$ as required. □

Theorem 2.1.2 (an ϵ of room)**.** *Suppose that x and y are real numbers such that $y \leq x + \epsilon$ for every positive number ϵ. Then $y \leq x$.*

Proof of Theorem 2.1.2. Suppose toward a contradiction that there exist real numbers x and y is as in the statement of Theorem 2.1.2, but such that the conclusion of Theorem 2.1.2 fails for x and y. By Axiom 13, we see that if $y \leq x$ fails then we have $y > x$.

- By Axioms 2, 3 and 4 we have
$$x + y - x = y \tag{2.1.8}$$
- By Axioms 6 and 9 we have
$$x + y - x = x + 1 \cdot (y - x) = x + \left(\frac{1}{2} + \frac{1}{2}\right)(y - x) = x + \frac{1}{2}(y - x) + \frac{1}{2}(y - x) \tag{2.1.9}$$
- By combining (2.1.8) and (2.1.9) we obtain
$$y = x + \frac{1}{2}(y - x) + \frac{1}{2}(y - x) \tag{2.1.10}$$
- The hypothesis that $y > x$ implies $y - x \in \mathbb{P}$. From our understanding of rational numbers, we see that $\frac{1}{2} \in \mathbb{P}$. Therefore Axiom 12 implies $\frac{1}{2}(y - x) \in \mathbb{P}$.
- By applying Corollary 2.1.1 to (2.1.10) we obtain
$$y > x + \frac{1}{2}(y - x) \tag{2.1.11}$$
- Since $\frac{1}{2}(y - x) \in \mathbb{P}$, the hypothesis that $y \leq x + \epsilon$ for all $\epsilon \in \mathbb{P}$ implies
$$y \leq x + \frac{1}{2}(y - x) \tag{2.1.12}$$
- The combination of (2.1.11) and (2.1.12) is inconsistent with the exclusivity of the trichotomy in Axiom 13. Thus we have obtain a contradiction and the proof of Theorem 2.1.2 is complete.

□

The proofs of Propositions 2.1.1 and 2.1.2 and of Theorem 2.1.2 have been written so that we can verify the arguments using the axioms in an essentially mechanical way. There is an extensive project to rewrite all mathematical literature in an analogous style so as to enable verification by computer. However, it is infeasible for human beings to think deeply about mathematics in this fashion. In the remainder of the text, we will write proofs to optimize human understanding rather than mechanical verification.

2.1.5 Existence of $\sqrt{2}$

Recall that our original motivation for developing the set of real numbers was that the set

$$\{q \in \mathbb{Q} : q \geq 0, q^2 \geq 2\}$$

had no greatest lower bound in the rational numbers. Since we are now working with real numbers, we will enlarge this set by letting

$$A = \{x \in \mathbb{R} : x \geq 0, x^2 \geq 2\}.$$

By Axiom 14, the set A has a greatest lower bound in the real numbers. It follows from the definition that there is at most one greatest lower bound for any set. Let s be the unique greatest lower bound of A.

Theorem 2.1.3. *We have* $s^2 = 2$.

We can interpret Theorem 2.1.3 as saying that Axiom 14 has provided us with the existence of the irrational real number $\sqrt{2}$.

Proof of Theorem 2.1.3. We proceed by ruling out each of the possibilities $s^2 > 2$ and $s^2 < 2$. If we do this, we will know that $s^2 - 2$ is neither positive nor negative. Then Axiom 13 will imply that $s^2 - 2 = 0$.

Case 1. In this case we assume $s^2 > 2$ and endeavor to obtain a contradiction. The way we will do this is by finding an element $x \in A$ such that $x < s$. This will contradict the hypothesis that s is a lower bound of A.

The number 1 is a lower bound for A since $x^2 \geq 2$ and $x \geq 0$ imply $x \geq 1$. Hence the assumption that s is the greatest lower bound for A implies $s \geq 1$. Therefore $2s \geq 2 > 0$.

Our hypothesis that $s^2 > 2$ implies that $s^2 - 2 > 0$. Therefore we have

$$\frac{s^2 - 2}{2s} > 0$$

Lemma 2.1.2 implies that there exists $\epsilon > 0$ such that

$$\frac{s^2 - 2}{2s} > \epsilon$$

or equivalently

$$s^2 - 2 > 2\epsilon s \tag{2.1.13}$$

Consider the number $s - \epsilon$. We claim that $s - \epsilon \in A$. Since $s - \epsilon < s$, this will violate the hypothesis that s is a lower bound for A.

In order to show $s - \epsilon \in A$ we need to show that $(s - \epsilon)^2 \geq 2$. We compute:

$$\begin{aligned} (s-\epsilon)^2 &= s^2 - 2\epsilon s + \epsilon^2 \\ &> s^2 - 2\epsilon s \qquad (2.1.14) \\ &> s^2 - (s^2 - 2) \qquad (2.1.15) \\ &= 2 \end{aligned}$$

Here in passing from (2.1.14) to (2.1.15) we used applied (2.1.13) to increase the size of the subtracted term and thereby decrease the overall size of the quantity. This completes the analysis of Case 1.

Case 2. In the case we assume that $s^2 < 2$ and again endeavor to obtain a contradiction. The way we will do this is by finding a real number $x > s$ such that x is a lower bound for A. This will contradict the hypothesis that s is the greatest lower bound of A.

Our assumption implies that $2 - s^2 > 0$. Therefore we can find a number $\epsilon > 0$ such that

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

Equivalently, we have

$$\epsilon(2s+1) < 2 - s^2. \tag{2.1.16}$$

We may assume without loss of generality that $\epsilon < 1$. We claim that $x = s + \epsilon$ is a lower bound for A. To verify this, let $y \in A$, so that we have $y^2 \geq 2$. We must show that $s + \epsilon \leq y$.

We compute:

$$\begin{aligned}
(s+\epsilon)^2 &= s^2 + 2\epsilon s + \epsilon^2 && (2.1.17)\\
&< s^2 + 2\epsilon s + \epsilon && (2.1.18)\\
&= s^2 + \epsilon(2s+1) && (2.1.19)\\
&< s^2 + 2 - s^2 && (2.1.20)\\
&= 2\\
&\leq y^2
\end{aligned}$$

Here in passing from (2.1.17) to (2.1.18) we used the assumption that $\epsilon < 1$ to conclude $\epsilon^2 < \epsilon$, and in passing from (2.1.19) to (2.1.20) we applied (2.1.16).

We have shown that $(s+\epsilon)^2 \leq y^2$. Therefore

$$0 \leq y^2 - (s+\epsilon)^2 = (y - s - \epsilon)(y + s + \epsilon).$$

Dividing by the positive number $y + s + \epsilon$ we obtain $0 \leq y - s - \epsilon$ so that $s + \epsilon \leq y$.

Since y was an arbitrary element of A, we see that $s + \epsilon$ is a lower bound for A. This contradicts the hypothesis that s was the greatest lower bound for A, completing the analysis of Case 2 and thereby completing the proof of Theorem 2.1.3. □

2.2 Uncountable sets

Recall that a set A is said to be countably infinite if there exists a bijection between A and the set $\mathbb{N}$ of natural numbers. A set is countable if it is finite or countably infinite. Intuitively, countable sets are no larger than the natural numbers. It is clear that a subset of a countable set is again countable. We saw that the rational numbers are countable, despite being a proper superset of the natural numbers.

We will say that a set is **uncountable** if it is not countable. Uncountable sets are infinite in a larger sense than countable sets - they cannot be enumerated with natural numbers. It is not obvious that uncountable sets exist. We will establish below in Theorem 2.2.1 that they do. This fact was considered very surprising when it was discovered in the late 19^{th} century. Most people have a misguided initial intuition that all infinities are equally large.

2.2.1 Uncountability of the set of binary sequences

An infinite binary sequence is a function α from $\mathbb{N}$ to the set $\{0, 1\}$. We can visualize an infinite binary sequence as a string of zeroes and ones which goes on indefinitely to the right, where the bit $\alpha(n)$ appears in the n^{th} position. We denote the set of all infinite binary sequences by $\{0, 1\}^{\mathbb{N}}$.

Theorem 2.2.1. *There exists an uncountable set. More specifically, the set $\{0,1\}^{\mathbb{N}}$ is uncountable.*

Proof of Theorem 2.2.1. Suppose toward a contradiction that $\{0,1\}^{\mathbb{N}}$ is countable. This means there exists a bijection f from $\mathbb{N}$ to $\{0,1\}^{\mathbb{N}}$. We will use the notation α_n for $f(n)$. Thus for each natural number n we have an infinite binary sequence α_n. Our hypothesis is that the list $\alpha_1, \alpha_2, \alpha_3, \ldots$ enumerates all the possible infinite binary sequences. We will obtain a contradiction by exhibiting an infinite binary sequence β which is not part of this enumerated list.

The technique for constructing β is referred to as diagonalization. Intuitively, we design β such that it differs from α_n in the n^{th} bit. More precisely, define

$$\beta(n) = \begin{cases} 0 & \text{if } \alpha_n(n) = 1 \\ 1 & \text{if } \alpha_n(n) = 0 \end{cases}$$

For every $n \in \mathbb{N}$ we have $\beta(n) \neq \alpha_n(n)$ and therefore $\beta \neq \alpha_n$. Since n was arbitrary, we conclude that β is not part of our enumerated list. This completes the proof of Theorem 2.2.1. $\square$

The proof of Theorem 2.2.1 is quite brief, but the underlying idea can be vastly generalized. An example of such a generalization is the existence of mathematical problems which cannot be completely solved by a computer program. For more information about such generalizations, research the unsolvability of the halting problem.

2.2.2 Uncountability of the real numbers

We will use Theorem 2.2.1 to establish that $\mathbb{R}$ is uncountable by exhibiting an injective function from $\{0,1\}^{\mathbb{N}}$ to $\mathbb{R}$. This means that the infinite size of the set of new quantities that we introduced when completing the rational numbers is larger than the infinite size of the rational numbers themselves.

Two lemmas on geometric sums

Before we prove that $\mathbb{R}$ is uncountable, we will establish a formula for a certain type of sum of real numbers that will used in mapping $\{0,1\}^{\mathbb{N}}$ to $\mathbb{R}$.

A **geometric progression** is a sequence $x_1, x_2, x_3, \ldots$ of positive real numbers such that the ratio $\frac{x_{n+1}}{x_n}$ of successive terms is constant. A geometric progression has the form $a, ab, ab^2, ab^3, \ldots$ for positive numbers $a = x_1$ and $b = \frac{x_{n+1}}{x_n}$.

Lemma 2.2.1 (Formula for geometric sums)**.** *For any positive number b different from 1 we have*

$$\sum_{k=0}^{n} b^k = \frac{1 - b^{n+1}}{1 - b}$$

Proof. We compute

$$(1-b)\sum_{k=0}^{n} b^k = \sum_{k=0}^{n} b^k - b\sum_{k=0}^{n} b^k$$

$$
\begin{aligned}
&= \sum_{k=0}^{n} b^k - \sum_{k=0}^{n} b^{k+1} && (2.2.1)\\
&= 1 - b^{n+1} && (2.2.2)
\end{aligned}
$$

Here, (2.2.2) follows from (2.2.1) since the terms $b, b^2, \ldots, b^{n-1}, b^n$ appear in both sums in (2.2.1) and hence are cancelled. □

The next lemma will be the key to our proof that $\mathbb{R}$ is uncountable.

Lemma 2.2.2. *For any $m, n \in \mathbb{N}$ we have*

$$\sum_{k=1}^{n} \frac{1}{10^{m+k}} \leq \frac{1}{9 \cdot 10^m}$$

Proof. We have

$$
\begin{aligned}
\sum_{k=1}^{n} \frac{1}{10^{m+k}} &= \frac{1}{10^{m+1}} \sum_{k=0}^{n-1} \frac{1}{10^k} && (2.2.3)\\
&= \frac{1}{10^{m+1}} \frac{1 - 10^{-n}}{1 - 10^{-1}} && (2.2.4)\\
&= \frac{1}{10^{m+1}} \frac{10(1 - 10^{-n})}{9} && (2.2.5)\\
&\leq \frac{1}{10^{m+1}} \frac{10}{9} && (2.2.6)\\
&= \frac{1}{9 \cdot 10^m}
\end{aligned}
$$

Here, (2.2.4) follows from (2.2.3) by applying Lemma 2.2.1 and (2.2.6) follows from (2.2.5) since $1 - 10^{-n} \leq 1$. □

Main argument

Theorem 2.2.2. *The set $\mathbb{R}$ is uncountable.*

Proof of Theorem 2.2.2. We establish Theorem 2.2.2 by constructing an injective function $f : \{0,1\}^{\mathbb{N}} \to \mathbb{R}$. Given an infinite binary sequence $\alpha \in \{0,1\}^{\mathbb{N}}$ and $n \in \mathbb{N}$, we define a rational number $R(\alpha, n)$ by setting

$$R(\alpha, n) = \sum_{k=1}^{n} \frac{\alpha(k)}{10^k}$$

Recall the notation $\sup A$ for the least upper bound of a set A of real numbers. Let $A(\alpha)$ be the set $\{R(\alpha, n) : n \in \mathbb{N}\}$ and define $f(\alpha) = \sup A(\alpha)$.

We claim that f is injective. To see this, let α and β be two distinct infinite binary sequences. We must show that $f(\alpha)$ is different from $f(\beta)$. Let $m \in \mathbb{N}$ be such that the m^{th} bit is the first place at which α differs from β. We may assume without loss of generality that $\alpha(m) = 0$ and $\beta(m) = 1$.

The number

$$x = \sum_{k=1}^{m} \frac{\beta(n)}{10^k}$$

is an element of the set $A(\beta)$. Since $f(\beta)$ is an upper bound for $A(\beta)$, we have $f(\beta) \geq x$.

Let now $j \in \mathbb{N}$ be such that $j > m$. We compute

$$\begin{aligned}
x - R(\alpha, j) &= \sum_{k=1}^{m} \frac{\beta(k)}{10^k} - \sum_{k=1}^{j} \frac{\alpha(k)}{10^k} && (2.2.7)\\
&= \frac{\beta(m)}{10^m} - \sum_{k=m}^{j} \frac{\alpha(k)}{10^k} && (2.2.8)\\
&= \frac{1}{10^m} - \sum_{k=m}^{j} \frac{\alpha(k)}{10^k} && (2.2.9)\\
&= \frac{1}{10^m} - \sum_{k=m+1}^{j} \frac{\alpha(k)}{10^k} && (2.2.10)\\
&= \frac{1}{10^m} - \sum_{k=1}^{j-m} \frac{\alpha(k+m)}{10^{k+m}} && (2.2.11)\\
&\geq \frac{1}{10^m} - \sum_{k=1}^{j-m} \frac{1}{10^{k+m}} && (2.2.12)\\
&\geq \frac{1}{10^m} - \frac{1}{9 \cdot 10^m} && (2.2.13)\\
&= \frac{8}{9 \cdot 10^m}
\end{aligned}$$

This computation can be justified as follows.

- (2.2.8) follows from (2.2.7) since we assumed that the m^{th} bit is the first place at which α differs from β and hence $\beta(k) = \alpha(k)$ for all $k \in \{1, \dots, m-1\}$
- (2.2.9) follows from (2.2.8) since we assumed $\beta(m) = 1$
- (2.2.10) follows from (2.2.9) since we assumed $\alpha(m) = 0$
- (2.2.11) follows from (2.2.10) by reindexing the sum
- (2.2.12) follows from (2.2.11) since $0 \leq \alpha(k) \leq 1$ for all k and hence

$$\sum_{k=1}^{j-m} \frac{\alpha(k+m)}{10^{k+m}} \leq \sum_{k=1}^{j-m} \frac{1}{10^{k+m}}$$

- and (2.2.13) follows from (2.2.12) by applying Lemma 2.2.2 with $n = j - m$.

The above computation shows that

$$x - \frac{8}{9 \cdot 10^m} \geq R(\alpha, j)$$

whenever $j > m$. Since $R(\alpha, \ell) \leq R(\alpha, j)$ when $\ell \leq j$ it follows that

$$x - \frac{8}{9 \cdot 10^m} \geq R(\alpha, j)$$

for all $j \in \mathbb{N}$. Therefore the number $x - \frac{8}{9 \cdot 10^m}$ is an upper bound for $A(\alpha)$.

Since $f(\beta) \geq x$, it follows that the number $f(\beta) - \frac{8}{9 \cdot 10^m}$ is an upper bound for $A(\alpha)$. Since $f(\beta) - \frac{8}{9 \cdot 10^m} < f(\beta)$ we see that $f(\beta)$ cannot be the least upper bound of $A(\alpha)$. By our definition of f, this implies that $f(\beta)$ is different from $f(\alpha)$. This completes the proof of Theorem 2.2.2 □

We emphasize that the function f was defined using least upper bounds. Since least upper bounds do not generally exist in the rational numbers the proof strategy of Theorem 2.2.2 fails when applied to the rational numbers. This is consistent with the countability of the rational numbers.

