Contents

I Relations and functions 1
1 Relations 1
2 Functions 4

II Elements of set theory 6
3 Finite, infinite and countable sets 6
4 Uncountable sets and the notion of continuum 12
5 Cardinality 16
6 The power set and the hierarchy of cardinalities 18

III Convergence and continuity 20
7 Subsequences and accumulation points 20
8 The Bolzano-Weierstrass theorem 24
9 Limit superior and limit inferior 26
10 Cauchy sequences 29
11 Uniformly continuous functions 31
12 Pointwise and Uniform convergence 35

IV The Riemann integral 38
13 Definition of the integral 38
14 Criterion of integrability 42
15 Classes of integrable functions 43
16 Inequalities and the mean-value property of the integral 45
17 Further properties of the integral 46
18 Integration as the inverse to differentiation 49
Notation

\( \mathbb{N} \) the set of natural numbers \( \{0, 1, 2, 3, \ldots \} \)

\( \mathbb{N}^+ \) the set of positive natural numbers \( \{1, 2, 3, \ldots \} \)

\( \mathbb{Z} \) the set of integer numbers \( \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots \} \)

\( \mathbb{Q} \) the set of rational numbers \( \{r = \frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}^+, \text{hcf}(p, q) = 1\} \)

\( \mathbb{R} \) the set of real numbers

\( \mathbb{R}_+ \) the set of non-negative real numbers \( \{x \in \mathbb{R} \mid x \geq 0\} \)

In the remainder of the text ‘Theorem A 1.2.3’ etc. refers to ‘Theorem 1.2.3’ in Lecture Notes in Analysis 1 (MATH 11006) by Vitali Liskevich.
Part I
Relations and functions

Recommended texts:

1 Relations

**Definition 1.1** Let \(X, Y\) be sets. A set \(R \subseteq X \times Y\) is called a relation from \(X\) to \(Y\).

If \((x, y) \in R\), we say that \(x\) is in relation \(R\) to \(y\). We also write \(xRy\).

**Example 1.2**
1. Let \(A = \{1, 2, 3\}\), \(B = \{3, 4, 5\}\). The set \(R = \{(1, 3), (1, 5), (3, 3)\}\) is a relation from \(A\) to \(B\) since \(R \subseteq A \times B\).
2. \(G = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x > y\}\) is a relation from \(\mathbb{R}\) to \(\mathbb{R}\).

**Definition 1.3** Let \(R\) be a relation from \(X\) to \(Y\). The domain of \(R\) is the set
\[
D(R) = \{x \in X \mid \exists y \in Y \ [(x, y) \in R]\}.
\]
The range of \(R\) is the set
\[
\text{Ran}(R) = \{y \in Y \mid \exists x \in X \ [(x, y) \in R]\}.
\]
The inverse of \(R\) is the relation \(R^{-1}\) from \(Y\) to \(X\) defined as follows
\[
R^{-1} = \{(y, x) \in Y \times X \mid (x, y) \in R\}.
\]

**Definition 1.4** Let \(R\) be a relation from \(X\) to \(Y\), \(S\) be a relation from \(Y\) to \(Z\). The composition of \(S\) and \(R\) is a relation from \(X\) to \(Z\) defined as follows
\[
S \circ R = \{(x, z) \in X \times Z \mid \exists y \in Y \ [(x, y) \in R] \land [(y, z) \in S]\}.
\]

**Theorem 1.5** Let \(R\) be a relation from \(X\) to \(Y\), \(S\) be a relation from \(Y\) to \(Z\), \(T\) be a relation from \(Z\) to \(V\). Then
1. \((R^{-1})^{-1} = R\).
2. $D(R^{-1}) = \text{Ran}(R)$.

3. $\text{Ran}(R^{-1}) = D(R)$.

4. $T \circ (S \circ R) = (T \circ S) \circ R$.

5. $(S \circ R)^{-1} = R^{-1} \circ S^{-1}$.

Proof. Exercise, or see [Stewart and Tall (1977)].

Next we take a look at some particular types of relations. Let us consider relations from $X$ to $X$, i.e. subsets of $X \times X$. In this case, we talk about a relation on $X$. A simple example of such a relation is the identity relation on $X$ which is defined by

$$i_X = \{(x, y) \in X \times X \mid x = y\}.$$

Definition 1.6 1. A relation $R$ on $X$ is said to be reflexive if

$$(\forall x \in X) (x, x) \in R.$$  

2. $R$ is said to be symmetric if

$$(\forall x \in X)(\forall y \in X) \{[x, y) \in R] \Rightarrow [(y, x) \in R]\}.$$

3. $R$ is said to be transitive if

$$(\forall x \in X)(\forall y \in X)(\forall z \in X)\{[(x, y) \in R] \land [(y, z) \in R] \Rightarrow [(x, z) \in R]\}.$$

Equivalence relations

A particularly important class of relations are equivalence relations.

Definition 1.7 A relation $R$ on $X$ is said to be an equivalence relation if it is reflexive, symmetric and transitive.

Example 1.8 1. Let $X$ be the set of students. Define a relation on $X$ by

$$R = \{(x, y) \in X \times X : x \text{ and } y \text{ are friends}\}.$$

Then $R$ is reflexive (assuming that each student is a friend to him or herself). It is also symmetric. But it is not transitive.

2. Let $X = \mathbb{R}$, $a$ be some positive number. Define $R \subseteq X \times X$ as

$$R = \{(x, y) \mid |x - y| \leq a\}.$$

$R$ is reflexive, symmetric, but not transitive.

3. Let $X = \mathbb{Z}$ and $m \in \mathbb{Z}$. Define

$$R := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : (\exists k \in \mathbb{Z})[x - y = km]\}.$$

Then $R$ is an equivalence relation on $\mathbb{Z}$. 