2.3 Absolute value and distance in the real line

In order to develop calculus we will need the notion of the absolute value of a real number. If $x \in \mathbb{R}$, the absolute value $|x|$ is defined through cases by

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

Here are some elementary observations about the notion of absolute value.

- $|x| \geq 0$ for all $x \in \mathbb{R}$
- $|xy| = |x||y|$
- $|x| = |-x|$
- $|x| = 0$ implies $x = 0$
- If $\epsilon > 0$, we have $|x| \leq \epsilon$ if and only if $-\epsilon \leq x \leq \epsilon$
- More generally, we have $|a - x| \leq \epsilon$ if and only if

$$a - \epsilon \leq x \leq a + \epsilon \tag{2.3.1}$$

The last property can be visualized as saying that the condition $|a - x| \leq \epsilon$ defines a symmetric interval of radius ϵ centered at x. An interval delimited by $\leq$ is called an **closed interval**. The condition $|a - x| < \epsilon$ defines a **open interval** of radius ϵ centered at a. If $a < b$ we can also define the closed interval $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ and the open interval $(a, b) = \{x \in \mathbb{R} : a < x < b\}$.

The following fact and its generalizations are the most important inequalites in all of mathematics.

Theorem 2.3.1 (triangle inequality)**.** *For all* $x, y \in \mathbb{R}$ *we have* $|x + y| \leq |x| + |y|$.

The triangle inequality gets its name because the two-dimensional version can be seen as asserting that the length of any side of a triangle is no larger than the sum of the lengths of the two other sides.

Proof of Theorem 2.3.1. First we observe that we have $-|x| \leq x \leq |x|$ and $-|y| \leq y \leq |y|$. (In each of these cases one inequality is in fact an equality.) Adding these inequalities, we obtain

$$-(|x| + |y|) \leq x + y \leq |x| + |y|.$$

We now consider two cases depending on the sign of $x + y$.

Suppose first that $x + y \geq 0$. Then $|x + y| = x + y \leq |x| + |y|$.

Suppose now that $x + y < 0$. Then $|x+y| = -(x+y)$. Since $-(|x|+|y|) \leq x+y$ we have $|x|+|y| \geq -(x+y)$. Thus we see that the theorem holds in both cases. □

We define the **distance** between two real numbers x and y to be the nonnegative quantity $|x - y|$. Here are the essential properties of the notion of distance.

- (distance is symmetric) For all x, y we have $|x - y| = |y - x|$
- (triangle inequality for distances) For all x, y, z we have

$$|x - z| \leq |x - y| + |y - z|$$

- (distance detects equality) We have $|x - y| = 0$ if and only if $x = y$.

These three properties define a structure called a metric. Metrics are ubiquitous in mathematics, as they allow one to generalize intuitions about distance in the real line.

The next property is a powerful method for proving equalities between real numbers.

Theorem 2.3.2 (an ϵ of room again)**.** *If* $|x - y| \leq \epsilon$ *for every* $\epsilon > 0$ *then* $x = y$.

Proof of Theorem 2.3.2. Theorem 2.1.2 asserted that if $z, w \in \mathbb{R}$ and $w \leq z + \epsilon$ for every positive number ϵ then $w \leq z$. If x and y are real numbers, we can apply the previous statement to $z = 0$ and $w = |x - y|$ to see that if $|x - y| \leq \epsilon$ for every positive number ϵ then $|x - y| \leq 0$. Since we always have $|x - y| \geq 0$ this implies $|x - y| = 0$. Since distance detects equality we obtain $x = y$. □

Chapter 3

Sequences and series

3.1 Sequences of real numbers

3.1.1 Basic information

We have now covered the basic theory of the real numbers themselves. The next level of structure we will consider is the idea of a function from the natural numbers to the real numbers.

A function $a : \mathbb{N} \to \mathbb{R}$ is called a sequence. We will typically write a_n instead of $a(n)$ and denote the whole sequence as $(a_n)_{n=1}^{\infty}$.

Here are some important examples of sequences.

- (constant sequence) Let $a_n = 3$ for all $n \in \mathbb{N}$
- (sequence of natural numbers) Let $a_n = n$ for all $n \in \mathbb{N}$.
- (sequence of reciprocals of natural numbers) Let $a_n = \frac{1}{n}$ for all $n \in \mathbb{N}$.

The choice of the number 3 in the first example is incidental; we could also consider a constant sequence with any other fixed value.

Since the natural numbers are infinite, it does not make sense to ask about the last element of a sequence. However, it is often possible to assert that the sequence approaches some final value as n goes to infinity. If this is the case, we say that the sequence converges.

We now present the formal definition of convergence of a sequence. This is the first aspect of our discussions which can genuinely be considered calculus.

Definition 3.1.1. *We say that a sequence* $(a_n)_{n=1}^{\infty}$ ***converges*** *to a real number* α *if for every* $\epsilon > 0$ *there exists* $N \in \mathbb{N}$ *such that for all* $n \geq N$ *we have* $|a_n - \alpha| \leq \epsilon$.

We emphasize that the number N depends on ϵ.

The intuition behind this definition is that for any fixed error width ϵ, if we wait long enough then the all remaining terms of the sequence will lie within the interval of radius ϵ around the limiting value α.

If the sequence $(a_n)_{n=1}^{\infty}$ converges to α, we denote this by writing $\lim_{n\to\infty} a_n = \alpha$.

3.1.2 Examples

We now analyze the convergence behavior of our example sequences. The first example is trivial but illustrative.

Proposition 3.1.1. *Let $a_n = 3$ for all n. We have $\lim_{n\to\infty} a_n = 3$.*

Proof of Proposition 3.1.1. Fix $\epsilon > 0$. Choose $N = 1$. For all $n \geq 1$ we have $|a_n - 3| = 0 \leq \epsilon$. □

Notice that this example would be mostly the same if we defined $a_1 = 7, a_2 = 9$ and $a_n = 3$ for all $n \geq 3$. In that case we would instead choose $N = 3$.

The next case is the essential example of a nontrivial convergent sequence.

Proposition 3.1.2. *Let $a_n = \frac{1}{n}$ for all n. We have $\lim_{n\to\infty} a_n = 0$.*

Proof of Proposition 3.1.2. Fix $\epsilon > 0$. We can consider the positive real number $\frac{1}{\epsilon}$.

We claim that there exists a natural number N such that $N \geq \frac{1}{\epsilon}$. Suppose toward a contradiction that the claim failed. Then $\frac{1}{\epsilon}$ would be an upper bound for the set of natural numbers.

By Axiom 14, there must exist a least upper bound x for $\mathbb{N}$. Since x is the least upper bound, $x - 1$ cannot be an upper bound. Therefore there must be a natural number m such that $x - 1 < m$. This implies that $m+1 > x$. Since $m+1$ is a natural number, this contradicts the hypothesis that x is an upper bound for $\mathbb{N}$.

Using the claim, we can choose a natural number N such that $N \geq \frac{1}{\epsilon}$. We need to show that for all $n \geq N$ we have $|a_n - 0| \leq \epsilon$.

Suppose $n \geq N$. Then

$$|a_n - 0| = |a_n| = \left|\frac{1}{n}\right| = \frac{1}{n} \leq \frac{1}{N} \leq \epsilon.$$

This verifies Definition 3.1.1 and completes the proof of Proposition 3.1.2 □

The next case is the essential example of a sequence which does not converge.

Proposition 3.1.3. *Let $a_n = n$ for all n. This sequence does not converge to any limit.*

Proof of Proposition 3.1.3. Suppose toward a contradiction that there exists some real number α such that $\lim_{n\to\infty} n = \alpha$.

Setting $\epsilon = 1$ in the definition of convergence, this implies that there exists N such that $|n - \alpha| \leq 1$ for all $n \geq N$. In particular, we must have $|N - \alpha| \leq 1$ so that the left side of (2.3.1) implies $N \geq \alpha - 1$.

Let $n = N + 3$. Then we have $n \geq \alpha - 1 + 3 = \alpha + 2$. This implies in particular that $n - \alpha$ is non-negative and so $|n - \alpha| = n - \alpha \geq 2$.

This contradicts our assumption that $|n - \alpha| \leq 1$ for all $n \geq N$. Since α was an arbitrary real number, the sequence $(a_n)_{n=1}^{\infty}$ does not converge to any limit. □

3.1.3 Uniqueness of limits of sequences

An important fact is that if a sequence has a limit, then that limit must be unique.

Theorem 3.1.1. *Suppose* $(a_n)_{n=1}^{\infty}$ *is a sequence of real numbers such that* $\lim_{n\to\infty} a_n = \alpha$ *and* $\lim_{n\to\infty} a_n = \beta$*. Then* $\alpha = \beta$*.*

Proof of Theorem 3.1.1. Let $(a_n)_{n=1}^{\infty}$ and α, β be as in the statement of Theorem 3.1.1.

Fix $\epsilon > 0$. Since $\lim_{n\to\infty} a_n = \alpha$, we can find N such that for all $n \geq N$ we have $|a_n - \alpha| \leq \frac{\epsilon}{2}$. Since $\lim_{n\to\infty} a_n = \beta$, we can find M such that for all $n \geq M$ we have $|a_n - \beta| \leq \frac{\epsilon}{2}$.

Choose $n \geq \max(N, M)$. Using the triangle inequality we have

$$|\alpha - \beta| \leq |\alpha - a_n| + |a_n - \beta| = |a_n - \alpha| + |a_n - \beta| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Since $\epsilon > 0$ was arbitrary, Theorem 2.3.2 implies $\alpha = \beta$ as required. □

The proof of Theorem 3.1.1 demonstrates two important analytic techniques. First of all, we see the fundamental utility of Theorem 2.3.2. Secondly, we used the idea of choosing an initial positive value of ϵ and then applying the definition of convergence to a smaller positive quantity, in this case $\frac{\epsilon}{2}$.

3.1.4 Arithmetic of sequences

We now state the basic rules that describe how limits of sequences interact with arithmetic.

Theorem 3.1.2. *Suppose* $(a_n)_{n=1}^{\infty}$ *and* $(b_n)_{n=1}^{\infty}$ *are sequences such that* $\lim_{n\to\infty} a_n = \alpha$ *and* $\lim_{n\to\infty} b_n = \beta$*.*

- *We have* $\lim_{n\to\infty}(a_n + b_n) = \alpha + \beta$*.*
- *We have* $\lim_{n\to\infty} a_n b_n = \alpha\beta$*.*

We will prove the second of these rules. The first can be proved in a similar way.

Proof of second clause in Theorem 3.1.2. Let $(a_n)_{n=1}^{\infty}, (b_n)_{n=1}^{\infty}$ and α, β be as in the statement of Theorem 3.1.2.

By applying the definition of convergence to $\epsilon = 1$, we see that there exists N such that $|a_n - \alpha| \leq 1$ for all $n \geq N$. This implies $||a_n| - |\alpha|| \leq 1$ so that from the right side of (2.3.1) we obtain $|a_n| \leq |\alpha| + 1$ for all $n \geq N$. Again applying the definition of convergence we can find N' such that

$$|a_n - \alpha| \leq \frac{\epsilon}{2(|\beta| + 1)}$$

for all $n \geq N'$. Similarly, we can find M such that

$$|b_n - \beta| \leq \frac{\epsilon}{2(|\alpha| + 1)}$$

for all $n \geq M$.

Let $K = \max(N, N', M)$ and consider $n \geq K$. We have

$$\begin{aligned}
|a_n b_n - \alpha\beta| &\leq |a_n b_n - a_n\beta| + |a_n\beta - \alpha\beta| && (3.1.1)\\
&= |a_n| \cdot |b_n - \beta| + |a_n - \alpha| \cdot |\beta| && (3.1.2)\\
&\leq (|\alpha| + 1) \cdot |b_n - \beta| + |a_n - \alpha| \cdot |\beta| && (3.1.3)\\
&\leq (|\alpha| + 1) \cdot \frac{\epsilon}{2(|\alpha| + 1)} + |a_n - \alpha| \cdot |\beta| && (3.1.4)\\
&\leq (|\alpha| + 1) \cdot \frac{\epsilon}{2(|\alpha| + 1)} + \frac{\epsilon}{2(|\beta| + 1)} \cdot |\beta| && (3.1.5)\\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

This computation can be justified as follows.

- The inequality in (3.1.1) follows from the triangle inequality (Theorem 2.3.1)
- (3.1.3) follows from (3.1.2) by our choice of N
- (3.1.4) follows from (3.1.3) by our choice of M
- and (3.1.5) follows from (3.1.4) by our choice of N'

Thus we have shown $\lim_{n\to\infty} a_n b_n = \alpha\beta$ as desired. □

In the above proof we used a more sophisticated version of the '$\frac{\epsilon}{2}$' idea by applying the definition of convergence to quantities such as $\frac{\epsilon}{2(|\beta|+1)}$.

3.1.5 Squeeze theorem

We now state a theorem which is crucial for computing many limits.

Theorem 3.1.3 (Squeeze theorem)**.** *Let $(a_n)_{n=1}^{\infty}$, $(b_n)_{n=1}^{\infty}$ and $(c_n)_{n=1}^{\infty}$ be sequences. Suppose there exists $M \in \mathbb{N}$ such that $a_n \leq b_n \leq c_n$ for all $n \geq M$. Then if $\lim_{n\to\infty} a_n = \alpha = \lim_{n\to\infty} c_n$ we also have $\lim_{n\to\infty} b_n = \alpha$.*

Proof of Theorem 3.1.3. Let $\epsilon > 0$. Since $\lim_{n\to\infty} a_n = \alpha$, we can find $N \in \mathbb{N}$ such that $|a_n - \alpha| \leq \epsilon$ for all $n \geq N$. In particular, we have $a_n \geq \alpha - \epsilon$. It follows that if $n \geq \max(M, N)$ then we have $b_n \geq a_n \geq \alpha - \epsilon$.

Since $\lim_{n\to\infty} c_n = \alpha$ we can find $N' \in \mathbb{N}$ such that $|c_n - \alpha| \leq \epsilon$ for all $n \geq N'$. In particular, we have $c_n \leq \alpha + \epsilon$. It follows that if $n \geq \max(M, N')$ then $a_n \leq c_n \leq \alpha + \epsilon$.

Therefore if $n \geq \max(M, N, N')$ then we will have $\alpha - \epsilon \leq a_n \leq \alpha + \epsilon$. By (2.3.1) this implies $|a_n - \alpha| \leq \epsilon$. □

We now consider an example of how we can apply the squeeze theorem.

Proposition 3.1.4. *Let*

$$b_n = \frac{(-1)^n + (-1)^{n^2}}{n}$$

for all n*. Then* $\lim_{n\to\infty} b_n = 0$.

Proof of Proposition 3.1.4. If we tried to work directly with the sequence $(b_n)_{n=1}^{\infty}$, this result would be difficult to prove due the oscillations of the numerator. Instead, we define sequences $a_n = -\frac{2}{n}$ and $c_n = \frac{2}{n}$. Observe that for all n we have

$$-2 \leq (-1)^n + (-1)^{n^2} \leq 2$$

so that $a_n \leq b_n \leq c_n$.

From Proposition 3.1.2 we have $\lim_{n\to\infty} \frac{1}{n} = 0$. By applying the second clause of Theorem 3.1.2 to the constant sequences with values -2 and 2 respectively, we see that

$$\lim_{n\to\infty} a_n = -2 \cdot 0 = 0 = 2 \cdot 0 = \lim_{n\to\infty} c_n$$

Thus we can apply Theorem 3.1.3 to conclude that $\lim_{n\to\infty} b_n = 0$. □

3.1.6 Monotone sequences

A special class of sequences that are particularly tractable are monotone sequences. We say that a sequence $(a_n)_{n=1}^{\infty}$ is **increasing** if $a_n \leq a_{n+1}$ for all $n \in \mathbb{N}$. Similarly, we say that $(a_n)_{n=1}^{\infty}$ is **decreasing** if $a_n \geq a_{n+1}$ for all $n \in \mathbb{N}$. We say a sequence is **monotone** if it is increasing or decreasing.

Theorem 3.1.4 (Convergence of monotone sequences)**.** *Suppose* $(a_n)_{n=1}^{\infty}$ *is an increasing sequence which has an upper bound for its terms. Then* $(a_n)_{n=1}^{\infty}$ *converges to its least upper bound. Similarly, if* $(a_n)_{n=1}^{\infty}$ *is a decreasing sequence which has a lower bound, then it converges to its greatest lower bound.*

Proof of Theorem 3.1.4. We will prove the increasing case of this theorem. The decreasing case follows by applying the increasing case to the sequence $(-a_n)_{n=1}^{\infty}$. Suppose $(a_n)_{n=1}^{\infty}$ is an increasing sequence with an upper bound. Let α be the least upper bound for the terms of $(a_n)_{n=1}^{\infty}$.