2
Definition 1.9 Let $R$ be an equivalence relation on $X$. Let $x \in X$. The equivalence class of $x$ with respect to $R$ is the set 

$$[x]_R = \{y \in X \mid (y, x) \in R\}.$$ 

Let us take a look at several properties of equivalence classes.

Proposition 1.10 Let $R$ be an equivalence relation on $X$. Then

1. $(\forall x \in X) \ x \in [x]_R.$

2. $(\forall x \in X)(\forall y \in X) \ [(y \in [x]_R) \iff ([y]_R = [x]_R)].$

Proof. 1. Since $R$ is reflexive, $(x, x) \in R$, hence $x \in [x]_R$.

2. First, let $y \in [x]_R$, so that $(y, x) \in R$. Suppose that $z \in [y]_R$. Then $(z, y) \in R$. By transitivity, $(z, x) \in R$; hence, $z \in [x]_R$. This shows that $[y]_R \subseteq [x]_R$. The reverse inclusion $[x]_R \subseteq [y]_R$ can be shown similarly. Therefore, $[x]_R = [y]_R$.

The implication $([y]_R = [x]_R) \Rightarrow (y \in [x]_R)$ is obvious. □

From the above proposition it follows that equivalence classes $[x]_R$ and $[y]_R$ either coincide or are disjoint, in other words,

$$[x]_R = [y]_R \text{ or } [x]_R \cap [y]_R = \emptyset.$$ 

Also, every element of the set $X$ belongs to an equivalence class. This may be expressed as 

$$\bigcup_{x \in X} [x]_R = X.$$ 

Example 1.11 Let $R$ be the equivalence class in Example 3 above. Then, for $x \in \mathbb{Z}$,

$$[x] = \{x + km : k \in \mathbb{Z}\}.$$ 

This is called the equivalence class of integers congruent to $x$ modulo $m$. Often we write $y \equiv x$ to mean $y \in [x]$.

Remark 1.12 For more on this topic, see [Stewart and Tall (1977)].
2 Functions

The notion of a function is of fundamental importance in all branches of mathematics. You
met functions in your previous study of mathematics, but without a precise definition. Here
we give a definition and make connections to examples you came across before.

Definition 2.1 Let $X$ and $Y$ be sets. Let $F$ be a relation from $X$ to $Y$. Then $F$ is called a
function from $X$ to $Y$, if the following properties are satisfied:

(i) $(\forall x \in X)(\exists y \in Y) [(x, y) \in F].$

(ii) $(\forall x \in X)(\forall y \in Y)(\forall z \in Y) \{([x, y] \in F] \land [(x, z) \in F] \Rightarrow (y = z)).$

It is customary to write $y = F(x)$, the image of $x$ under $F$. $X$ is called the domain of $F$
and $Y$ is called codomain.

Let us consider several examples.

Example 2.2 (i) Let $X = \{1, 2, 3\}$, $Y = \{4, 5, 6\}$. Define $F \subseteq X \times Y$ as

$$F = \{(1, 4), (2, 5), (3, 5)\}.$$ 

Then $F$ is a function. In contrast to this, define $G \subseteq X \times Y$ via

$$G = \{(1, 4), (1, 5), (2, 6), (3, 6)\}.$$ 

Then $G$ is not a function.

(ii) Let $X = \mathbb{R}$ and $Y = \mathbb{R}$. Define $F \subseteq X \times Y$ to be

$$F = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid y = x^2\}.$$ 

Then $F$ is a function from $\mathbb{R}$ to $\mathbb{R}$. In contrast to this, define $G \subseteq X \times Y$ to be

$$G = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x^2 + y^2 = 1\}.$$ 

Then $G$ is not a function.

We often use the notation

$$F : X \to Y$$

for a function $F$ from $X$ to $Y$.

Note that in order to define a function $F$ from $X$ to $Y$ we have to define $X$, $Y$ and a subset
of $X \times Y$ satisfying (i) and (ii) of the definition.

Note: one can assign a value $y$ to each value $x \in X$ by means of a “rule” or “formula”. But
this notion is rather vague. The present definition is more precise.

Theorem 2.3 Let $X, Y$ be sets and $F, G$ functions from $X$ to $Y$. Then

$$[(\forall x \in X)(F(x) = G(x))] \iff (F = G).$$
Proof. 1. \((\Rightarrow)\). Assume the LHS above. We first show that \(F \subseteq G\). Let \((x, y) \in F\). Then \(y = F(x)\). So \(y = G(x);\) or, \((x, y) \in G\). The proof of the inclusion \(G \subseteq F\) is similar. The implication \((\Leftarrow)\) is trivial. \(\Box\)

Thus, two functions are equal if they share the same domain and codomain and elements in the domain have equal images.

**Example 2.4** Let \(f : \mathbb{R} \to \mathbb{R}, \ g : \mathbb{R} \to \mathbb{R}, \ h : \mathbb{R} \to \mathbb{R}_+\). Here, \(\mathbb{R}_+\) denotes the set of non-negative real numbers. Suppose that for all \(x \in \mathbb{R},\)

\[
f(x) = (x + 1)^2, \quad g(x) = x^2 + 2x + 1, \quad h(x) = (x + 1)^2.
\]

Then \(f\) and \(g\) are equal, but \(f\) and \(h\) are not since they have different codomains.

The definition of composition of relations can be also applied to functions. If \(f : X \to Y\) and \(g : Y \to Z\) then

\[
g \circ f = \{(x, z) \in X \times Z \mid (\exists y \in Y)[[(x, y) \in f] \land [(y, z) \in g]]\}.
\]

**Theorem 2.5** Let \(f : X \to Y\) and \(g : Y \to Z\). Then \(g \circ f : X \to Z\) and

\[
(\forall x \in X) \ [(g \circ f)(x) = g(f(x))].
\]

*Proof.* We know that \(g \circ f\) is a relation. So we must prove that for every \(x \in X\) there exists a unique element \(z \in Z\) such that \((x, z) \in g \circ f\).