Let $\epsilon > 0$. We must find $N \in \mathbb{N}$ such that $|a_n - \alpha| \leq \epsilon$ for all $n \geq N$. Notice that since $a_n \leq \alpha$ for all n, it suffices to show that $a_n \geq \alpha - \epsilon$ for all $n \geq N$.

Since α is the least upper bound for the terms of $(a_n)_{n=1}^{\infty}$, we know that $\alpha - \epsilon$ cannot be an upper bound. Therefore there exists some a_N such that $a_N \geq \alpha - \epsilon$. Since $(a_n)_{n=1}^{\infty}$ is increasing, this implies that $a_n \geq \alpha - \epsilon$ for all $n \geq N$ as desired. □

3.1.7 Subsequences and the Bolzano-Weierstrass theorem

A subsequence of the natural numbers is an infinite list

$$n_1 < n_2 < n_3 < \cdots < n_k < n_{k+1} < \cdots$$

where each $n_k \in \mathbb{N}$. For example, we could choose $n_k = 3k$ or $n_k = 2^k$.

If $(a_n)_{n=1}^{\infty}$ is a sequence of real numbers, a subsequence of $(a_n)_{n=1}^{\infty}$ is a sequence which has the form $(a_{n_k})_{k=1}^{\infty}$ for some subsequence $(n_k)_{k=1}^{\infty}$ of the natural numbers. Intuitively, a subsequence of $(a_n)_{n=1}^{\infty}$ is obtained by forgetting some of the terms of $(a_n)_{n=1}^{\infty}$.

For example, consider the sequence $a_n = 1/n$. The sequence

$$1, \frac{1}{3}, \frac{1}{9}, \frac{1}{27}, \cdots, \frac{1}{3^k}, \cdots$$

is a subsequence of $(a_n)_{n=1}^{\infty}$. On the other hand, the sequence

$$1, \frac{1}{3}, \frac{1}{2}, \frac{1}{9}, \frac{1}{27}, \frac{1}{81}, \cdots$$

is not a subsequence of $(a_n)_{n=1}^{\infty}$. Even though each term of this latter sequence is also a term of $(a_n)_{n=1}^{\infty}$, the fact that $1/3$ and $1/2$ are in the wrong order prevents this from being a subsequence.

Lemma 3.1.1 (existence of monotone subsequences)**.** *Every sequence of real numbers has a monotone subsequence.*

Proof of Lemma 3.1.1. Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers. We need to find a monotone subsequence. The key idea is the following definition. We say that a term a_n is a **peak** of the sequence if $a_n \geq a_m$ for all $m \geq n$. In other words, a_n is at least as big as any term which follows it.

For example, every term of a decreasing sequence is a peak. On the other hand, an increasing sequence has a peak only if it is eventually constant. We will consider two cases, depending on whether the set of peaks is finite or infinite.

Case 1 (finitely many peaks) In this case we assume that the set of peaks of $(a_n)_{n=1}^{\infty}$ is finite. We will recursively build an increasing subsequence. Since the set of peaks is finite, we can find N such that if $n \geq N$ then a_n is not a peak. We choose the first index for our subsequence as $n_1 = N$.

Since $n_1 > N$, we know that a_{n_1} is not a peak. Therefore there exists $n_2 > n_1$ such that $a_{n_2} > a_{n_1}$. Since n_2 is also greater than N, we know that a_{n_2} is not a peak. Therefore there exists $n_3 > n_2$ such that $a_{n_3} > a_{n_2}$. Continuing in this way, we find an increasing subsequence.

Case 2 (infinitely many peaks) In this case we assume that the set of peaks of $(a_n)_{n=1}^{\infty}$ is infinite. We will build a decreasing subsequence.

Let $n_1 < n_2 < n_3 < \cdots$ be the sequence of indexes of peaks. We claim that the sequence $a_{n_1}, a_{n_2}, a_{n_3}, \ldots$ is decreasing. This follows from the definition of a peak: since each a_{n_k} is a peak, we have $a_{n_k} \geq a_m$ for all $m \geq n_k$. In particular, this implies $a_{n_{k+1}} \leq a_{n_k}$.

In either case, we have found a monotone subsequence. This completes the proof of Lemma 3.1.1 □

We will say that a sequence is **bounded** if there is an upper bound for the absolute values of its terms. Equivalently, there is both an upper and lower bound for the terms. The following theorem will be crucial to many later arguments.

Theorem 3.1.5 (Bolzano-Weierstrass theorem)**.** *Every bounded sequence of real numbers has a convergent subsequence.*

Proof of Theorem 3.1.5. Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers. By Lemma 3.1.1 there exists a monotone subsequence $(a_{n_k})_{k=1}^{\infty}$. Since $(a_n)_{n=1}^{\infty}$ is bounded the sequence $(a_{n_k})_{k=1}^{\infty}$ is also bounded. Therefore Theorem 3.1.5 follows from Theorem 3.1.4 $\square$

3.1.8 Cauchy sequences

In order to verify the definition of convergence directly, one needs to know in advance what the limit should be. It would useful to know a criterion for convergence which does not have this requirement. We now present such a criterion

Definition 3.1.2. *We say that a sequence $(a_n)_{n=1}^{\infty}$ is a **Cauchy sequence** if for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have $|a_n - a_m| \leq \epsilon$.*

The importance of Cauchy sequences is captured by the following theorem.

Theorem 3.1.6. *A sequence of real numbers converges if and only if it is a Cauchy sequence*

Proof of Theorem 3.1.6. The theorem asserts an equivalence, so we have two implications to prove.

Convergence implies Cauchy Suppose that $(a_n)_{n=1}^{\infty}$ is a convergent sequence with $\lim_{n\to\infty} a_n = \alpha$. Let $\epsilon > 0$. By the definition of convergence, we can find N such that if $n \geq N$ then $|a_n - \alpha| \leq \frac{\epsilon}{2}$.

In particular, if $n, m \geq N$ then we have $|a_n - \alpha| \leq \frac{\epsilon}{2}$ and $|a_m - \alpha| \leq \frac{\epsilon}{2}$ so that

$$|a_n - a_m| \leq |a_n - \alpha| + |a_m - \alpha| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus we have verified the definition of a Cauchy sequence.

Cauchy implies convergence We claim first that every Cauchy sequence is bounded. Let $(a_n)_{n=1}^{\infty}$ be a Cauchy sequence. By the definition of a Cauchy sequence, we can find $N \in \mathbb{N}$ such that if $n \geq N$ then $|a_N - a_n| \leq 1$. In particular, if $n \geq N$ then we have $|a_n| \leq |a_N| + 1$. Define $C = \max(|a_1|, |a_2|, \ldots, |a_N|) + 1$. Then $|a_n| \leq C$ for all n so $(a_n)_{n=1}^{\infty}$ is bounded.

We now claim that if a Cauchy sequence has a convergent subsequence, then the original sequence converges. Let $(a_n)_{n=1}^{\infty}$ be a Cauchy sequence and let $(a_{n_k})_{k=1}^{\infty}$ be a subsequence with $\lim_{k\to\infty} a_{n_k} = \alpha$.

Let $\epsilon > 0$. By the definition of a Cauchy sequence, we can find $N \in \mathbb{N}$ such that if $n, m \geq N$ then $|a_n - a_m| \leq \frac{\epsilon}{2}$. By the definition of convergence, we can find $K \in \mathbb{N}$ such that if $k \geq K$ then $|a_{n_k} - \alpha| \leq \frac{\epsilon}{2}$.

Since $(n_k)_{k=1}^{\infty}$ is a subsequence of the natural numbers, we can find $k \geq K$ such that $n_k \geq N$. Then for any $m \geq N$ we have

$$|a_m - \alpha| \leq |a_m - a_{n_k}| + |a_{n_k} - \alpha| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This completes the proof of the second claim.

Now, let $(a_n)_{n=1}^{\infty}$ be a Cauchy sequence. Since $(a_n)_{n=1}^{\infty}$ is bounded, Theorem 3.1.5 implies that $(a_n)_{n=1}^{\infty}$ has a convergent subsequence. Therefore the second claim implies that $(a_n)_{n=1}^{\infty}$ converges. $\square$

3.2 Infinite sums

3.2.1 Basic information

The associativity axiom for addition allows us to unambiguously define the sum of any finite list of real numbers. However, it is important in many contexts to consider infinite sums. We can give a precise definition of an infinite sum using our definition of convergence for sequences.

Definition 3.2.1. *Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers. We say that the infinite sum*

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

converges *if the sequence of finite partial sums*

$$b_k = \sum_{n=1}^{k} a_n$$

converges as $k \to \infty$.

Suppose that this is the case, and write α for $\lim_{k\to\infty} b_k$. We think of α as the exact value of the infinite sum, and we write

$$\sum_{n=1}^{\infty} a_n = \alpha$$

Often an infinite sum is called a series.

3.2.2 Zero test

Proposition 3.2.1. *A necessary condition for a series $\sum_{n=1}^{\infty} a_n$ to converge is that $\lim_{n\to\infty} a_n = 0$.*

Proof of Proposition 3.2.1. Suppose toward a contradiction that Proposition 3.2.1 fails. Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers which does not converge to zero. Negating the definition of convergence, we see that there exists some $\epsilon > 0$ such that for an infinite set $J \subseteq \mathbb{N}$ and all $k \in J$ we have $|a_k| \geq \epsilon$.

This implies that if $k \in J$ we have

$$\left|\sum_{n=1}^{k-1} a_n - \sum_{n=1}^{k} a_n\right| = |a_k| \geq \epsilon.$$

Therefore the partial sums $\sum_{n=1}^{k} a_n$ cannot form a Cauchy sequence and so Theorem 3.1.6 implies the series does not converge. □

3.2.3 Examples

Geometric series

Proposition 3.2.2. *Let $0 < b < 1$. Then we have*

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

Proof of Proposition 3.2.2. In order to prove Proposition 3.2.2, we need to examine the finite partial sums. By applying Lemma 2.2.1 we see that

$$\sum_{n=1}^{k} b^k = \frac{1-b^{k+1}}{1-b} - 1 = \frac{b-b^{k+1}}{1-b}$$

Since $b < 1$ we have $\lim_{k\to\infty} b^{k+1} = 0$, so

$$\lim_{k\to\infty} \frac{b-b^{k+1}}{1-b} = \frac{b}{1-b}$$

Thus we have shown

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

□

Harmonic series

The following example shows that the condition that $\lim_{n\to\infty} a_n = 0$ is not sufficient for the series $\sum_{n=1}^{\infty} a_n$ to converge.

Proposition 3.2.3. *The series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.*

Proof of Proposition 3.2.3. We group the terms of the series into blocks whose lengths are powers of two:

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \left(\frac{1}{2}+\frac{1}{3}\right) + \left(\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}\right) + \left(\frac{1}{8}+\cdots+\frac{1}{15}\right) + \cdots$$

Each term in the k^{th} block is at least $\frac{1}{2^k}$. Thus we see that

$$\sum_{n=1}^{\infty} \frac{1}{n} \geq 1 + \left(\frac{1}{4}+\frac{1}{4}\right) + \left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right) + \left(\frac{1}{16}+\cdots+\frac{1}{16}\right) + \cdots$$

Since there are 2^{k-1} terms in the k^{th} block, we see that the sum of each block is at least $\frac{1}{2}$ and therefore

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{n} &\geq 1 + 2\cdot\frac{1}{4} + 4\cdot\frac{1}{8} + 8\cdot\frac{1}{16} + \cdots \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots = \infty.\end{aligned}$$

□

Basel series

In the next example we will need to exploit the Cauchy criterion in order to prove convergence.

Proposition 3.2.4. *The series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges.*

For historical reasons, this is called the Basel series.

Proof of Proposition 3.2.4. Let $\epsilon > 0$ and choose $N \in \mathbb{N}$ such that $\frac{1}{N} \leq \epsilon$. We claim that if $k, m \geq N$ then

$$\left|\sum_{n=1}^{k} \frac{1}{n^2} - \sum_{n=1}^{m} \frac{1}{n^2}\right| \leq \epsilon.$$

This will show that the partial sums of the series form a Cauchy sequence, and therefore we can conclude that the series converges.

Let $k, m \geq N$ and assume without loss of generality that $m > k$. We compute:

$$\begin{aligned}
\left|\sum_{n=1}^{k} \frac{1}{n^2} - \sum_{n=1}^{m} \frac{1}{n^2}\right| &= \sum_{n=k+1}^{m} \frac{1}{n^2} \\
&\leq \sum_{n=k+1}^{m} \frac{1}{n(n-1)} \\
&= \sum_{n=k+1}^{m} \left(\frac{1}{n-1} - \frac{1}{n}\right) \\
&= \sum_{n=k+1}^{m} \frac{1}{n-1} - \sum_{n=k+1}^{m} \frac{1}{n} \\
&= \sum_{n=k}^{m-1} \frac{1}{n} - \sum_{n=k+1}^{m} \frac{1}{n} \qquad (3.2.1) \\
&= \frac{1}{k} - \frac{1}{m} \qquad (3.2.2) \\
&\leq \frac{1}{N} \leq \epsilon
\end{aligned}$$

Here in passing from (3.2.1) to (3.2.2) we cancelled a telescoping sum. □

We have shown that

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

converges, but we haven't computed the value of the sum. Determining this value was considered one of the most significant challenges in mathematics during the eighteenth century. The problem was solved by a man named Leonhard Euler, who is universally regarded as one of the greatest mathematicians of all time.

The solution is

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

3.2.4 Convergence criteria

We now develop some general techniques for analyzing the convergence behavior of infinite series.

Absolute convergence

Theorem 3.2.1. *Suppose $\sum_{n=1}^{\infty} |a_n|$ converges. Then $\sum_{n=1}^{\infty} a_n$ converges.*

Proof of Theorem 3.2.1. Assume that $\sum_{n=1}^{\infty} |a_n|$ converges. Then the partial sums $\sum_{n=1}^{k} |a_n|$ form a Cauchy sequence.

Let $\epsilon > 0$. We can find N such that if $k, m \geq N$ we have

$$\left| \sum_{n=1}^{k} |a_n| - \sum_{n=1}^{m} |a_n| \right| \leq \epsilon.$$

Assume without loss of generality that $m > k$. Using the triangle inequality we have

$$\begin{aligned} \left| \sum_{n=1}^{m} a_n - \sum_{n=1}^{k} a_n \right| = \left| \sum_{n=k+1}^{m} a_n \right| &\leq \sum_{n=k+1}^{m} |a_n| \\ &= \sum_{n=1}^{m} |a_m| - \sum_{n=1}^{k} |a_m| \\ &= \left| \sum_{n=1}^{m} |a_m| - \sum_{n=1}^{k} |a_m| \right| \leq \epsilon \end{aligned}$$

Therefore the partial sums of $\sum_{n=1}^{\infty} a_n$ form a Cauchy sequence and so the series converges. □

If $\sum_{n=1}^{\infty} |a_n|$ converges we say that $\sum_{n=1}^{\infty} a_n$ **converges absolutely**.

Comparison test

Theorem 3.2.2. *Suppose that there exists $N \in \mathbb{N}$ such that $0 \leq a_n \leq b_n$ for all $n \geq N$. If $\sum_{n=1}^{\infty} b_n$ converges then $\sum_{n=1}^{\infty} a_n$ converges. If $\sum_{n=1}^{\infty} a_n$ diverges then $\sum_{n=1}^{\infty} b_n$ diverges.*

We leave the proof of the validity of this test as an optional exercise.

The comparison test can be used in a similar way to the squeeze theorem to convert more complicated series to simpler ones. We now give an example of this.

Proposition 3.2.5. *The series*

$$\sum_{n=1}^{\infty} \frac{1}{n^2 + 2n - 5}$$

converges.

Proof of Proposition 3.2.5. If $n \geq 3$ then $2n - 5 > 0$ so that

$$\frac{1}{n^2 + 2n - 5} \leq \frac{1}{n^2}.$$

Since

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

converges, we see that

$$\sum_{n=1}^{\infty} \frac{1}{n^2 + 2n - 5}$$

converges. $\square$

Ratio test

Theorem 3.2.3. *Let $(a_n)_{n=1}^{\infty}$ be a sequence of nonzero real numbers. Suppose that there exists $N \in \mathbb{N}$ and $r < 1$ such that $|\frac{a_{n+1}}{a_n}| \leq r$ for all $n \geq N$. Then $\sum_{n=1}^{\infty} a_n$ converges. On the other hand, suppose that there exists $N \in \mathbb{N}$ such that $|\frac{a_{n+1}}{a_n}| \geq 1$ for all $n \geq N$. Then $\sum_{n=1}^{\infty} a_n$ does not converge.*

Proof of Theorem 3.2.3. Suppose first that there exists $N \in \mathbb{N}$ and $r < 1$ such that $|\frac{a_{n+1}}{a_n}| \leq r$ for all $n \geq N$.