*Existence:* Let \(x \in X\) be arbitrary. As \(f\) is a function, \(\exists y \in Y\) such that \((x, y) \in f\). As \(g\) is a function, \(\exists z \in Z\) such that \((y, z) \in g\). So

\[
(\exists y \in Y)[(x, y) \in f \land (y, z) \in g]
\]

and \((x, z) \in g \circ f\). In short,

\[
(\forall x \in X)(\exists z \in Z)[(x, z) \in g \circ f].
\]

*Uniqueness:* Let \(x \in X\). Suppose that \((x, z_1) \in g \circ f\) and \((x, z_2) \in g \circ f\). As \((x, z_1) \in g \circ f\), \((\exists y_1 \in Y)[(x, y_1) \in f \land (y_1, z_1) \in g]\). As \((x, z_2) \in g \circ f\), \((\exists y_2 \in Y)[(x, y_2) \in f \land (y_2, z_2) \in g]\). As \(f\) is a function, \(y_1 = y_2\). As \(g\) is a function, \(z_1 = z_2\). \(\Box\)

**Example 2.6** Let \(f : \mathbb{R} \to \mathbb{R}, \ g : \mathbb{R} \to \mathbb{R},\)

\[
f(x) = \frac{1}{x^2 + 2}, \quad g(x) = 2x - 1.
\]

Find \((f \circ g)(x)\) and \((g \circ f)(x)\).

*Solution.*

\[
(f \circ g)(x) = f(g(x)) = \frac{1}{(g(x))^2 + 2} = \frac{1}{(2x - 1)^2 + 2},
\]

\[
(g \circ f)(x) = g(f(x)) = 2f(x) - 1 = \frac{2}{x^2 + 2} - 1.
\]

*Warning:* As is clear from the above, \(f \circ g \neq g \circ f\) in general.
Part II

Elements of set theory

Recommended texts:


In this part of the course we study the notion of *cardinality*. The aim is to formulate a notion of size of an infinite set, or to count its elements. It is possible to ask questions such as *how many rational numbers are there?* or *is the set of real numbers bigger then the set of natural numbers?* To do so, however, it is helpful to be able to compare the number of elements in a given set with those in another. We say that sets $X$ and $Y$ have the same *cardinality* (intuitively, number of elements) if there is a bijection $f : X \rightarrow Y$. This idea was first conceived by the German mathematician Georg Cantor (1845 – 1918) at the end of the 19th century. His “Naive Set Theory” gives rise to various *paradoxes* (e.g. Russell’s Paradox, see p. 19), which are resolved in the subsequent *Axiomatic Set Theory*. For our purposes, we take the “naive” notion of *set* as a *collection of elements*.

3 Finite, infinite and countable sets

Let $X$, $Y$ be sets and $f : X \rightarrow Y$ a function from $X$ to $Y$. Recall that the function $f$ is an injection if

$$(\forall x_1 \in X)(\forall x_2 \in X)[(f(x_1) = f(x_2)) \Rightarrow (x_1 = x_2)].$$

This means that $f$ is a one-to-one correspondence between $X$ and the range $\text{Ran}(f)$. The function $f$ is a surjection if

$$(\forall y \in Y)(\exists x \in X)[f(x) = y].$$

This means that $\text{Ran}(f) = Y$. Finally, $f$ is a bijection if $f$ is an injection and surjection. In this case, $f$ establishes a one-to-one correspondence between $X$ and $Y$.

How do we decide how many elements there are in a set? One answer is to count them. But what does this mean? This can be described as follows. We take an element of the set, and think *one*. We then look at a different element and think *two*, and so on. We need to be careful to make sure that all elements have been counted, and none has been counted twice. If this process terminates, we say that the set is finite.

**Definition 3.1** A set $X$ is said to be *finite* if either $X = \emptyset$ or for some $n \in \mathbb{N}^+$ there exists a bijection

$$f : \{1, 2, \ldots, n\} \rightarrow X.$$
In the former case, we say that $X$ has zero elements; in the latter, that $X$ has $n$ elements. $X$ is said to be **infinite** if it is not finite.

A natural way to extend this process of *counting* to infinite sets is as follows.

**Definition 3.2** A set $X$ is said to be **countable** if there exists a bijection $f : \mathbb{N}^+ \to X$.

**Remark 3.3** By Theorem A1.7.6, a function $f : \mathbb{N}^+ \to X$ is a bijection if and only if the inverse function $f^{-1} : X \to \mathbb{N}^+$ exists. In this case, the inverse $f^{-1}$ is a bijection too. In particular, $X$ is countable if and only if there exists a bijection $f : X \to \mathbb{N}^+$.

**Example 3.4** $\mathbb{N} = \{0, 1, 2, \ldots\}$ is countable.

**Proof.** Define $f : \mathbb{N}^+ \to \mathbb{N}$ by $f(n) = n - 1$.

\[
\begin{array}{ccccccc}
1 & 2 & 3 & 4 & 5 & \ldots & n & \ldots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\
0 & 1 & 2 & 3 & 4 & \ldots & n - 1 & \ldots \\
\end{array}
\]

Then $f$ is a bijection. \qed

In this example, we observe the first remarkable property of infinite sets. $\mathbb{N}^+$ is a subset of $\mathbb{N}$, so intuitively it should have fewer elements. Yet $\mathbb{N}^+$ and $\mathbb{N}$ are both countable. The next example, constructed by Galileo in 1638, was considered as paradox for more than two centuries.

**Example 3.5** The set of perfect squares $S := \{1^2, 2^2, 3^2, \ldots, n^2, \ldots\}$ is countable.

**Proof.** Define $f : \mathbb{N}^+ \to S$ by $f(n) = n^2$.

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & \ldots & n & \ldots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
1 & 4 & 9 & 16 & 25 & \ldots & n^2 & \ldots \\
\end{array}
\]

It is clear that $f$ is a bijection. \qed

The reason Galileo’s example might be considered a paradox is that the set $S$ of perfect squares seems *smaller* then $\mathbb{N}^+$. But $S$ is still countable.