By induction we can see that this implies $|a_n| \leq r^{n-N}|a_N|$ for all $n \geq N$. Therefore we have

$$\begin{aligned}
\sum_{n=1}^{\infty} |a_n| &= \sum_{n=1}^{N} |a_n| + \sum_{n=N+1}^{\infty} |a_n| \\
&\leq \sum_{n=1}^{N} |a_n| + \sum_{n=N+1}^{\infty} r^{n-N}|a_N| \\
&= \sum_{n=1}^{N} |a_n| + |a_N| \cdot \sum_{m=1}^{\infty} r^m
\end{aligned}$$

Since $\sum_{n=1}^{N} |a_n|$ is a finite number, it does not affect the convergence behavior of the series. The series $\sum_{n=1}^{\infty} r^m$ is a geometric series and is hence convergent by Proposition 3.2.2. Therefore $\sum_{n=1}^{\infty} a_n$ converges.

Suppose now that there exists $N \in \mathbb{N}$ such that $|\frac{a_{n+1}}{a_n}| \geq 1$ for all $n \geq N$. Equivalently, this means that $|a_{n+1}| \geq |a_n|$. Therefore $|a_n| \geq |a_N| > 0$ for all $n \geq N$. Since the individual terms of the sequence do not approach zero, we see that $\sum_{n=1}^{\infty} a_n$ cannot converge. $\square$

We now give an example of how to apply Theorem 3.2.3.

Proposition 3.2.6. *Let $a_n = \frac{n}{4^{n+3}}$. Then the series $\sum_{n=1}^{\infty} a_n$ converges.*

Proof of Proposition 3.2.6. We compute

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{n+1}{4^{n+4}} \cdot \frac{4^{n+3}}{n}\right| = \frac{n+1}{n} \cdot \frac{1}{4} = \left(1 + \frac{1}{n}\right) \cdot \frac{1}{4}.$$

Since $1 + \frac{1}{n} \leq 2$, we see that

$$\left| \frac{a_{n+1}}{a_n} \right| \leq \frac{1}{2}$$

for all n. Therefore Theorem 3.2.3 implies that

$$\sum_{n=1}^{\infty} \frac{n}{4^{n+3}}$$

converges. □

Root test

Theorem 3.2.4. *Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers. Suppose that there exists $N \in \mathbb{N}$ and $r < 1$ such that $|a_n| \leq r^n$ for all $n \geq N$. Then $\sum_{n=1}^{\infty} a_n$ converges.*

Theorem 3.2.4 can be established in a similar way to Theorem 3.2.3, using comparison with a geometric series.

We now consider an example of how to apply Theorem 3.2.4.

Proposition 3.2.7. *Let $a_n = 3^{11n} n^{-n}$. Then the $\sum_{n=1}^{\infty} a_n$ converges.*

Proof of Proposition 3.2.7. Suppose $n \geq 3^{11} + 1$. Then we have

$$a_n = \left(\frac{3^{11}}{n} \right)^n \leq \left(\frac{3^{11}}{3^{11} + 1} \right)^n .$$

Letting $r = \frac{3^{11}}{3^{11}+1} < 1$ we can apply the root test to conclude that

$$\sum_{n=1}^{\infty} \frac{3^{11n}}{n^n}$$

converges. □

Chapter 4

Limits and continuity

4.1 Functions of a real variable

We will now begin to discuss the main subject of this text: functions from a subset of the real numbers to the real numbers.

If $A \subseteq \mathbb{R}$ and f is a function from A to $\mathbb{R}$, we refer to A as the domain of the function f. We think of A as being the set of real numbers for which the function f is defined. In many cases, A will be the whole set of real numbers. In this case we say that the function is defined everywhere.

4.1.1 Examples

We now give some basic examples of classes of functions.

Polynomials

Polynomials are functions which have the form

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1} + c_n x^n$$

where $c_0, \ldots, c_n$ are fixed real numbers. A polynomial is defined everywhere. Below are some explicit examples of polynomials.

- $f(x) = x$
- $f(x) = 7x^5 - 3x - 19$
- $f(x) = \displaystyle\sum_{k=0}^{153} \frac{(-1)^k}{k^{13}+1} x^{4k}$

The highest power of the variable which appears in a polynomial is called the degree of the polynomial. The degrees of the examples above are $1, 5$ and $612 = 153 \cdot 4$ respectively.

Constant functions are considered polynomials of degree zero.

Rational functions

Rational functions have the form $f(x) = \frac{p(x)}{q(x)}$ where p and q are polynomials. Such a rational function is defined for all x with $q(x) \neq 0$.

Taking q to be the constant function with value 1, we see that polynomials are a special case of rational functions. Below are some explicit examples of rational functions which are not polynomials.

- $f(x) = \dfrac{1}{x}$
- $f(x) = \dfrac{7x^3 - 11x + 4}{(x-3)(x+8)}$
- $f(x) = \dfrac{-19x^4}{x^2+1}$

The first example is defined for all nonzero x. The second example is defined for all x except 3 and -8. The third example is defined everywhere, since the quantity $x^2 + 1$ is always positive.

Absolute value

We can consider the function $f(x) = |x|$. More generally, if a is a fixed real number, we can consider the function $f(x) = |x - a|$. We think of this function as expressing the distance from x to the point a.

Indicator functions

If $B \subseteq \mathbb{R}$, we can define a function $f : \mathbb{R} \to \mathbb{R}$ by setting

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

This function is called the indicator of B. As the name suggests, indicator functions give a way of identifying subsets of the real numbers.

4.1.2 Operations on functions

There are some basic operations we can perform to make new functions out of old ones.

Arithmetic operations

If $f : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$, we can define the sum $f + g$ by $(f + g)(x) = f(x) + g(x)$. Similarly, we can define the product fg by $(fg)(x) = f(x)g(x)$. We can also define the quotient $\frac{f}{g}$ by $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$ for inputs x with $g(x) \neq 0$.

All rational functions can be built from constant functions and the function $f(x) = x$ using these operations.

Compositions

Suppose $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ are functions such that $f(A) \subseteq B$. Then we can define the composition $g \circ f : A \to \mathbb{R}$ by $(g \circ f)(x) = g(f(x))$.

For example, let $f(x) = x^3 - 3$ and $g(x) = x^2 + x + 7$. Then we have

$$\begin{aligned}(g \circ f)(x) &= g(f(x)) \\ &= g(x^3 - 3) \\ &= (x^3 - 3)^2 + (x^3 - 3) + 7 \\ &= x^6 - 5x^3 + 13\end{aligned}$$

Piecewise definitions

Suppose $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ are functions such that $f(x) = g(x)$ whenever $x \in A \cap B$. Then we can define a piecewise function $h : A \cup B \to \mathbb{R}$ by letting

$$h(x) = \begin{cases} f(x) & \text{if } x \in A \\ g(x) & \text{if } x \in B \end{cases}$$

In particular, if A and B are disjoint, then we can make a piecewise definition of a function on $A \cup B$ for any pair of functions $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$.

More generally, if $A_1, \ldots, A_n$ are pairwise disjoint sets and $f_k : A_k \to \mathbb{R}$ are functions, then we can make a piecewise definition of a function $g : A_1 \cup \cdots \cup A_n \to \mathbb{R}$ by letting $g(x) = f_k(x)$ for $x \in A_k$.

4.2 Limits

4.2.1 Basic information

Definition 4.2.1. *Let $A \subseteq \mathbb{R}$. We say that a point $a \in \mathbb{R}$ is a **cluster point** of A if for every $\delta > 0$ there exists $x \in A$ with $|x - a| \leq \delta$.*

Intuitively, a point a is a cluster point of a set A if there exist points in A which are arbitrarily close to a.

Definition 4.2.2. *Let $f : A \to \mathbb{R}$ be a function and let $a \in \mathbb{R}$ be a cluster point of A. We say that the **limit** of f as x approaches a is equal to α if the the following condition is satisfied. For every $\epsilon > 0$ there exists $\delta > 0$ such that if $x \in A$ and $|x - a| \leq \delta$ then $|f(x) - \alpha| \leq \epsilon$. If this is the case, we write $\lim_{x \to a} f(x) = \alpha$.*

We emphasize that the number δ depends on ϵ.

The hypothesis that a is a cluster point of A is necessary in order to rule out the possibility that the conditions $|x - a| \leq \delta$ and $x \in A$ are incompatible for some positive value of δ. When we discuss limits, we will always implicitly assume that the limit point is a cluster point of the domain.

4.2.2 Uniqueness of limits

Theorem 4.2.1. *Let $f : A \to \mathbb{R}$ be a function and suppose $\lim_{x\to a} f(x) = \alpha$ and $\lim_{x\to a} f(x) = \beta$. Then $\alpha = \beta$.*

Proof of Theorem 4.2.1. Let $\epsilon > 0$. Applying the definition of the limit to α, we see that there exists $\delta > 0$ such that if $x \in A$ and $|x - a| \leq \delta$ then $|f(x) - \alpha| \leq \frac{\epsilon}{2}$.

Applying the definition of the limit to β, we see that there exists $\delta' > 0$ such that if $x \in A$ and $|x - a| \leq \delta'$ then $|f(x) - \beta| \leq \frac{\epsilon}{2}$.

Let $\delta'' = \min(\delta, \delta') > 0$. Since a is a cluster point of A, we can choose $x \in A$ such that $|x - a| \leq \delta''$. We have

$$|\alpha - \beta| \leq |\alpha - f(x)| + |f(x) - \beta| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

4.2.3 Examples

We now give some examples of limits of functions.

Identity function

The first example is trivial but illustrative.

Proposition 4.2.1. *Let $f(x) = x$. Then for any point $a \in \mathbb{R}$ we have $\lim_{x\to a} f(x) = a$.*

Proof of Proposition 4.2.1. Let $\epsilon > 0$ and choose $\delta = \epsilon$. Then $|x - a| \leq \delta$ implies $|f(x) - a| = |x - a| \leq \epsilon$. □

Cube function

Proposition 4.2.2. *Let $f(x) = x^3$. Then $\lim_{x\to 2} f(x) = 8$.*

Proof of Proposition 4.2.2. Let $\epsilon > 0$. We need to find $\delta > 0$ such that if $|x - 2| \leq \delta$ then $|x^3 - 8| \leq \epsilon$.

Set $\delta = \min\left(\frac{\epsilon}{19}, 1\right)$ and suppose $|x - 2| \leq \delta$. In particular, this implies $|x - 2| \leq 1$ so that $|x| \leq 3$. We have

$$\begin{aligned}
|x^3 - 8| &= |(x-2)(x^2 + 2x + 4)| \\
&= |x - 2| \cdot |x^2 + 2x + 4| \\
&\leq |x - 2| \cdot (|x|^2 + 2|x| + 4) \\
&\leq |x - 2| \cdot (3^2 + 2 \cdot 3 + 4) \\
&= 19 \cdot |x - 2| \leq 19\delta \leq \epsilon
\end{aligned}$$

□

A rational function

Proposition 4.2.3. *Let*

$$f(x) = \frac{x^3 - 3x^2 - 4x + 12}{x - 3}$$

Then $\lim_{x \to 3} f(x) = 5$.

The domain of the function f in Proposition 4.2.3 is $A = \{x \in \mathbb{R} : x \neq 3\}$. We want to evaluate $\lim_{x \to 3} f(x)$. Since the limit point 3 is a cluster point of A the limit is well-defined.

Proof of Proposition 4.2.3. Let $\epsilon > 0$. We need to find $\delta > 0$ such that if $|x - 3| \leq \delta$ then

$$\left| \frac{x^3 - 3x^2 - 4x + 12}{x - 3} - 5 \right| \leq \epsilon.$$

We set $\delta = \min\left(\frac{\epsilon}{7}, 1\right)$. This implies in particular that if $|x - 3| \leq \delta$ then $|x - 3| \leq 1$ and hence $|x| \leq 4$. Suppose $x \in A$ and $|x - 3| \leq \delta$. We compute

$$\begin{aligned} \left| \frac{x^3 - 3x^2 - 4x + 12}{x - 3} - 5 \right| &= \left| \frac{(x-3)(x^2-4)}{x - 3} - 5 \right| && (4.2.1) \\ &= |x^2 - 9| = |x - 3| \cdot |x + 3| && (4.2.2) \\ &\leq |x - 3| \cdot (|x| + 3) \\ &\leq 7 \cdot |x - 3| \leq 7\delta \leq \epsilon. \end{aligned}$$

Here our cancellation of the $x - 3$ factors in passing from (4.2.1) to (4.2.2) is justified since the assumption that $x \in A$ implies $x - 3 \neq 0$. □

Sign function

We now consider an example where a limit does not exist.

Proposition 4.2.4. *Define*

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

Then $\lim_{x \to 0} f(x)$ *does not exist.*

The function f in Proposition 4.2.4 is sometimes called the sign function.

Proof of Proposition 4.2.4. Suppose toward a contradiction that there were $\alpha \in \mathbb{R}$ such that $\lim_{x \to 0} f(x) = \alpha$. Applying the definition of the limit with $\epsilon = \frac{1}{2}$, we see that there exists $\delta > 0$ such that if $|x| \leq \delta$ then $|f(x) - \alpha| \leq \frac{1}{2}$.
In particular, this implies that

$$|f(\delta) - f(-\delta)| \leq |f(\delta) - \alpha| + |\alpha - f(-\delta)| \leq \frac{1}{2} + \frac{1}{2} = 1. \tag{4.2.3}$$

On the other hand, we have $|f(\delta) - f(-\delta)| = |1 - (-1)| = 2$, which contradicts (4.2.3). □

Note that the fact $f(x) = 0$ played no role in the proof of Proposition 4.2.4. We could define $f(0)$ to be any real number and the limit would still fail to exist. We could also leave f undefined at 0. This illustrates an important idea: limits depend on the behavior of the function nearby the limit point, and not on the value at the limit point itself.

4.2.4 Manipulation of limits

The rules for arithmetically manipulating limits of functions are identical to the rules for manipulating limits of sequences. The proofs are similar.

Theorem 4.2.2. *Suppose $f : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$ are functions, suppose a is a cluster point of A, suppose $\lim_{x\to a} f(x) = \alpha$ and suppose $\lim_{x\to b} g(x) = \beta$.*

- *We have $\lim_{x\to a}(f + g)(x) = \alpha + \beta$.*
- *We have $\lim_{x\to a}(fg)(x) = \alpha\beta$.*
- *If $\beta \neq 0$ we have $\lim_{x\to a} \left(\frac{f}{g}\right)(x) = \frac{\alpha}{\beta}$.*

Proof of Theorem 4.2.2. We prove the second of these rules; the proof of the first and third are similar. Let $\epsilon > 0$. Since $\lim_{x\to a} f(x) = \alpha$ we see there exists $\delta > 0$ such that if $|x - a| \leq \delta$ then $|f(x) - \alpha| \leq 1$ and so $|f(x)| \leq |\alpha| + 1$. Again using the hypothesis that $\lim_{x\to a} f(x) = \alpha$ we see that there exists $\delta' > 0$ such that if $|x - a| \leq \delta'$ then

$$|f(x) - \alpha| \leq \frac{\epsilon}{2(|\beta| + 1)}$$

Finally, using the hypothesis that $\lim_{x\to a} g(x) = \beta$ we see that there exists $\gamma > 0$ such that $|x - a| \leq \gamma|$ then

$$|g(x) - \beta| \leq \frac{\epsilon}{2(|\alpha| + 1)}$$

Suppose $|x - a| \leq \min(\delta, \delta', \gamma)$. We compute

$$\begin{aligned}
|f(x)g(x) - \alpha\beta| &\leq |f(x)g(x) - f(x)\beta| + |f(x)\beta - \alpha\beta| && (4.2.4)\\
&= |f(x)| \cdot |g(x) - \beta| + |f(x) - \alpha|\beta && (4.2.5)\\
&\leq (|\alpha| + 1)|g(x) - \beta| + |f(x) - \alpha|\beta && (4.2.6)\\
&\leq (|\alpha| + 1)\frac{\epsilon}{2(|\alpha| + 1)} + |f(x) - \alpha|\beta && (4.2.7)\\
&\leq (|\alpha| + 1)\frac{\epsilon}{2(|\alpha| + 1)} + \frac{\epsilon\beta}{2(|\beta| + 1)} && (4.2.8)\\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

This computation can be justified as follows.

- The inequality in (4.2.4) follows from the triangle inequality (Theorem 2.3.1)
- (4.2.6) follows from (4.2.5) by our choice of δ
- (4.2.7) follows from (4.2.6) by our choice of δ'
- and (4.2.8) follows from (4.2.7) by our choice of γ.

□

4.2.5 Squeeze theorem

We also have an analog of the squeeze theorem.

Theorem 4.2.3. *Suppose $f : A \to \mathbb{R}, g : A \to \mathbb{R}$ and $h : A \to \mathbb{R}$ are functions. Further suppose that there exists $\gamma > 0$ such that if $|x - a| \leq \gamma$ then $f(x) \leq g(x) \leq h(x)$. If $\lim_{x\to a} f(x) = \alpha$ and $\lim_{x\to a} h(x) = \alpha$ then we also have $\lim_{x\to a} g(x) = \alpha$.*

Proof of Theorem 4.2.3. Let $\epsilon > 0$. Since $\lim_{x\to a} f(x) = \alpha$ we see there exists $\delta > 0$ such that if $|x - a| \leq \delta$ then $|f(x) - \alpha| \leq \epsilon$. In particular, this implies $f(x) \geq \alpha - \epsilon$. Therefore if $|x - a| \leq \min(\gamma, \delta)$ we will have $g(x) \geq f(x) \geq \alpha - \epsilon$.

Since $\lim_{x\to a} h(x) = \beta$ we see there exists $\delta' > 0$ such that if $|x - a| \leq \delta'$ then $|h(x) - \alpha| \leq \epsilon$. In particular, this implies $h(x) \leq \alpha + \epsilon$. Therefore if $|x - a| \leq \min(\gamma, \delta')$ we will have $f(x) \leq h(x) \leq \alpha + \epsilon$.

It follows that if $|x - a| \leq \min(\gamma, \delta, \delta')$ we will have $\alpha - \epsilon \leq f(x) \leq \alpha + \epsilon$. By (2.3.1) we see this implies $|f(x) - \alpha| \leq \epsilon$. □

4.2.6 Extended limits

We now discuss some important extensions of the limit concept.

Limits to infinity

Definition 4.2.3. *Let $f : A \to \mathbb{R}$ be a function. We say that $\lim_{x\to a} f(x) = \infty$ if for every $M < \infty$ there exists $\delta > 0$ such that if $|x - a| \leq \delta$ then $f(x) \geq M$.*

Similarly, we say that $\lim_{x\to a} f(x) = -\infty$ if for every $M > -\infty$ there exists $\delta > 0$ such that if $|x - a| \leq \delta$ then $f(x) \leq M$.

We now consider an example of a limit to infinity.

Proposition 4.2.5. *Let $f(x) = \frac{1}{|x|}$. Then $\lim_{x\to 0} f(x) = \infty$.*

Proof of Proposition 4.2.5. Let $M < \infty$. We need to find $\delta > 0$ such that if $|x| \leq \delta$ and $x \neq 0$ then $\frac{1}{|x|} \geq M$.

We can choose $\delta = \frac{1}{M}$. Then if $|x| \leq \delta$ and $x \neq 0$ we have $\frac{1}{|x|} \geq \frac{1}{\delta} \geq M$. We can show with a parallel argument that $\lim_{x\to 0} -\frac{1}{|x|} = -\infty$. □

Limits at infinity

If $c \in \mathbb{R}$ we define the right infinite ray $[c, \infty) = \{x \in \mathbb{R} : x \geq c\}$. We also define the left infinite ray $(-\infty, c] = \{x \in \mathbb{R} : x \leq c\}$.

Definition 4.2.4. *Suppose $c \in \mathbb{R}$ and let $f : [c, \infty) \to \mathbb{R}$ be a function. We say that $\lim_{x\to\infty} f(x) = \alpha$ if for every $\epsilon > 0$ there exists $M \geq c$ such that if $x \geq M$ then $|f(x) - \alpha| \leq \epsilon$.*

Similarly, if $f : (-\infty, c] \to \mathbb{R}$ is a function, we say that $\lim_{x\to-\infty} f(x) = \alpha$ if for every $\epsilon > 0$ there exists $M \leq c$ such that if $x \leq M$ then $|f(x) - \alpha| \leq \epsilon$.

We now consider an example of a limit at infinity.

Proposition 4.2.6. *Let $f(x) : [1,\infty) \to \mathbb{R}$ be given by $f(x) = \frac{1}{x}$. We have* $\lim_{x\to\infty} f(x) = 0$.

Proof of Proposition 4.2.6. Let $\epsilon > 0$. We need to find $M \geq 1$ such that if $x \geq M$ then $|\frac{1}{x}| \leq \epsilon$. Assume without loss of generality that $\epsilon < 1$.

We can choose $M = \frac{1}{\epsilon}$. Then if $x \geq M$ we have $\frac{1}{x} \leq \frac{1}{M} \leq \epsilon$. □

We can show with a parallel argument that $\lim_{x\to-\infty} f(x) = 0$.

Infinite limits at infinity

Definition 4.2.5. *Let $f : [c,\infty) \to \mathbb{R}$ be a function. We say that* $\lim_{x\to\infty} f(x) = \infty$ *if for every $M < \infty$ there exists $K \geq c$ such that if $x \geq K$ then $f(x) \geq M$. We can make analogous definitions involving negative infinity.*

We now consider an example of an infinite limit at infinity.

Proposition 4.2.7. *Let $f(x) = x^2 - 10^{100}x$. We have* $\lim_{x\to\infty} f(x) = \infty$.

Intuitively, Proposition 4.2.7 because once x is sufficiently large the extra power in the x^2 term outweighs the huge negative coefficient in the x term.

Proof of Proposition 4.2.7. Let $M < \infty$. We need to find $K \in \mathbb{R}$ such that it $x \geq K$ then $f(x) \geq M$. Choose $K = \max(M, 10^{100} + 1)$ and suppose $x \geq K$. We have

$$f(x) = x(x - 10^{100}) \geq M(10^{100} + 1 - 10^{100}) = M$$

□

4.2.7 One-sided limits

Definition 4.2.6. *Let $f : A \to \mathbb{R}$ be a function. We say that the limit of f as x approaches a* ***from above*** *is equal to α if for every $\epsilon > 0$ there exists $\delta > 0$ so that if $a \leq x \leq a + \delta$ then $|f(x) - \alpha| \leq \epsilon$. If this is the case we write* $\lim_{x\downarrow a} f(x) = \alpha$.

We say that the limit of f as x approaches a ***from below*** *is equal to α if for every $\epsilon > 0$ there exists $\delta > 0$ so that if $a - \delta \leq x \leq a$ then $|f(x) - \alpha| \leq \epsilon$. If this is the case we write* $\lim_{x\uparrow a} f(x) = \alpha$.

Proposition 4.2.8. *Let f be the sign function defined in Proposition 4.2.4. Then $\lim_{x\downarrow 0} f(x) = 1$ and* $\lim_{x\uparrow 0} f(x) = -1$.

Proof of Proposition 4.2.8. Let $\epsilon > 0$ and choose $\delta = 1$. Then if $0 \leq x \leq \delta$ we have $|f(x) - 1| = |1 - 1| = 0 \leq \epsilon$. The proof of the second claim is similar. □

4.3 Continuity

4.3.1 Basic information

In general, functions of a real variable can exhibit extremely strange behavior and it is nearly impossible to prove nontrivial theorems about arbitrary functions. In order to avoid these pathological cases and have a chance to say interesting things, we will need to restrict our attention to special classes of 'nice' functions.

The most fundamental of these is the class of continuous functions.

Definition 4.3.1. *Let $f : A \to \mathbb{R}$ be a function and let a be a cluster point of A such that $a \in A$. We will say that f is **continuous** at a if $\lim_{x \to a} f(x) = f(a)$.*

Writing this out in terms of the definition of a limit, we see that f is continuous at a if for every $\epsilon > 0$ there exists $\delta > 0$ such that if $x \in A$ and $|x - a| \le \delta$ then $|f(x) - f(a)| \le \epsilon$.

If $B \subseteq A$, we will say that f is continuous on B if f is continuous at every point of B. If we assert that f is continuous without further qualification, we mean that f is continuous on its entire domain. Intuitively, a function is continuous on a set B if making a small change to an input which lies in B results in a small change in the corresponding output.

Proposition 4.3.1. *Assume that f and g are continuous at a. Then $f + g$ and fg are continuous at a. If $g(a) \neq 0$ then $\frac{f}{g}$ is continuous at a.*

Proof of Proposition 4.3.1. By Theorem 4.2.2 we have

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

and

$$\lim_{x \to a} (fg)(x) = f(a)g(a) = (fg)(a)$$

and if $g(a) \neq 0$ then

$$\lim_{x \to a} \left(\frac{f}{g}\right)(x) = \frac{f(a)}{g(a)} = \left(\frac{f}{g}\right)(a)$$

□

4.3.2 Examples

Polynomials

Proposition 4.3.2. *Every polynomial is continuous everywhere. Every rational function is continuous on its domain.*

Proof of Proposition 4.3.2. By Proposition 4.2.1 we see that if $f(x) = x$ then $\lim_{x \to a} f(x) = a$ for every $a \in \mathbb{R}$. Since $f(a) = a$, we can interpret this as saying that f is a continuous function. It is straightforward to see that constant functions are also continuous.

Since every polynomial can be built up from constant functions and the function $f(x) = x$ using sums and products, the first two clauses of Proposition 4.3.1 imply that every polynomial is continuous. The third clause in Proposition 4.3.1 implies that rational functions are continuous wherever they are defined. □

In particular, we see from Proposition 4.3.2 that $\lim_{x\to 2} x^3 = 2^3 = 8$. We proved this already using the definition of a limit, but that proof involved some algebraic tricks. This illustrates how developing some theory can make mathematics much more efficient.

Piecewise definition

Proposition 4.3.3. *Let $f : [0, 1] \to \mathbb{R}$ be given by*

$$f(x) = \begin{cases} 5x & \text{if } 0 \le x \le \frac{1}{2} \\ 5 - 5x & \text{if } \frac{1}{2} < x \le 1 \end{cases}$$

Then f is continuous on its whole domain.

Proof of Proposition 4.3.3. If $a \neq \frac{1}{2}$, then f is equal to a polynomial near the point a and is therefore continuous at a. Therefore we only need to check for continuity at the point $\frac{1}{2}$.

Let $\epsilon > 0$. We have $f(\frac{1}{2}) = \frac{5}{2}$. Thus we need to find $\delta > 0$ such that if $|x - \frac{1}{2}| \le \delta$ then $|f(x) - \frac{5}{2}| \le \epsilon$. Let $\delta = \frac{\epsilon}{5}$ and suppose $|x - \frac{1}{2}| \le \delta$. Because f has a piecewise definition, we will need to consider two cases.

First suppose that $x \le \frac{1}{2}$. Then

$$\left|f(x) - \frac{5}{2}\right| = \left|5x - \frac{5}{2}\right| = 5 \cdot \left|x - \frac{1}{2}\right| \le 5\delta = \epsilon.$$

Now suppose $x > \frac{1}{2}$. Then

$$\left|f(x) - \frac{5}{2}\right| = \left|5 - 5x - \frac{5}{2}\right| = \left|\frac{5}{2} - 5x\right| = 5 \cdot \left|\frac{1}{2} - x\right| \le 5\delta = \epsilon.$$

Since both cases have been verified, we see that f is continuous at the point $\frac{1}{2}$ and is therefore continuous on its whole domain. □

Single discontinuity

Proposition 4.3.4. *Define*

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases}$$

Then f is not continuous at the point 0.

Proof of Proposition 4.3.4. Suppose toward a contradiction that f were continuous at 0. Letting $\epsilon = \frac{1}{2}$, we see that there must exist $\delta > 0$ such that if $|x| \le \delta$ then $|f(0) - f(x)| = |1 - f(x)| \le \frac{1}{2}$.

On the other hand, we have $|f(0) - f(\delta)| = |1 - 0| = 1$, which is the desired contradiction. □

This behavior of this function at 0 is representative of typical points of discontinuity: if a function is discontinuous at a point a then its graph will have some sort of jump at a.

Discontinuity everywhere

Proposition 4.3.5. *Define*

$$f(x) = \begin{cases} 1 & \textit{if } x \textit{ is rational} \\ 0 & \textit{if } x \textit{ is irrational} \end{cases}$$

Then f is nowhere continuous.

Proof of Proposition 4.3.5. Let a be an arbitrary real number and suppose toward a contradiction that f were continuous at a. Letting $\epsilon = \frac{1}{2}$, we see that there must exist $\delta > 0$ such that if $|x - a| \leq \delta$ then $|f(a) - f(x)| \leq \frac{1}{2}$.

We consider two cases, depending on whether or not a is rational. Suppose first that a is irrational. Recall that for any real numbers c and d with $c < d$ there exists a rational number x such that $c \leq x \leq d$. In particular, we can find a rational number x such that $|a - x| \leq \delta$. Then we have $|f(a) - f(x)| = |0 - 1| = 1$ and thus we have obtained a contradiction.

Suppose now that a is rational. We claim that for any real numbers c and d with $c < d$ there exists an irrational number x such that $c \leq x \leq d$. To see this, choose a nonzero rational number q such that $\frac{c}{\sqrt{2}} \leq q \leq \frac{d}{\sqrt{2}}$. Then $c \leq \sqrt{2}q \leq d$. Since an irrational number multiplied by an nonzero rational number must be irrational, we have verified the claim.

In particular we can choose an irrational number x such that $|x - a| \leq \delta$. Then we have $|f(a) - f(x)| = |1 - 0| = 1$ and so we again have a contradiction. ☐

4.3.3 Continuity and composition

Theorem 4.3.1. *Suppose $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ are functions with $f(A) \subseteq B$. If f is continuous at a point a and g is continuous at $f(a)$ then $g \circ f$ is continuous at a.*

In particular, if both functions are continuous on their whole domains then the composition is continuous on its whole domain.

Proof of Theorem 4.3.1. Suppose that f is continuous at a and that g is continuous at $f(a)$. Let $\epsilon > 0$. We need to find $\delta > 0$ such that if $|x - a| \leq \delta$ then $|(g \circ f)(x) - (g \circ f)(a)| \leq \epsilon$.

Applying the definition of continuity to g at the point $f(a)$, we see that there exists $\delta > 0$ such that if $|y - f(a)| \leq \delta'$ then $|g(y) - g(f(a))| \leq \epsilon$. Applying the definition of continuity to f at the point a, we see that there exists $\delta' > 0$ such that if $|x - a| \leq \delta'$ then $|f(x) - f(a)| \leq \delta$.

Suppose that $|x - a| \leq \delta'$ Then $|f(x) - f(a)| \leq \delta$, so that $f(x)$ qualifies as a point y with $|y - f(a)| \leq \delta'$. Therefore $|g(f(x)) - g(f(a))| \leq \epsilon$ as desired. ☐

4.3.4 Continuity and sequences

We can reformulate the notion of continuity in terms of compatibility with taking limits of sequences.

Theorem 4.3.2. *Let $f : A \to \mathbb{R}$ be a function. Then f is continuous at a if and only if* $\lim_{n\to\infty} f(a_n) = f(a)$ *for any sequence $(a_n)_{n=1}^{\infty}$ in A such that* $\lim_{n\to\infty} a_n = a$.

Proof of Theorem 4.3.2. As usual, when proving an equivalence we treat each implication separately.

Continuity implies compatibility with sequences Suppose that f is continuous at a point a and let $(a_n)_{n=1}^{\infty}$ be a sequence in the domain of f such that $\lim_{n\to\infty} a_n = a$. Let $\epsilon > 0$. We need to find $N \in \mathbb{N}$ such that if $n \geq N$ then $|f(a_n) - f(a)| \leq \epsilon$.

Since f is continuous at a, we can find $\delta > 0$ such that if $|x - a| \leq \delta$ then $|f(x) - f(a)| \leq \epsilon$. Since $\lim_{n\to\infty} a_n = a$, we can find $N \in \mathbb{N}$ such that if $n \geq N$ then $|a_n - a| \leq \delta$. Putting these two conditions together, we see that if $n \geq N$ then $|f(a_n) - f(a)| \leq \epsilon$ as desired.

Compatibility with sequences implies continuity We prove this implication by establishing the contrapositive. Suppose that f is not continuous at a. We will exhibit a sequence $(a_n)_{n=1}^{\infty}$ which converges to a such that the sequence $(f(a_n))_{n=1}^{\infty}$ does not converge to $f(a)$.

Negating the definition of continuity, we see that there exists some $\epsilon > 0$ such that for every $\delta > 0$ there exists $x \in A$ such that $|x - a| \leq \delta$ but $|f(a) - f(x)| > \epsilon$. Fix $n \in \mathbb{N}$. Applying the previous criterion with $\delta = \frac{1}{n}$ we obtain a number x such that $|x - a| \leq \frac{1}{n}$ but $|f(a) - f(x)| > \epsilon$.

Making explicit the dependence of x on n, we write $x = a_n$. Thus we have constructed a sequence $(a_n)_{n=1}^{\infty}$. Since $|a_n - a| \leq \frac{1}{n}$ we have $\lim_{n\to\infty} |a_n - a| \leq \lim_{n\to\infty} \frac{1}{n} = 0$. It is clear that $\lim_{n\to\infty} |a_n - a| = 0$ if and only $\lim_{n\to\infty} a_n = a$.