**Example 3.6** The set of integers $\mathbb{Z}$ is countable.

**Proof.** Define $f : \mathbb{N}^+ \to \mathbb{Z}$ by

\[
f(n) = \begin{cases} 
n \cdot \frac{n}{2}, & \text{if } n \text{ is even}, \\
\frac{1-n^2}{2}, & \text{if } n \text{ is odd}. \end{cases}
\]

The following diagram illustrates $f$:

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & \ldots & 2n & 2n + 1 & \ldots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
0 & 1 & -1 & 2 & -2 & \ldots & n & -n & \ldots \\
\end{array}
\]
The inverse of \( f \) is the function \( f^{-1} : \mathbb{Z} \to \mathbb{N}^+ \) given by

\[
f^{-1}(m) = \begin{cases} 
2m, & \text{if } m \geq 1, \\
1 - 2m, & \text{if } m \leq 0.
\end{cases}
\]

By Theorem A1.7.6 we conclude that \( f \) is a bijection.

**Remark 3.7** Notice that although \( f \) in the last example is a bijection, it does not preserve order, in the sense that \( m < n \) does not imply \( f(m) < f(n) \).

**Example 3.8** The set \( \mathbb{N}^+ \times \mathbb{N}^+ \) is countable.

**Proof.** We write elements of \( \mathbb{N}^+ \times \mathbb{N}^+ \) in a rectangular array (figure on the left). We then read them off along the cross diagonals (figure on the right): first \((1,1)\), next \((1,2), (2,1)\), then \((1,3), (2,2), (3,1)\) and so on.

<table>
<thead>
<tr>
<th>((n,m))</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(1,1)</td>
<td>(1,2)</td>
<td>(1,3)</td>
<td>(1,4)</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>(2,1)</td>
<td>(2,2)</td>
<td>(2,3)</td>
<td>...</td>
<td>(f(n,m))</td>
</tr>
<tr>
<td>3</td>
<td>(3,1)</td>
<td>(3,2)</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>(4,1)</td>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This correspondence can also be written

\[
(1,1) \quad (1,2) \quad (2,1) \quad (1,3) \quad (2,2) \quad (3,1) \quad ... \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad ... \\
\]

In this way we establish a bijection between \( \mathbb{N}^+ \times \mathbb{N}^+ \) and \( \mathbb{N}^+ \). This bijection \( f : \mathbb{N}^+ \times \mathbb{N}^+ \to \mathbb{N}^+ \) can be represented analytically by the formula

\[
f(n,m) = \frac{(n+m-2)(n+m-1)}{2} + n.
\]

Therefore \( \mathbb{N}^+ \times \mathbb{N}^+ \) is countable.

**Remark 3.9** Let \( X \) be a countable set and \( f : \mathbb{N}^+ \to X \) a bijection. Define \( x_1 := f(1), x_2 := f(2), x_3 := f(3), \ldots, x_n := f(n), \ldots \).

\[
1 \quad 2 \quad 3 \quad 4 \quad 5 \quad ... \quad n \quad ... \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad ... \quad x_n \quad ...
\]

Thus all elements of \( X \) can be listed in a sequence:

\[
X = \{x_1, x_2, x_3, x_4, x_5, \ldots, x_n, \ldots\}.
\]

We now establish several important properties of countable sets.

**Lemma 3.10** If a set \( X \) is countable then any subset \( A \subseteq X \) is finite or countable.
Proof. Assume $A$ is infinite. We need to show that $A$ is countable. Since $X$ is countable there exists a bijection $f : \mathbb{N}^+ \to X$. Define $g : \mathbb{N}^+ \to A$ inductively as follows. Set
\[
g(1) := f(\min\{k \in \mathbb{N}^+ : f(k) \in A\}); \quad g(2) := f(\min\{k \in \mathbb{N}^+ : f(k) \in A \setminus \{g(1)\}\}); \quad \ldots \quad g(n+1) := f(\min\{k \in \mathbb{N}^+ : f(k) \in A \setminus \{g(1), \ldots, g(n)\}\}).
\]

It is clear that $g$ is a bijection. \hfill \Box

**Remark 3.11** We say that a set $X$ is at most countable if $X$ is finite or countable.

**Remark 3.12** Lemma 3.10 is useful for proving that a given set is countable. For example, let $P$ be the set of all prime numbers. Then $P \subseteq \mathbb{N}^+$. And it is known that the set of primes is infinite. We may then conclude that $P$ is countable.

**Lemma 3.13** Any infinite set $X$ has a countable subset $A \subseteq X$.

**Proof.** First, $X \neq \emptyset$ as $X$ is infinite. So we may choose $x_1 \in X$. Define
\[
g_1 : \{1\} \to X; 1 \mapsto x_1.
\]
Then $g_1$ is not onto, for otherwise $X$ would be finite. So we may choose $x_2 \in X \setminus \{x_1\}$. Define
\[
g_2 : \{1, 2\} \to X; j \mapsto x_j.
\]
Then $g_2$ is not onto, for otherwise $X$ would be finite. So we may choose $x_3 \in X \setminus \{x_1, x_2\}$. Continuing in this way, we obtain a sequence $(x_n)_{n \in \mathbb{N}^+}$ of distinct points in $X$. The mapping
\[
g : \mathbb{N}^+ \to A := \{x_n \in X | n \in \mathbb{N}^+\}; n \mapsto x_n
\]
is then a bijection from $\mathbb{N}^+$ to $A \subseteq X$. Thus $A$ is a countable subset of $X$. \hfill \Box

**Definition 3.14** A set $Y$ is said to be a proper subset of a set $X$ if $Y \subseteq X$ and $Y \neq X$.

**Proposition 3.15** If a set $X$ is infinite then there exists a proper subset $Y \subseteq X$ and a bijection $f : X \to Y$.