On the other hand, by hypothesis $|f(a_n) - f(a)| > \epsilon$ for all n. Therefore $(f(a_n))_{n=1}^{\infty}$ does not converge to $f(a)$. □

4.3.5 Intermediate value theorem

The following theorem that illustrates the power of assuming continuity.

Theorem 4.3.3 (Intermediate value theorem)**.** *Suppose $f : A \to \mathbb{R}$ is a continuous function and $[c, d]$ is an interval contained in A such that $f(c) \leq f(d)$. Then for any α with $f(c) \leq \alpha \leq f(d)$ there exists $t \in [c, d]$ such that $f(t) = \alpha$. If $f(c) \geq f(d)$ then we have the same statement for any α such that $f(c) \geq \alpha \geq f(d)$.*

This theorem gives precise meaning to the intuition that a continuous function cannot make 'jumps'.

Proof of Theorem 4.3.3. We claim that it suffices to prove the theorem in the case that $\alpha = 0$. Indeed, the case of the general theorem when $f(c) \leq f(d)$ follows by applying the special case to the function $f - \alpha$. Then the case of the general theorem when $f(c) \geq f(d)$ follows by applying the previous case to $-f$. Thus we focus on proving the theorem in the case $\alpha = 0$.

Let f and $[c, d]$ be as in the hypotheses of Theorem 4.3.3 and assume $f(c) \leq 0 \leq f(d)$. Let $S = \{x \in [c, d] : f(x) \leq 0\}$. Since we have assumed that $f(c) \leq 0$, we know that S is nonempty. Let $t = \sup S$. We claim that $f(t) = 0$.

Let $\epsilon > 0$. Since f is continuous, there exists $\delta > 0$ such that if $|x - t| \leq \delta$ then $|f(x) - f(t)| \leq \epsilon$. Since $t = \sup S$ we see that $t - \delta$ cannot be an upper bound for S and so there exists $x \in S$ so $x > t - \delta$. Since t is an upper bound for S we have $t \geq x$ and so $|x - t| \leq \delta$.

This implies that $|f(x) - f(t)| \leq \epsilon$ and since $f(x) \leq 0$ we see that $f(t) \leq \epsilon$. Thus in order to prove the theorem it suffices to show that $f(t) \geq -\epsilon$.

Since $S \subseteq [c, d]$ we see that d is an upper bound for S and therefore $t \leq d$. We now consider two cases, depending on whether or not $t = d$. Suppose first that $t = d$. Then $f(t) = f(d) \geq 0$ and so certainly we have $f(t) \geq -\epsilon$.

Now, suppose that $t < d$. Then there exists $x \in [c, d]$ such that $t < x \leq t + \delta$. Since t is an upper bound for S, we must have $x \notin S$. Since $|t - x| \leq \delta$ we see that $|f(x) - f(t)| \leq \epsilon$. The fact that $x \notin S$ implies $f(x) > 0$ and so we see that $f(t) \geq -\epsilon$ as desired. □

We emphasize that the proof of Theorem 4.3.3 involved taking the supremum of S and hence ultimately relies on Axiom 14. The analog of Theorem 4.3.3 fails for the rational numbers. The validity of Theorem 4.3.3 for real numbers is one of the primary motivations for passing from the rational numbers to the real numbers.

The intermediate value theorem has numerous interesting consequences. We will give two examples.

Theorem 4.3.4. *Every polynomial of odd degree is surjective.*

Proof of Theorem 4.3.4. Let

$$p(x) = c_0 + c_1 x + c_2 x_2 + \cdots + c_{n-1} x^{n-1} + c_n x^n$$

be a polynomial and assume that n is odd. Since p is surjective if and only if $-p$ is surjective, we may assume without loss of generality that $c_n > 0$.

We claim that $\lim_{x\to\infty} p(x) = \infty$. Since this statement only depends on large positive values of x, we can ignore the case $x = 0$. Thus we can rewrite

$$p(x) = x^n \left(\frac{c_0}{x^n} + \frac{c_1}{x^{n-1}} + \frac{c_2}{x^{n-2}} + \cdots + \frac{c_{n-1}}{x} + c_n \right).$$

Then we have

$$\begin{aligned} \lim_{x\to\infty} p(x) &= \left(\lim_{x\to\infty} x^n \right) \cdot \lim_{x\to\infty} \left(\frac{c_0}{x^n} + \frac{c_1}{x^{n-1}} + \frac{c_2}{x^{n-2}} + \cdots + \frac{c_{n-1}}{x} + c_n \right) \\ &= c_n \cdot \lim_{x\to\infty} x^n = \infty \end{aligned}$$

Similarly, we can see that

$$\begin{aligned} \lim_{x\to-\infty} p(x) &= \left(\lim_{x\to-\infty} x^n \right) \cdot \lim_{x\to-\infty} \left(\frac{c_0}{x^n} + \frac{c_1}{x^{n-1}} + \frac{c_2}{x^{n-2}} + \cdots + \frac{c_{n-1}}{x} + c_n \right) \\ &= c_n \cdot \lim_{x\to-\infty} x^n = -\infty \end{aligned}$$

Now, let α be an arbitrary real number. Since $\lim_{x\to\infty} p(x) = \infty$, we can find $d > 0$ such that $p(d) \geq \alpha$. On the other hand, since $\lim_{x\to-\infty} p(x) = -\infty$, we can find $c < 0$ such that $p(c) \leq \alpha$.

Applying Theorem 4.3.3 we see that there must exist $t \in [c, d]$ such that $p(t) = \alpha$. □

The next example is not mathematical and should not be taken too seriously.

Statement 4.3.1. *At any time there exist two antipodal locations on the earth's equator which have the same temperature.*

It is reasonable to assume that the temperature at a fixed instant in time is a continuous function of location. Let $T(\ell)$ be the temperature at a location ℓ.

Given a location ℓ on the equator, let $\tilde{\ell}$ be its antipode. Fix any location ℓ_0 on the equator. If $x \geq 0$, define ℓ_x to be the location obtained by traveling x miles west along the equator from ℓ_0.

Define a continuous function f on the real numbers by letting $f(x) = T(\ell_x) - T(\widetilde{\ell_x})$. If c is half the circumference of the earth, then $\widetilde{\ell_c} = \ell_0$ and similarly $\widetilde{\ell_0} = \ell_c$. Therefore

$$\begin{aligned} f(c) &= T(\ell_c) - T(\widetilde{\ell_c}) = T(\ell_c) - T(\ell_0) \\ &= -(T(\ell_0) - T(\ell_c)) = -(T(\ell_0) - T(\widetilde{\ell_0})) = -f(0). \end{aligned}$$

By the intermediate value theorem there must exist $x \in [0, c]$ such that $f(x) = 0$. For such an x, the location ℓ_x has the same temperature as $\widetilde{\ell_x}$.

4.3.6 Continuity and compactness

We now discuss two more important theorems about continuous functions. In addition to continuity, these theorems crucially exploit the properties of closed intervals of finite length. A closed interval of finite length is referred to as a compact interval.

Boundedness theorem

We say that a function $f : A \to \mathbb{R}$ is **bounded** if there exists a positive constant $M < \infty$ such that $|f(x)| \leq M$ for all $x \in A$.

Theorem 4.3.5. *Let $f : [c, d] \to \mathbb{R}$ be a continuous function. Then f is bounded.*

Before we prove this theorem, let us examine the necessity of its hypotheses.

- (Closure is necessary) The function $f(x) = \frac{1}{x}$ is continuous on the open interval $(0, 1)$ but is not bounded.
- (Finite length is necessary) The function $f(x) = x$ is continuous on the right infinite ray $[0, \infty)$ but is not bounded.
- (Continuity is necessary) The discontinuous function

$$f(x) = \begin{cases} \dfrac{1}{x} & \text{if } 0 < x \leq 1 \\ 0 & \text{if } x = 0 \end{cases}$$

is defined on the closed interval $[0, 1]$ but is not bounded.

In order to prove Theorem 4.3.5, we will need the following lemma.

Lemma 4.3.1. *Suppose $[c, d]$ is a closed interval and $(a_n)_{n=1}^{\infty}$ is a sequence in $[c, d]$ such that $\lim_{n\to\infty} a_n = \alpha$. Then $\alpha \in [c, d]$.*

Proof of Lemma 4.3.1. Suppose toward a contradiction that the lemma failed, so that $\alpha \notin [c, d]$.

Assume that $\alpha > d$. Applying the definition of convergence with $\epsilon = \frac{\alpha-d}{2}$, we see that there must exist $N \in \mathbb{N}$ such that for all $n \geq N$ we have $|a_n - \alpha| \leq \frac{\alpha-d}{2}$. In particular, if $n \geq N$ we have

$$a_n \geq \alpha - \frac{\alpha - d}{2} = \frac{\alpha + d}{2} > d,$$

contradicting the hypothesis that $a_n \in [c, d]$.

If $\alpha < c$ we can obtain a contradiction by a parallel argument with $\epsilon = \frac{c-\alpha}{2}$. Thus we have established the lemma. ☐

Proof of Theorem 4.3.5. Suppose toward a contradiction that $f : [c, d] \to \mathbb{R}$ were continuous and unbounded. Then for every $n \in \mathbb{N}$ there must exist a point $x \in [c, d]$ such that $|f(x)| \geq n$.

Making explicit the dependence of x on n, write a_n for this x. Since $a_n \in [c, d]$ for every n, the sequence $(a_n)_{n=1}^{\infty}$ is bounded. Therefore we can apply the Bolzano-Weierstrass theorem (Theorem 3.1.5) to find a subsequence $(a_{n_k})_{k=1}^{\infty}$ and a number $\alpha \in \mathbb{R}$ such that $\lim_{k\to\infty} a_{n_k} = \alpha$. By Lemma 4.3.1, we must have $\alpha \in [c, d]$.

By Theorem 4.3.2 we must have $\lim_{k\to\infty} f(a_{n_k}) = f(\alpha)$. Applying the definition of convergence with $\epsilon = 1$, we see that there must exist $K \in \mathbb{N}$ such that for all $k \geq K$ we have $|f(a_{n_k}) - f(\alpha)| \leq 1$ and therefore $|f(a_{n_k})| \leq |f(\alpha)| + 1$.

Since $(n_k)_{k=1}^{\infty}$ is a subsequence of the natural numbers, we can find $k \geq K$ such that $n_k \geq |f(\alpha)| + 2$. By construction, $|f(a_{n_k})| \geq n_k$, so we see that $|f(a_{n_k})| \geq |f(\alpha)| + 2$ and have obtained a contradiction. ☐

Extreme value theorem

Definition 4.3.2. *Let $f : A \to \mathbb{R}$ be a function. We say that f attains its **maximum value** at a point a if $f(a) = \sup_{x\in A} f(x)$. We say that f attains its minimum value at a point a if $f(a) = \inf_{x\in A} f(x)$.*

Theorem 4.3.6 (Extreme value theorem)**.** *Let $f : [c, d] \to \mathbb{R}$ be a continuous function. Then f attains its maximum and minimum values at points in $[c, d]$.*

Before we prove Theorem 4.3.6, we examine the necessity of its hypotheses.

- (Closure is necessary) The function $f(x) = x$ is continuous on the open interval $(0, 1)$. We have $\sup_{0<x<1} f(x) = 1$ and $\inf_{0<x<1} f(x) = 0$, but these values are not attained by f.
- (Finite length is necessary) The function $f(x) = \frac{x}{1+|x|}$ is continuous on the whole real line. We have $\sup_{x\in\mathbb{R}} f(x) = 1$ and $\inf_{x\in\mathbb{R}} f(x) = -1$ but f never attains these values.

- (Continuity is necessary) The discontinuous function $f : [0, 1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x & \text{if } 0 < x < 1 \\ \frac{1}{2} & \text{if } x = 0 \text{ or } x = 1 \end{cases}$$

  satisfies $\sup_{0 \leq x \leq 1} f(x) = 1$ and $\inf_{0 \leq x \leq 1} f(x) = 0$ but f never attains these values.

Proof of Theorem 4.3.6. Let $f : [c, d] \to \mathbb{R}$ be a continuous function. We will prove that f attains its maximum. Assuming we have done this, we can see that f attains its minimum by applying the maximum case to $-f$.

Let $S = \{f(x) : c \leq x \leq d\}$ be the range of f. We need to find $\alpha \in [c, d]$ such that $f(\alpha) = \sup S$. If $n \in \mathbb{N}$ then $-\frac{1}{n} + \sup S$ cannot be an upper bound for S and hence there exists $x \in S$ with $\sup S - \frac{1}{n} \leq x$. By the definition of S, we see that there exists $y \in [c, d]$ such that $-\frac{1}{n} + \sup S \leq f(y)$. Making explicit the dependence of this y on n, write $y = a_n$.

Thus we have defined a sequence $(a_n)_{n=1}^{\infty}$ in $[c, d]$. By the Bolzano-Weierstrass theorem (Theorem 3.1.5) we can find a subsequence $(a_{n_k})_{k=1}^{\infty}$ and a number α such that $\lim_{k\to\infty} a_{n_k} = \alpha$. By Lemma 4.3.1, we see that $\alpha \in [c, d]$. We claim that $f(\alpha) = \sup S$.

By Theorem 4.3.2 we see that $\lim_{k\to\infty} f(a_{n_k}) = f(\alpha)$. By construction, we have

$$f(a_{n_k}) \geq \sup S - \frac{1}{n_k}$$

Since $f(a_{n_k}) \in S$, automatically we have $f(a_{n_k}) \leq \sup S$. Applying the squeeze theorem we see that

$$\sup S \geq \lim_{k\to\infty} f(a_{n_k}) \geq \sup S - \lim_{k\to\infty} \frac{1}{n_k} = \sup S.$$

Here we have used the fact that $(n_k)_{k=1}^{\infty}$ is a subsequence of the natural numbers to conclude that $\lim_{k\to\infty} \frac{1}{n_k} = 0$. So we see $f(\alpha) = \sup S$ as desired. □

4.3.7 Uniform continuity

Basic information

We now introduce a stronger notion of continuity.

Definition 4.3.3. *A function $f : A \to \mathbb{R}$ is* ***uniformly continuous*** *if for every $\epsilon > 0$ there exists $\delta > 0$ such that for every $a \in A$ we have that $|x - a| \leq \delta$ implies $|f(x) - f(a)| \leq \epsilon$.*

In order to make a comparison, we write out our original definition of continuity.

Definition 4.3.4. *A function $f : A \to \mathbb{R}$ is* ***continuous*** *if for every $a \in A$ and every $\epsilon > 0$ there exists $\delta > 0$ such that if $|x - a| \leq \delta$ then $|f(x) - f(a)| \leq \epsilon$.*

The difference between these definitions is that for the usual notion of continuity, the number δ is allowed to depend on both ϵ and a. In the definition of uniform continuity, we require that for every ϵ we can choose a value of δ which works simultaneously for every point $a \in A$.

In other words, we must allow for the possibility that a depends on δ. In many contexts, uniformly continuous functions are easier to work with than general continuous functions. We will see instances of this phenomenon later in the course.

Examples

Proposition 4.3.6. *Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = 100x - 3$.Then f is uniformly continuous.*

Proof of Proposition 4.3.6. Let $\epsilon > 0$. In order to show that f is uniformly continuous, we need to find $\delta > 0$ such that for every $a \in \mathbb{R}$ we have that $|x - a| \leq \delta$ implies that $|f(x) - f(a)| \leq \epsilon$. We can choose $\delta = \frac{\epsilon}{100}$. With this choice, for any $a \in \mathbb{R}$, if $|x - a| \leq \delta$ then we have

$$|f(x) - f(a)| = |100x - 3 - 100a + 3| = 100|x - a| \leq 100\delta = \epsilon.$$

□

The previous example can be generalized considerably. We say that a function $f : A \to \mathbb{R}$ is **Lipschitz continuous** if there exists a positive constant $C < \infty$ such that $|f(x) - f(y)| \leq C|x - y|$ for all $x, y \in A$. If this is the case, we call C the Lipschitz constant of f.

Every Lipschitz continuous function f is uniformly continuous. Indeed, if C is the Lipschitz constant of f then we can always choose $\delta = \frac{\epsilon}{C}$ to verify the definition of uniform continuity.

Proposition 4.3.7. *The function $f : [0, \infty) \to \mathbb{R}$ given by $f(x) = x^2$ is continuous but not uniformly continuous.*

Proof of Proposition 4.3.7. Proposition 4.3.2 implies that f is continuous. Suppose toward a contradiction that f were uniformly continuous. Letting $\epsilon = 1$ we could find $\delta > 0$ such that for all $a \in [0, \infty)$ we have that $|x - a| \leq \delta$ implies $|f(x) - f(a)| \leq 1$.

Set $a = \frac{1}{\delta}$ and $x = \frac{1}{\delta} + \delta$. Then certainly we have $|x - a| \leq \delta$. On the other hand,

$$\begin{aligned} |f(x) - f(a)| &= \left| \left(\frac{1}{\delta}\right)^2 - \left(\frac{1}{\delta} + \delta\right)^2 \right| \\ &= \left| \frac{1}{\delta} - \left(\frac{1}{\delta} + \delta\right) \right| \cdot \left| \frac{1}{\delta} + \frac{1}{\delta} + \delta \right| \\ &= \delta\left(\frac{2}{\delta} + \delta\right) \geq 2. \end{aligned}$$

□

Note that the proof of Proposition 4.3.7 relied on the dependence of a on δ.

Proposition 4.3.8. *Define $f : (0, 1) \to \mathbb{R}$ by $f(x) = \frac{1}{x}$. Then f is continuous but not uniformly continuous.*

Proof. Proposition 4.3.2 implies that f is continuous. Suppose toward a contradiction that f were uniformly continuous. Let $\epsilon = 1$ we could we could find $\delta > 0$ such that for all $a \in (0, 1)$ we have that $|x - a| \leq \delta$ implies $|f(x) - f(a)| \leq 1$.

We distinguish two cases, depending on whether or not $\delta \leq \frac{1}{4}$. If $\delta > \frac{1}{4}$ then setting $a = \frac{1}{2}$ and $x = \frac{1}{4}$ we have $|x - a| = \frac{1}{4} \leq \delta$ but

$$|f(x) - f(a)| = |2 - 4| = 2.$$

On the other hand, if $\delta \leq \frac{1}{4}$ then setting $a = \delta$ and $x = 2\delta$ we have

$$|f(x) - f(a)| = \left|\frac{1}{\delta} - \frac{1}{2\delta}\right| = \frac{1}{2\delta} \geq 2.$$

In either case we have obtained a contradiction. □

Again, the proof of Proposition 4.3.8 relied on the dependence of a on δ.

Uniform continuity and compactness

Theorem 4.3.7. *If $f : [c, d] \to \mathbb{R}$ is a continuous function then f is uniformly continuous.*

Proof of Theorem 4.3.7. Suppose toward a contradiction that Theorem 4.3.7 fails. Negating the definition of uniform continuity, we see that there must exist $\epsilon > 0$ such that for every $\delta > 0$ there exists a point $a \in [c, d]$ and a point $x \in [c, d]$ with $|a - x| \leq \delta$ but $|f(a) - f(x)| \geq \epsilon$. We fix such an ϵ for the remainder of the proof.

Setting $\delta = \frac{1}{n}$ we see that there exist $a, x \in [c, d]$ with $|a - x| \leq \frac{1}{n}$ but $|f(a) - f(x)| \geq \epsilon$. Making explicit the dependence on n, we write a_n and x_n for these points. Thus for each $n \in \mathbb{N}$ we have $|a_n - x_n| \leq \frac{1}{n}$ but $|f(a_n) - f(x_n)| \geq \epsilon$.

Applying the Bolzano-Weierstrass theorem (Theorem 3.1.5), we can find a subsequence $(a_{n_k})_{k=1}^{\infty}$ and a number $\alpha \in \mathbb{R}$ with $\lim_{k\to\infty} a_{n_k} = \alpha$. By Lemma 4.3.1, we must have $\alpha \in [c, d]$. We must also have $\lim_{k\to\infty} x_{n_k} = \alpha$.

By Theorem 4.3.2 we obtain

$$\begin{aligned}
0 &= |f(\alpha) - f(\alpha)| \\
&= \left| f\left(\lim_{k\to\infty} a_{n_k}\right) - f\left(\lim_{k\to\infty} x_{n_k}\right)\right| \\
&= \left|\left(\lim_{k\to\infty} f(a_{n_k})\right) - \left(\lim_{k\to\infty} f(x_{n_k})\right)\right| \\
&= \left|\lim_{k\to\infty} \left(f(a_{n_k}) - f(x_{n_k})\right)\right| && (4.3.1) \\
&= \lim_{k\to\infty} |f(a_{n_k}) - f(x_{n_k})| && (4.3.2) \\
&\geq \epsilon,
\end{aligned}$$

which contradicts the hypothesis that $\epsilon > 0$. Here in passing from (4.3.1) to (4.3.2) we used the fact that the absolute value function is continuous. □

Chapter 5

Differentiation

5.1 Fundamentals

5.1.1 Basic information

Definition 5.1.1. *Let $f : [c, d] \to \mathbb{R}$ be a function. We define f to be **differentiable** at a point $a \in [c, d]$ if the limit*

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \tag{5.1.1}$$

*exists. If this is the case, we write $f'(a)$ for the value of the limit and say that $f'(a)$ is the **derivative** of f at the point a. If we say that a function is differentiable without further qualification, we mean that it is differentiable on its whole domain.*

Sometimes the limit in (5.1.1) is written as

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

We can see that these are equivalent by substituting h for $x - a$.

The intuition behind Definition 5.1.1 as follows. If $x \neq a$, then the slope of the line passing through $f(x)$ and $f(a)$ is $\frac{f(x)-f(a)}{x-a}$ and the equation of the line is

$$\ell(y) = \frac{f(x) - f(a)}{x - a}(y - a) + f(a)$$

If x is close to a, we think of this line as being an approximation to the graph of f at a.

If this slope has a limiting value as x approaches a, then the limit is the slope of a tangent line to the graph of f at the point a. The equation of this tangent is

$$\ell(y) = f'(a)(y - a) + f(a).$$

We think of this as the 'best possible' linear approximation to f at a.

It is useful to think of the derivative as a function of the same variable as the original function. Thus we will frequently use the notation $f'(x)$ for the derivative of a function $f(x)$.

5.1.2 Differentiation and continuity

Theorem 5.1.1. *Suppose $f : [c, d] \to \mathbb{R}$ is differentiable at a point a. Then f is continuous at a.*

Proof of Theorem 5.1.1. Suppose f is differentiable at a. Then we have

$$\begin{aligned}\left(\lim_{x\to a} f(x)\right) - f(a) &= \lim_{x\to a}(f(x) - f(a)) \\ &= \lim_{x\to a} \frac{f(x) - f(a)}{x - a}(x - a) \qquad (5.1.2) \\ &= \left(\lim_{x\to a} \frac{f(x) - f(a)}{x - a}\right)\left(\lim_{x\to a}(x - a)\right) \qquad (5.1.3) \\ &= f'(a)\left(\lim_{x\to a}(x - a)\right) = 0\end{aligned}$$

Here, in passing from (5.1.2) to (5.1.3) we used the hypothesis that f is differentiable allows us to apply Theorem 4.2.2. □

5.1.3 Examples

Proposition 5.1.1. *Let $f(x) = sx + t$ for fixed $s, t \in \mathbb{R}$. Then f is differentiable and $f'(x) = s$ for all x.*

Proof. For any $a \in \mathbb{F}$ we compute

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} = \lim_{x\to a} \frac{sx + t - sa - t}{x - a} = \lim_{x\to a} s = s$$

□

We can see that if f is as in Proposition 5.1.1 then the equation of the tangent to f is

$$\begin{aligned}\ell(y) = f'(a)(y - a) + f(a) &= s(y - a) + sa + t \\ &= sy + t = f(y).\end{aligned}$$

This can be interpreted as asserting that if f is already a linear function then the best linear approximation to f is f itself.

Proposition 5.1.2. *Let $f(x) = x^2$. Then f is differentiable and $f'(x) = 2x$ for all x.*

Proof of Proposition 5.1.2. We compute

$$\begin{aligned}\lim_{x\to a} \frac{f(x) - f(a)}{x - a} &= \lim_{x\to a} \frac{x^2 - a^2}{x - a} \\ &= \lim_{x\to a} \frac{(x - a)(x + a)}{x - a} \\ &= \lim_{x\to a}(x + a) = 2a.\end{aligned}$$

□

Theorem 5.1.1 implies that any discontinuous function is not differentiable. However, the following example shows that are continuous functions which are not differentiable.

Proposition 5.1.3. *Let $f(x) = |x|$. Then f is not differentiable at 0.*

Proof of Proposition 5.1.3. For $x \neq 0$ we have

$$\frac{f(x) - f(0)}{x - 0} = \frac{|x|}{x} = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

Proposition 4.2.4 shows this limit does not exist as $x \to 0$. □

5.1.4 Rules for taking derivatives

Arithmetic rules

Theorem 5.1.2. *Suppose f and g are functions which are differentiable at a point a.*

- *(constant rule) If $\alpha \in \mathbb{R}$ then $(\alpha f)'(a) = \alpha f'(a)$.*
- *(sum rule) We have $(f + g)'(a) = f'(a) + g'(a)$.*
- *(product rule) We have $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.*
- *(quotient rule) If $g(a) \neq 0$ then*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Implicit in these statements in Theorem 5.1.2 is that the assumptions imply all of the functions above are differentiable at a.

Proof of Theorem 5.1.2. We will prove the product rule and the quotient rule, as the proofs of the first two rules are straightforward.

Proof of product rule Write $h = fg$. We compute

$$\begin{aligned}
\lim_{x\to a} \frac{h(x) - h(a)}{x - a} &= \lim_{x\to a} \frac{f(x)g(x) - f(a)g(a)}{x - a} \\
&= \lim_{x\to a} \frac{f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)}{x - a} \\
&= \left(\lim_{x\to a} \frac{f(x)g(x) - f(a)g(x)}{x - a}\right) + \left(\lim_{x\to a} \frac{f(a)g(x) - f(a)g(a)}{x - a}\right) \\
&= \left(\lim_{x\to a} \frac{f(x) - f(a)}{x - a} g(x)\right) + f(a)\left(\lim_{x\to a} \frac{g(x) - g(a)}{x - a}\right) \\
&= \left(\lim_{x\to a} \frac{f(x) - f(a)}{x - a}\right)\left(\lim_{x\to a} g(x)\right) + f(a)g'(a) && (5.1.4) \\
&= f'(a)g(a) + f(a)g'(a) && (5.1.5)
\end{aligned}$$

Here in passing from (5.1.4) to (5.1.5) we used the fact that since g is differentiable at a it must be continuous at a and therefore $\lim_{x\to a} g(x) = g(a)$.

Proof of quotient rule Now, suppose $h = \frac{f}{g}$ and assume that $g(a) \neq 0$. Since g is continuous at a, this implies that for x sufficiently close to a we have $g(x) \neq 0$. Therefore we are to free to divide by $g(x)$ in our calculation of limits at a. We have

$$\begin{aligned}
\lim_{x\to a} \frac{h(x) - h(a)}{x - a} &= \lim_{x\to a} \frac{1}{x-a}\left(\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}\right) \\
&= \lim_{x\to a} \frac{f(x)g(a) - f(a)g(x)}{g(x)g(a)(x-a)} \\
&= \lim_{x\to a} \frac{f(x)g(a) - f(a)g(a) + f(a)g(a) - f(a)g(x)}{g(x)g(a)(x-a)} \\
&= \lim_{x\to a} \frac{1}{g(x)g(a)}\left(g(a)\frac{f(x)-f(a)}{x-a} + f(a)\frac{g(a)-g(x)}{x-a}\right) \\
&= \left(\lim_{x\to a} \frac{1}{g(x)g(a)}\right)\left(g(a)\lim_{x\to a}\frac{f(x)-f(a)}{x-a} - f(a)\lim_{x\to a}\frac{g(x)-g(a)}{x-a}\right) \qquad (5.1.6) \\
&= \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2} \qquad (5.1.7)
\end{aligned}$$

Here in passing from (5.1.6) to (5.1.7) we used the fact that g is continuous at a to conclude that

$$\lim_{x\to a} \frac{1}{g(x)g(a)} = \frac{1}{g(a)^2}$$

□

Derivatives of polynomials

Using the product rule, we can compute the derivative of any polynomial. We first compute the derivatives of monomials.

Proposition 5.1.4. *Let $n \in \mathbb{N}$ and let $m_n(x) = x^n$. Then $m_n'(x) = nx^{n-1}$.*

Proof of Proposition 5.1.4. We establish Proposition 5.1.4 by induction on n. Let $P(n)$ be the statement that $m_n'(x) = nx^{n-1}$. We have seen in Proposition 5.1.1 that $P(1)$ is true.

Suppose we know that $P(n)$ is true. We compute

$$\begin{aligned}
m_{n+1}'(x) = (m_n m_1)'(x) &= m_n'(x)m_1(x) + m_1'(x)m_n(x) \\
&= nx^{n-1} \cdot x + x^n = (n+1)x^n.
\end{aligned}$$

Therefore $P(n+1)$ holds and by induction we see that $P(n)$ holds for all n. □

Combining Proposition 5.1.4 with the sum and constant rules we see that if $p(x) = \sum_{n=0}^{k} c_n x^n$ then

$$p'(x) = \sum_{n=0}^{k} nc_n x^{n-1} = \sum_{n=0}^{k-1} (n+1)c_{n+1}x^n.$$

Chain rule

We now state the rule for how differentiation interacts with composition of functions.

Theorem 5.1.3 (Chain rule)**.** *Let $f : [c, d] \to \mathbb{R}$ and $g : [s, t] \to \mathbb{R}$ be functions such that $f(x) \in [s, t]$ for all $x \in [c, d]$. Suppose that f is differentiable at $a \in [c, d]$ and g is differentiable at $f(a)$. Then $g \circ f$ is differentiable at a and $(g \circ f)'(a) = g'(f(a))f'(a)$.*

The proof of Theorem 5.1.3 will rely on the following lemma.

Lemma 5.1.1. *A function $f : [c, d] \to \mathbb{R}$ is differentiable at $a \in [c, d]$ if and only if there exists a function $\phi : [c, d] \to \mathbb{R}$ which is continuous at a such that $f(x) - f(a) = \phi(x)(x - a)$ for all $x \in [c, d]$. In this case, we have $f'(a) = \phi(a)$.*

Proof of Lemma 5.1.1. Suppose first that f is differentiable at a. Then we can define

$$\phi(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{if } x \neq a \\ f'(a) & \text{if } x = a \end{cases}.$$

The continuity of ϕ at a follows from the differentiability of f.

Suppose now that there exists a function ϕ as in the statement of the lemma. For $x \neq a$ we can divide by $x - a$ to obtain

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \phi(x) = \phi(a)$$

since ϕ is continuous at a. □

We now establish Theorem 5.1.3.

Proof of Theorem 5.1.3. Since f is differentiable at a, there exists a function $\phi : [c, d] \to \mathbb{R}$ which is continuous at a with $\phi(a) = f'(a)$ and such that $f(x) - f(a) = \phi(x)(x - a)$. Since g is differentiable at $f(a)$, there exists a function $\psi : [s, t] \to \mathbb{R}$ which is continuous at $f(a)$ with $\psi(f(a)) = g'(a)$ and such that $g(x) - g(a) = \psi(x)(x - a)$.

We compute

$$g(f(x)) - g(f(a)) = \psi(f(x))(f(x) - f(a)) = \psi(f(x))\phi(x)(x - a).$$

If $x \neq a$ we can divide by $x - a$ to obtain

$$\begin{aligned} \lim_{x \to a} \frac{g(f(x)) - g(f(a))}{x - a} &= \lim_{x \to a} \psi(f(x))\phi(x) \\ &= \psi(f(a))\phi(a) = g'(f(a))f'(a). \end{aligned}$$

□

5.2 Essential theorems about derivatives

In Section 5.2 we establish three essential theorems about derivatives. Each of these theorems will build on the previous one.

5.2.1 Interior extremum theorem

Definition 5.2.1. *If $f : [c, d] \to \mathbb{R}$ is a function, we say that f has a* ***relative maximum*** *at a point $a \in [c, d]$ if there exists $\epsilon > 0$ such that $|x - a| \le \epsilon$ implies $f(a) \ge f(x)$. We say that f has a* ***relative minimum*** *at a if the previous statement holds with the final inequality reversed. We say that f has a* ***relative extremum*** *at a if it has a relative maximum or a relative minimum at a.*

Theorem 5.2.1. *Suppose $f : [c, d] \to \mathbb{R}$ is differentiable and that f has a relative extremum at a point a such that $c < a < d$. Then $f'(a) = 0$.*

Proof of Theorem 5.2.1. We prove the case when f has a relative maximum at a. The case when f has a relative minumum at a can be established by applying the maximum case to $-f$.

Assume f has a relative maximum at $a \in (c, d)$ and let $\epsilon > 0$ be such that $|x - a| \le \epsilon$ implies $f(a) \ge f(x)$. Suppose toward a contradiction that $f'(a) > 0$. Then there exists $\delta > 0$ such that $|x - a| \le \delta$ implies

$$\left| \frac{f(x) - f(a)}{x - a} - f'(a) \right| \le \frac{f'(a)}{2}.$$

In particular, we have

$$\frac{f(x) - f(a)}{x - a} \ge \frac{f'(a)}{2}.$$

Since $a \in (c, d)$ there exists $x \in (c, d)$ with

$$a < x < \min(d, a + \epsilon, a + \delta).$$

Then

$$f(x) - f(a) \ge \frac{f'(a)(x - a)}{2} > 0,$$

contradicting the hypothesis that f has a relative maximum at a.