**Proof.** Let $A := \{x_1, x_2, \ldots, x_n, \ldots\}$ be a countable subset of $X$, constructed as in Lemma 3.13. Let $B := \{x_1\} \subseteq A$. Set $Y := X \setminus B$. Define $f : X \to Y$ by
\[
f(x_n) = x_{n+1} \quad \text{for } x_n \in A
\]
and
\[
f(x) = x \quad \text{for } x \in X \setminus A.
\]
Then $f$ is a bijection. \hfill \Box
Remark 3.16 We proved that if $X$ is infinite and $B = \{x_1\} \subseteq X$ then $Y := X \setminus B$ is infinite and there is a bijection $f : X \to Y$. Thus removing one element from an infinite set does not change its "size". In a similar way, it is possible to select a countable subset $B \subseteq X$, set $Y := X \setminus B$ and yet construct a bijection $f : X \to Y$. To do so, select a countable subset $A := \{x_1, x_2, \ldots, x_n, \ldots\}$ of $X$, then let $B := \{x_{2k-1} | k \in \mathbb{N}^+\} \subseteq A$, set $Y := X \setminus B$ and define $f : X \to Y$ by

$$f(x_n) = x_{2n} \text{ for } x_n \in A$$

and

$$f(x) = x \text{ for } x \in X \setminus A.$$ 

Then $f$ is a bijection.

Lemma 3.17 Let $X$ and $Y$ be countable sets. Then the Cartesian product $X \times Y$ is countable.

Proof. Let $f : X \to \mathbb{N}^+$ and $g : Y \to \mathbb{N}^+$ be bijections. Define the function $h : X \times Y \to \mathbb{N}^+ \times \mathbb{N}^+$ via

$$h(x, y) := (f(x), g(y)).$$

To see that $h$ is an injection suppose that $h(x_1, y_1) = h(x_2, y_2)$. This means that

$$(f(x_1), g(y_1)) = (f(x_2), g(y_2)).$$

So $f(x_1) = f(x_2)$ and $g(y_1) = g(y_2)$. Since $f$ and $g$ are injections it follows that $x_1 = x_2$ and $y_1 = y_2$, that is $(x_1, y_1) = (x_2, y_2)$. To check that $h$ is a surjection, suppose $(n, m) \in \mathbb{N}^+ \times \mathbb{N}^+$. Then since $f$ and $g$ are both surjections, we can choose $x \in X$ and $y \in Y$ such that $f(x) = n$ and $g(y) = m$. Therefore, $h(x, y) = (n, m)$, as required. Hence $h : X \times Y \to \mathbb{N}^+ \times \mathbb{N}^+$ is a bijection. By Example 3.8, there is a bijection $\varphi : \mathbb{N}^+ \times \mathbb{N}^+ \to \mathbb{N}^+$. It is easy to see that the composition $\varphi \circ h : X \times Y \to \mathbb{N}^+$ is also a bijection. Therefore $X \times Y$ is countable. \qed

Lemma 3.18 Let $\{X_k\}_{k \in \mathbb{N}^+}$ be a countable collection of pairwise disjoint countable sets. Then the union

$$\bigcup_{k=1}^{\infty} X_k$$

is countable.

Proof. Since every set $X_k$ is countable we can write the elements of $X_k$ in a list:

$X_1 = \{x_{11}, x_{12}, x_{13}, \ldots, x_{1n}, \ldots\},$

$X_2 = \{x_{21}, x_{22}, x_{23}, \ldots, x_{2n}, \ldots\},$

$X_3 = \{x_{31}, x_{32}, x_{33}, \ldots, x_{3n}, \ldots\},$

$\ldots$

$X_k = \{x_{k1}, x_{k2}, x_{k3}, \ldots, x_{kn}, \ldots\},$

$\ldots$

By definition of the union

$$\bigcup_{k=1}^{\infty} X_k = \{x_{kn} | k \in \mathbb{N}^+, n \in \mathbb{N}^+\}.$$
Define a function \( h : \bigcup_{k=1}^{\infty} X_k \to \mathbb{N}^+ \times \mathbb{N}^+ \) by
\[
h(x_kn) := (k, n).
\]
Clearly \( h \) is a bijection. Since \( \mathbb{N}^+ \times \mathbb{N}^+ \) is countable we conclude that \( \bigcup_{k=1}^{\infty} X_k \) is countable too.

2 Example 3.19 The set of rational numbers \( \mathbb{Q} \) is countable.

Proof. For \( n \in \mathbb{N}^+ \), let
\[
X_k := \left\{ \frac{a}{k} \mid a \in \mathbb{Z} \setminus \{0\}, (a, k) = 1 \right\}
\]
be the set of all (non-zero) fractions with denominator \( k \) written in lowest possible terms. Now for each \( k \in \mathbb{N}^+ \) the set \( X_k \) is countable, and for \( k \neq k' \), \( X_k \) and \( X_{k'} \) are disjoint. Let \( X := \bigcup_{k=1}^{\infty} X_k \). By Lemma 3.18 we conclude that the set \( X \) is countable. But \( \mathbb{Q} = X \cup \{0\} \) so \( \mathbb{Q} \) is countable by Remark 3.16. \( \square \)
4 Uncountable sets and the notion of continuum

We have shown that the sets $\mathbb{N}^+ \times \mathbb{N}^+$, $\mathbb{Q}$, and others, although seemingly larger than $\mathbb{N}^+$, are nevertheless countable. So maybe all infinite sets are countable? The answer is no.

**Definition 4.1** A set $X$ is said to be **uncountable** if $X$ is infinite but not countable.

Roughly speaking, an uncountable set has so many elements that they cannot be listed, nor arranged in a sequence.