The possibility that $f'(a) < 0$ can be excluded using a parallel argument. □

5.2.2 Rolle's theorem

Theorem 5.2.2 (Rolle's theorem)**.** *Suppose $c < d$ and suppose that $f : [c, d] \to \mathbb{R}$ is differentiable with $f(c) = f(d)$. Then there exists $a \in (c, d)$ such that $f'(a) = 0$.*

Proof of Theorem 5.2.2. By replacing f with $f - f(c)$, we may assume without loss of generality that $f(c) = f(d) = 0$. If f is identically 0, then f' is also identically 0 so the theorem is true in this case. Thus we may assume that f is not identically 0.

By replacing f with $-f$, we may assume that $f(b) > 0$ for some $b \in (c, d)$. Since f is continuous, the extreme value theorem implies that there exists $a \in [c, d]$ such that $f(a) = \sup_{c \le x \le d} f(x)$. We have $f(a) \ge f(b) > 0 = f(c) = f(d)$, so that $a \in (c, d)$. Then the interior extemum theorem implies that $f'(a) = 0$. □

5.2.3 Mean value theorem

Theorem 5.2.3 (Mean value theorem)**.** *Suppose $c < d$ and suppose that $f : [c, d] \to \mathbb{R}$ is differentiable. Then there exists $a \in [c, d]$ such that*

$$f'(a) = \frac{f(d) - f(c)}{c - d}$$

Proof of Theorem 5.2.3. Define

$$g(x) = f(x) - f(c) - \frac{f(d) - f(c)}{d - c}(x - c).$$

We can compute that $g(c) = g(d) = 0$. By Rolle's theorem (Theorem 5.2.2), there exists a point $a \in [c.d]$ such that $g'(a) = 0$. We have

$$g'(a) = f'(a) - \frac{f(d) - f(c)}{d - c}$$

which establishes Theorem 5.2.3. □

Chapter 6

Integration

6.1 Fundamental definitions

Definition 6.1.1. *Let $[a,b]$ be a closed interval of finite length. A* ***partition*** *of $[a,b]$ is a strictly increasing finite list of points $x_0 < x_1 < \cdots < x_{n-1} < x_n$ in $[a,b]$ such that $x_0 = a$ and $x_n = b$. The* ***norm*** *of a partition is the positive number*

$$\max(x_1 - x_0, x_2 - x_1, \ldots, x_n - x_{n-1})$$

We denote the norm of a partition $\mathcal{P}$ by $||\mathcal{P}||$.

Definition 6.1.2. *Given a partition $\mathcal{P} = \{x_0, \ldots, x_n\}$ of $[a,b]$, a* ***set of tags*** *for $\mathcal{P}$ is a list $t_1, \ldots, t_n$ of real numbers such that $x_{k-1} \leq t_k \leq x_k$ for all $k \in \{1, \ldots, n\}$. A* ***tagged partition*** *is a partition with a distinguished set of tags.*

Definition 6.1.3. *If $\mathcal{P}$ is a tagged partition of $[a,b]$ and a function $f : [a,b] \to \mathbb{R}$, the* ***Riemann sum*** *of f over $\mathcal{P}$ is denoted $S(\mathcal{P}, f)$ and is defined by*

$$S(\mathcal{P}, f) = \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1})$$

Definition 6.1.4. *A function $f : [a,b] \to \mathbb{R}$ is integrable if there exists a real number α such that for every $\epsilon > 0$ there exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of $]a,b]$ such that $||\mathcal{P}|| \leq \delta$ then $|S(\mathcal{P}, f) - \alpha| \leq \epsilon$. If this is the case we say that α is the integral of f and write*

$$\alpha = \int_a^b f(x)\,\mathrm{d}x$$

6.2 Uniqueness of the integral

Theorem 6.2.1. *The integral of an integrable function is unique.*

Proof of Theorem 6.2.1. Let $f : [a,b] \to \mathbb{R}$ be an integrable function and assume that α and β satisfy the condition in Definition 6.1.4. Let $\epsilon > 0$. We can find $\delta > 0$ such that if $\mathcal{P}$ is a partition of $[a,b]$ with $||\mathcal{P}|| \leq \delta$ then $|S(\mathcal{P}, f) - \alpha| \leq \frac{\epsilon}{2}$. Similarly, we can find $\delta' > 0$ such that if $\mathcal{P}$ is a partition of $[a,b]$ with $||\mathcal{P}|| \leq \delta'$ then $|S(\mathcal{P}, f) - \beta| \leq \delta$.

Let $n \in \mathbb{N}$ be such that $\frac{b-a}{n} \leq \min(\delta, \delta')$ and consider the partition

$$\mathcal{P} = \left\{a, a + \frac{b-a}{n}, a + \frac{2(b-a)}{n}, \ldots, a + \frac{(n-1)(b-a)}{b}, b\right\}$$

We have $||\mathcal{P}|| \leq \min(\delta, \delta')$. Thus we can use the triangle inequality to obtain

$$|\alpha - \beta| \leq |\alpha - S(\mathcal{P}, f)| + |S(\mathcal{P}, f) - \beta| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

□

6.3 Criteria for integrability

Theorem 6.3.1 (Cauchy criterion)**.** *A function $f : [a,b] \to \mathbb{R}$ is integrable if and only if the following condition holds. For every $\epsilon > 0$ there exists $\delta > 0$ such that if $\mathcal{P}$ and $\mathcal{Q}$ are partitions of $[a,b]$ with* $\max(||\mathcal{P}||, ||\mathcal{Q}||) \leq \delta$ *then* $|S(\mathcal{P}, f) - S(\mathcal{Q}, f)| \leq \epsilon$.

Proof of Theorem 6.3.1. We treat each implication separately.

Integrability implies Cauchy Assume that f is integrable and let $\epsilon > 0$. Writing α for the integral of f, we can find $\delta > 0$ such that if $\mathcal{P}$ is a partition with $||\mathcal{P}|| \leq \delta$ then we have $|S(\mathcal{P}, f) - \alpha| \leq \frac{\epsilon}{2}$. Then if $\mathcal{Q}$ is another partition with $||\mathcal{Q}|| \leq \delta$ we also have $|S(\mathcal{Q}, f) - \alpha| \leq \delta$. Using the triangle inequality we obtain

$$|S(\mathcal{P}, f) - S(\mathcal{Q}, f)| \leq |S(\mathcal{P}, f) - \alpha| + |\alpha - S(\mathcal{Q}, f)| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Cauchy implies integrability Assume that f satisfies the Cauchy criterion. For each $n \in \mathbb{N}$ we can find δ_n such that if $\mathcal{P}$ and $\mathcal{Q}$ are partitions with $\max(||\mathcal{P}||, ||\mathcal{Q}||) \leq \delta_n$ then $|S(\mathcal{P}, f) - S(\mathcal{Q}, f)| \leq \frac{1}{n}$. We may assume without loss of generality that $\delta_{n+1} \leq \delta_n$.

For each n, let $\mathcal{P}_n$ be a partition with $||\mathcal{P}_n|| \leq \delta_n$. If $m > n$ then since $||\mathcal{P}_m|| \leq \delta_m \leq \delta_n$ we have $|S(\mathcal{P}_m, f) - S(\mathcal{P}_n, f)| \leq \frac{1}{n}$. This implies that $(S(\mathcal{P}_n, f))_{n=1}^{\infty}$ is a Cauchy sequence of real numbers. By Theorem 3.1.6 this sequence is convergent. Let $\alpha = \lim_{n\to\infty} S(\mathcal{P}_n, f)$.

Now, let $\epsilon > 0$. We can find $N \in \mathbb{N}$ such that $n \geq N$ implies $|S(\mathcal{P}_n, f) - \alpha| \leq \frac{\epsilon}{2}$. We may assume that $\frac{1}{N} \leq \frac{\epsilon}{2}$. Let $\mathcal{Q}$ be a partition with $||\mathcal{Q}|| \leq \delta_n$, so that

$$|S(\mathcal{Q}, f) - S(\mathcal{P}_n, f)| \leq \frac{1}{n} \leq \frac{1}{N} \leq \frac{\epsilon}{2}$$

Using the triangle inequality we have

$$|S(\mathcal{Q}, f) - \alpha| \leq |S(\mathcal{Q}, f) - S(\mathcal{P}_n, f)| + |S(\mathcal{P}_n, f) - \alpha| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

□

Definition 6.3.1. *Let $f : [a, b] \to \mathbb{R}$ be an integrable function and let $\mathcal{P} = \{x_0, \ldots, x_n\}$ be a partition of $[a, b]$. For each $k \in \{1, \ldots, n\}$ define $u_k = \sup_{x_{k-1} \le x \le x_k} f(x)$ and $l_k = \inf_{x_{k-1} \le x \le x_k} f(x)$. The **upper Riemann sum** of f over $\mathcal{P}$ is denoted $U(\mathcal{P}, f)$ and is defined by*

$$U(\mathcal{P}, f) = \sum_{k=1}^{n} u_k(x_k - x_{k-1})$$

*The **lower Riemann sum** of f over $\mathcal{P}$ is denoted $L(\mathcal{P}, f)$ and is defined by*

$$L(\mathcal{P}, f) = \sum_{k=1}^{n} l_k(x_k - x_{k-1})$$

Theorem 6.3.2 (Max-min criterion)**.** *A function $f : [a, b] \to \mathbb{R}$ is integrable if and only if for every $\epsilon > 0$ there exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of $[a, b]$ with $||\mathcal{P}|| \le \delta$ then $U(\mathcal{P}, f) - L(\mathcal{P}, f) \le \epsilon$.*

Proof of Theorem 6.3.2. It is immediate that the Cauchy criterion in Theorem 6.3.1 implies the criterion in Theorem 6.3.2. We now show that the criterion in Theorem 6.3.2 implies the Cauchy criterion in Theorem 6.3.1. Let $\epsilon > 0$ and let $\delta > 0$ be such that if $\mathcal{P}$ is a partition with $||\mathcal{P}|| \le \delta$ then $U(\mathcal{P}, f) - L(\mathcal{P}, f) \le \frac{\epsilon}{2}$. Let $\mathcal{Q}$ be another partition with $||\mathcal{Q}|| \le \delta$. It follows directly from the definition that if $S(\mathcal{P}, f)$ is a Riemann sum with an arbitrary set of tags then $U(\mathcal{P}, f) \ge S(\mathcal{P}, f) \ge L(\mathcal{P}, f)$ and $U(\mathcal{Q}, f) \ge S(\mathcal{Q}, f) \ge L(\mathcal{Q}, f)$.

Now, let $\mathcal{R}$ be the common refinement of $\mathcal{P}$ and $\mathcal{Q}$. This means that $\mathcal{R}$ is the partition whose set of subinterval endpoints is the union of the set of $\mathcal{P}$ endpoints and $\mathcal{Q}$ endpoints. We have $U(\mathcal{R}, f) \le \min(U(\mathcal{P}, f), U(\mathcal{Q}, f))$ and $L(\mathcal{R}, f) \ge \max(L(\mathcal{P}, f), L(\mathcal{Q}, f))$. Therefore the intervals $[L(\mathcal{P}, f), U(\mathcal{P}, f)]$ and $[L(\mathcal{Q}, f), U(\mathcal{Q}, f)]$ must overlap. Since these intervals have length at most $\frac{\epsilon}{2}$ we see that $|S(\mathcal{P}, f) - S(\mathcal{Q}, f)| \le \epsilon$ as required. ☐

6.4 Integrability of continuous functions

Theorem 6.4.1. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f is integrable.*

Proof of Theorem 6.4.1. Let $\epsilon > 0$. By Theorem 4.3.7 there exists $\delta > 0$ such that if $|x - y| \le \delta$ then $|f(x) - f(y)| \le \frac{\epsilon}{b-a}$. Let $\mathcal{P} = \{x_0, \ldots, x_n\}$ be a partition with $||\mathcal{P}|| \le \delta$ and for each $k \in \{1, \ldots, n\}$ let u_k and l_k be as Definition 6.3.1.

By the extreme value theorem (Theorem 4.3.6) there exist $s_k \in [x_{k-1}, x_k]$ and $v_k \in [x_{k-1}, x_k]$ such that $f(s_k) = u_k$ and $f(v_k) = l_k$. Since $x_k - x_{k-1} \le \delta$ we have $|s_k - v_k| \le \delta$ and therefore $u_k - l_k \le \frac{\epsilon}{b-a}$. Thus we can estimate

$$\begin{aligned} U(\mathcal{P}, f) - L(\mathcal{P}, f) &= \sum_{k=1}^{n} u_k(x_k - x_{k-1}) - \sum_{k=1}^{n} l_k(x_k - x_{k-1}) \\ &= \sum_{k=1}^{n} (u_k - l_k)(x_k - x_{k-1}) \\ &\le \sum_{k=1}^{n} \frac{\epsilon}{b-a}(x_k - x_{k-1}) \end{aligned}$$

$$
\begin{aligned}
&= \frac{\epsilon}{b-a}\sum_{k=1}^{n}(x_k - x_{k-1}) \\
&= \frac{\epsilon}{b-a}(b-a) \\
&= \epsilon
\end{aligned}
$$

Thus Theorem 6.3.2 implies that f is integrable. □

6.5 First fundamental theorem

Theorem 6.5.1 (First fundamental theorem of calculus)**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Define a function $g : [a, b] \to \mathbb{R}$ by*

$$
g(y) = \int_a^y f(x)\,\mathrm{d}x
$$

Then $g'(y) = f(y)$ for all $y \in [a, b]$

Proof of Theorem 6.5.1. We have

$$
\begin{aligned}
\frac{g(y+h) - g(y)}{h} &= \frac{1}{h}\left(\int_a^{y+h} f(x)\,\mathrm{d}x - \int_a^y f(x)\,\mathrm{d}x\right) \\
&= \frac{1}{h}\int_y^{y+h} f(x)\,\mathrm{d}x
\end{aligned}
$$

Let $\epsilon > 0$. Since f is continuous, we can find $\delta > 0$ such that if $|y - x| \leq \delta$ then $|f(x) - f(y)| \leq \epsilon$. We have

$$
\begin{aligned}
\left|\frac{1}{h}\int_y^{y+h} f(x)\,\mathrm{d}x - f(y)\right| &= \left|\frac{1}{h}\int_y^{y+h} f(x)\,\mathrm{d}x - \frac{1}{h}\int_y^{y+h} f(y)\,\mathrm{d}y\right| \\
&= \left|\frac{1}{h}\int_y^{y+h} f(x) - f(y)\,\mathrm{d}x\right| \\
&\leq \frac{1}{h}\int_y^{y+h} |f(x) - f(y)|\,\mathrm{d}x \\
&\leq \frac{1}{h}\int_y^{y+h} \epsilon\,\mathrm{d}x \\
&= \epsilon
\end{aligned}
$$

□

6.6 Second fundamental theorem

Theorem 6.6.1 (Second fundamental theorem of calculus)**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Let $F : [a, b] \to \mathbb{R}$ satisfy $F'(x) = f(x)$ for all $x \in [a, b]$. Then we have*

$$
\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)
$$

Proof of Theorem 6.6.1. Let $\epsilon > 0$. By Theorem 4.3.7 there exists $\delta > 0$ such that if $|x-y| \leq \delta$ then $|f(x) - f(y)| \leq \frac{\epsilon}{b-a}$. Let $\mathcal{P} = \{x_0, \dots, x_n\}$ be a partition of $[a,b]$ with $||\mathcal{P}|| \leq \delta$ and let $t_1, \dots, t_n$ be a set of tags for $\mathcal{P}$.

By the mean value theorem (Theorem 5.2.3) for every $k \in \{1, \dots, n\}$ there exists $s_k \in [x_{k-1}, x_k]$ such that

$$\frac{F(x_k) - F(x_{k-1})}{x_k - x_{k-1}} = f(s_k)$$

or equivalently

$$F(x_k) - F(x_{k-1}) = f(s_k)(x_k - x_{k-1})$$

We have

$$\begin{aligned}
\left| \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}) - \sum_{k=1}^{n} f(s_k)(x_k - x_{k-1}) \right| &= \left| \sum_{k=1}^{n} (f(t_k) - f(s_k)(x_k - x_{k-1}) \right| \\
&\leq \sum_{k=1}^{n} |f(t_k) - f(s_k)|(x_k - x_{k-1}) \\
&\leq \frac{\epsilon}{b-a} \sum_{k=1}^{n} (x_k - x_{k-1}) \\
&= \epsilon
\end{aligned}$$

On the other hand

$$\sum_{k=1}^{n} f(s_k)(x_k - x_{k-1}) = \sum_{k=1}^{n} (F(x_k) - F(x_{k-1})) = F(b) - F(a)$$

Therefore $|S(\mathcal{P}, f) - (F(b) - F(b)| \leq \epsilon$. □

Department of Mathematics
The University of Texas at Austin
pjburton@math.utexas.edu