**Theorem 4.2** The open interval $(0, 1)$ is uncountable.

**Proof.** We prove this by contradiction. Assume that the interval $(0, 1)$ is countable and $f : \mathbb{N}^+ \rightarrow (0, 1)$ is a bijection. Then we can arrange all elements of $(0, 1)$ in a sequence

$$\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \ldots, \alpha_k, \ldots\}.$$ 

Represent each $\alpha_k \in (0, 1)$ as a decimal expansion $
\alpha_k = 0.a_{k1}a_{k2}a_{k3}a_{k4}a_{k5}\ldots,n$

where we agree that if the decimal expansion terminates $^1$ we will write it ending with a sequence of zeros, not a sequence of nines. Then elements of the interval $(0, 1)$ may be listed:

\[
\begin{align*}
\alpha_1 &= 0. [a_{11}] a_{12} a_{13} a_{14} a_{15} \ldots a_{1k} \ldots \\
\alpha_2 &= 0. a_{21} [a_{22}] a_{23} a_{24} a_{25} \ldots a_{2k} \ldots \\
\alpha_3 &= 0. a_{31} a_{32} [a_{33}] a_{34} a_{35} \ldots a_{3k} \ldots \\
\alpha_4 &= 0. a_{41} a_{42} a_{43} [a_{44}] a_{45} \ldots a_{4k} \ldots \\
\alpha_5 &= 0. a_{51} a_{52} a_{53} a_{54} [a_{55}] \ldots a_{5k} \ldots \\
\ldots &= 0. \ldots \\
\alpha_k &= 0. a_{k1} a_{k2} a_{k3} a_{k4} a_{k5} \ldots [a_{kk}] \ldots \\
\ldots &= 0. \ldots 
\end{align*}
\]

Let $\beta \in (0, 1)$ be the real number with the decimal expansion

$$\beta := 0. b_1 b_2 b_3 b_4 b_5 \ldots b_k \ldots$$

where

$$b_k := \begin{cases} 
1, & \text{if } a_{kk} \neq 1, \\
2, & \text{if } a_{kk} = 1.
\end{cases}$$

Then $\beta$ is different from $\alpha_k$ for any $k \in \mathbb{N}^+$. The reason is that, for each $k \in \mathbb{N}^+$ the decimal expansion of $\beta$ differs from that of $\alpha_k$ in its $k$th place. We thus arrive at a contradiction. Hence $(0, 1)$ is uncountable. $\Box$

**Remark 4.3** The method used in the proof of Theorem 4.2 is often referred to as *(Cantor’s)* diagonalisation argument. It is a powerful technique that plays a role in many other proofs.

\[^1\text{See Proposition 4.11 below and the discussion after it.}\]
**Definition 4.4** A set $X$ is said to be a **continuum** if there exists a bijection

$$f : (0, 1) \rightarrow X.$$ 

**Example 4.5** Let $a, b \in \mathbb{R}$ and $a < b$. Then the open interval $(a, b)$ is a continuum.

**Proof.** The function $f : (0, 1) \rightarrow (a, b)$ defined by the formula

$$f(x) := (b - a)x + a$$

is a bijection with inverse $f^{-1} : (a, b) \rightarrow (0, 1)$ given by

$$f^{-1}(y) = \frac{y - a}{b - a}.$$

**Example 4.6** The interval $[0, 1)$ is a continuum (as is $(0, 1]$ and $[0, 1]$).

**Proof.** (Compare this proof with the proof of Proposition 3.15!) The function $f : [0, 1) \rightarrow (0, 1)$ defined by the formula

$$f(x) := \begin{cases} 
1 - \frac{1}{n+1}, & \text{if } x = 1 - \frac{1}{n}, \ n \in \mathbb{N}^+,
\end{cases}$$

is a bijection with inverse

$$f^{-1}(y) := \begin{cases} 
1 - \frac{1}{n}, & \text{if } y = 1 - \frac{1}{n+1}, \ n \in \mathbb{N}^+.
\end{cases}$$

The intervals $[0, 1]$ and $[0, 1]$ can be considered in a similar way.

**Example 4.7** The real line $\mathbb{R}$ is a continuum.

**Proof.** The function $g : (-1, 1) \rightarrow \mathbb{R}$ defined by

$$g(x) := \frac{x}{1 - |x|}$$

is a bijection with inverse $g^{-1} : \mathbb{R} \rightarrow (-1, 1)$ given by

$$g^{-1}(y) = \frac{y}{1 + |y|}.$$ 

By Example 4.5 there is a bijection $f : (0, 1) \rightarrow (-1, 1)$. Then the composition $g \circ f : (0, 1) \rightarrow \mathbb{R}$ is a bijection.

**Example 4.8** $[0, 1) \times [0, 1)$ is a continuum.

**Idea of proof.** Any $x \in [0, 1)$ has a decimal expansion

$$x = 0.a_1 a_2 a_3 \cdots.$$ 

We assume that the decimal expansion does not contain a string of consecutive 9’s with infinite length. The expansion is then unique. The expansion may be rewritten in blocks

$$x = 0.a_1 a_2 a_3 \cdots.$$ 

13
where each block $\alpha_j$ has the form
\[
\alpha_j = *
\]
where * stands for one of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8; or
\[
\alpha_j = 99 \cdots 9 *
\]
(a finite block of 9’s followed by a non-nine digit).

Define $f : [0, 1) \times [0, 1) \rightarrow [0, 1)$ via
\[
f(x, y) := 0. \alpha_1 \beta_1 \alpha_2 \beta_2 \alpha_3 \beta_3 \ldots .
\]
Here, we write
\[
x = 0.a_1 a_2 a_3 \cdots = 0. \alpha_1 \alpha_2 \alpha_3 \cdots ;
\]
\[
y = 0.b_1 b_2 b_3 \cdots = 0. \beta_1 \beta_2 \beta_3 \cdots ;
\]
first in decimal form, then in block form, as described above. Then $f$ is a bijection.

To conclude, compose $f$ with the bijection in Example 4.6. \qed

The interval $(0, 1)$ is uncountable – another proof.\footnote{The content of this paragraph will be excluded from the examination.} Our proof of Theorem 4.2 relies on the decimal expansion of the real numbers. Cantor’s original proof of Theorem 4.2, published in 1874, was different. It was based on the following “nested intervals” theorem, which is similar to Theorem A 3.2.3 from Analysis.

**Theorem 4.9** Let $([b_k, c_k])_{k \in \mathbb{N}^+}$ be a sequence of closed nonempty bounded intervals such that
\[
(\forall k \in \mathbb{N}^+) ([b_{k+1}, c_{k+1}] \subset [b_k, c_k]) .
\]
Then $\cap_{k \in \mathbb{N}^+} [b_k, c_k] \neq \emptyset$.

**Exercise 4.10** Prove Theorem 4.9.

**Hint.** Modify the proof of Theorem A 3.2.3 from Analysis I.

Cantor’s original proof of Theorem 4.2. Let $E \subset (0, 1)$ be a countable subset of $(0, 1)$, so $E$ can be represented as
\[
E = \{x_1, x_2, x_3, \ldots \} .
\]
Observe, that given a point $x \in (0, 1)$ there exists a nonempty interval $[b, c] \subset (0, 1)$ such that $x \notin [b, c]$.

Hence, we can choose a nonempty interval $[b_1, c_1] \subset (0, 1)$ such that $x_1 \notin [b_1, c_1]$. Now we proceed inductively. Let $n \in \mathbb{N}^+$. Having chosen a nonempty interval $([b_k, c_k])_{k=1}^n$ such that $x_k \notin [b_k, c_k]$ for $k = 1, \ldots, n$, and such that $[b_{k+1}, c_{k+1}] \subset (b_k, c_k)$ for $k = 1, \ldots, n-1$, we can choose nonempty interval $[b_{n+1}, c_{n+1}]$ such that $x_{n+1} \notin [b_{n+1}, c_{n+1}]$ and such that $[b_{n+1}, c_{n+1}] \subset (b_n, c_n)$.

In such a way we inductively defined a “nested” sequence of closed nonempty bounded intervals $([b_k, c_k])_{k \in \mathbb{N}^+}$ so that
\[
(\forall k \in \mathbb{N}^+) ([b_{k+1}, c_{k+1}] \subset [b_k, c_k]) \land (x_k \notin [b_k, c_k]) .
\]
According to Theorem 4.9, there exists a point $x_\ast \in \cap_{k \in \mathbb{N}^+} [b_k, c_k]$. But since $x_n \notin \cap_{n \in \mathbb{N}^+} [b_k, c_k]$ for every $n \in \mathbb{N}^+$, we conclude that $x_\ast \notin E$. Therefore $E \neq (0, 1)$, which means that the interval $(0, 1)$ is uncountable. \qed
Decimal expansions. Let \((a_k)_{k \in \mathbb{N}^+}\) be a sequence of numbers chosen from the set \(\{0, 1, 2, \ldots, 9\}\). Then the series
\[
\sum_{k=1}^{\infty} \frac{a_k}{10^k}
\]
converges by comparison with a geometric progression to a sum \(\alpha \in [0, 1]\). We write
\[
\alpha = 0.a_1a_2a_3a_4a_5 \ldots,
\]
and say that the right-hand side is a decimal expansion of the real number \(\alpha\). For example,
\[
\frac{1}{3} = 0.33333 \ldots, \quad \frac{1}{7} = 0.142857 \ldots, \quad \pi = 0.785398163397448 \ldots.
\]

**Proposition 4.11** Every real number \(\alpha \in [0, 1]\) has a unique decimal expansion unless it is a rational number of the form \(\alpha = \frac{m}{10^n}\) for some \(m \in \mathbb{N}^+\) and \(n \in \mathbb{N}^+\) in which case \(\alpha\) has precisely two decimal expansions.

**Sketch of the proof.** Let \(\alpha \in [0, 1]\). Suppose that \(\{a_1, a_2, \ldots, a_k\}\) have been chosen from \(\{0, 1, 2, \ldots, 9\}\) in such a way that
\[
\epsilon_k := \alpha - \left(\frac{a_1}{10} + \frac{a_2}{10^2} + \cdots + \frac{a_k}{10^k}\right)
\]
satisfies
\[
0 \leq \epsilon_k < \frac{1}{10^k}.
\]
Let \(a_{k+1}\) be chosen from \(\{0, 1, 2, \ldots, 9\}\) so that \(a_{k+1}\) is maximized subject to the constraint \(\epsilon_{k+1} \geq 0\). If \(a_{k+1} < 9\) then
\[
\epsilon_{k+1} < \frac{1}{10^{k+1}},
\]
because otherwise we could replace \(a_{k+1}\) by \(a_{k+1} + 1\). If \(a_{k+1} = 9\) then
\[
\epsilon_{k+1} = \epsilon_k - \frac{9}{10^{k+1}} < \frac{1}{10^k} - \frac{9}{10^{k+1}} = \frac{1}{10^{k+1}}.
\]
By the “sandwich rule” (A3.1.6) we conclude that \(\lim_{k \to \infty} \epsilon_k = 0\). Therefore
\[
\alpha = \sum_{k=1}^{\infty} \frac{a_k}{10^k},
\]
which gives the required decimal expansion of \(\alpha\).

We do not consider the issue of uniqueness of the decimal expansion except for the observation that, where two distinct expansions of \(\alpha\) exist then \(\alpha = \frac{m}{10^n}\) for some \(m \in \mathbb{N}^+\) and \(n \in \mathbb{N}^+\) and one of the expansions consists of zeros from some point and the other consists of nines from some point on. In this case we say that a decimal expansion of \(\alpha\) terminates. For example,
\[
\frac{1}{2} = 0.5000 \cdots = 0.4999 \ldots.
\]
If we agree that when decimal expansion terminates we will write it ending with a sequence of zeros, not a sequence of nines, then every real number \(\alpha \in [0, 1]\) has a unique decimal expansion.

---

\(^3\)The content of this paragraph will not be included in the examination paper.
5 Cardinality

In this section, we discuss a method for comparing the size of arbitrary sets.

**Definition 5.1** Let \( X, Y \) be sets. We say that \( X \) has the same cardinality as \( Y \) if there exists a bijection

\[
f : X \to Y.
\]

In this case, we write \( X \sim Y \) or \( \text{card}(X) = \text{card}(Y) \).

**Remark 5.2** Intuitively, \( X \sim Y \) or \( \text{card}(X) = \text{card}(Y) \) means that the sets \( X \) and \( Y \) have the same size or the same number of elements.

**Example 5.3** We already know that \( \mathbb{N}^+ \sim \mathbb{Z} \sim \mathbb{Q} \sim \mathbb{N}^+ \times \mathbb{N}^+ \) and \( (0,1) \sim [0,1] \sim \mathbb{R} \sim (0,1) \times (0,1) \).

**Proposition 5.4** For any sets \( X, Y, Z \):
(a) \( X \sim X \) (reflexivity);
(b) if \( X \sim Y \) then \( Y \sim X \) (symmetry);
(c) if \( X \sim Y \) and \( Y \sim Z \) then \( X \sim Z \) (transitivity).

**Remark 5.5** Properties (a), (b), (c) entail that “\( \sim \)” is an equivalence relation on the collection of all sets. We will see however that the collection of all sets cannot be a set. This means that “\( \sim \)” is not a relation in the sense defined before, as its domain is not a set. Even so, we can still consider the equivalence class \([X]\) of a given set \( X \): the collection of sets which are equivalent to \( X \). For example, the equivalence class \([\mathbb{N}^+]\) of the set of natural numbers \( \mathbb{N}^+ \) is the collection of all countable sets.

**Proof.** (a) The identity function \( i_X : X \to X \) is a bijection.
(b) Suppose \( X \sim Y \). Let \( f : X \to Y \) be a bijection. Hence by Theorem A1.7.6, the inverse function \( f^{-1} : Y \to X \) exists. But now note that \((f^{-1})^{-1} = f : X \to Y \) is an inverse to \( f^{-1} \). So by Theorem A1.7.6 again, \( f^{-1} : Y \to X \) is a bijection. Therefore \( Y \sim X \).
(c) Suppose \( X \sim Y \) and \( Y \sim Z \). Let \( f : X \to Y \) and \( g : Y \to Z \) be bijections. Hence by Theorem A1.7.16 there exist inverse functions \( f^{-1} : Y \to X \) and \( g^{-1} : Z \to Y \). Consider the composition \( g \circ f : X \to Z \). It is easy to check that \((g \circ f)^{-1} = f^{-1} \circ g^{-1} : Z \to X \) is the inverse function of \( g \circ f : X \to Z \). Hence \( g \circ f : X \to Z \) is a bijection. Therefore \( X \sim Z \). \( \square \)

**Exercise 5.6** Suppose \( A \sim B \) and \( C \sim D \). Prove that:
i) \((A \times C) \sim (B \times D)\);
ii) if \( A \) and \( C \) are disjoint and \( B \) and \( D \) are disjoint, then \((A \cup C) \sim (B \cup D)\).

**Proof.** See Exercises 3, Q 5. \( \square \)

**Definition 5.7** Let \( X, Y \) be sets. We say that the cardinality of \( Y \) is greater than or equal to the cardinality of \( X \) (\( Y \) dominates \( X \)), if there exists an injection

\[
f : X \to Y.
\]

In this case, we write \( X \preceq Y \) or \( \text{card}(X) \leq \text{card}(Y) \). We say that the cardinality of \( Y \) is strictly greater then the cardinality of \( X \), if there exists an injection

\[
f : X \to Y,
\]
but there is no injection (and hence no bijection) from \( Y \) to \( X \). In this case, we write \( X \prec Y \) or \( \text{card}(X) < \text{card}(Y) \).

**Remark 5.8** (a) Intuitively, \( X \preceq Y \) (or \( \text{card}(X) \leq \text{card}(Y) \)) means that the set \( X \) has fewer elements than the set \( Y \).

(b) If \( X \sim Y \) then \( X \preceq Y \) and \( Y \preceq X \). To see this, consider a bijection \( f : X \to Y \). Then \( f : X \to Y \) is an injection and \( f^{-1} : Y \to X \) is an injection too.

(c) If \( A \subseteq X \) then the identity map \( i_A : A \to X \) is an injection. Hence if \( A \subseteq X \) then \( A \preceq X \).

(d) If \( X \prec Y \) then \( X \preceq Y \) and \( X \not\sim Y \).

**Example 5.9** \( \emptyset \prec \{1\} \prec \{1,2\} \prec \{3,4\} \prec \{1,2,\ldots,n\} \prec \mathbb{N}^+ \preceq \mathbb{Q} \preceq \mathbb{R} \prec (0,1) \). However \( \{1,2\} \sim \{3,4\}, \mathbb{N}^+ \sim \mathbb{Q} \) and \( \mathbb{R} \sim (0,1) \).

**Exercise 5.10** Prove that for any sets \( X, Y, Z \):

(a) \( X \preceq X \) (reflexivity);

(b) if \( X \preceq Y \) and \( Y \preceq Z \) then \( X \preceq Z \) (transitivity).

**Proof.** Similar to the proof of Proposition 5.4.

**Exercise 5.11** Suppose \( A \preceq B \) and \( C \preceq D \). Prove that:

i) \( (A \times C) \preceq (B \times D) \);

ii) If \( A \) and \( C \) are disjoint and \( B \) and \( D \) are disjoint, then \( (A \cup C) \preceq (B \cup D) \).

**Proof.** Similar to Exercises 3, Q5.

**Remark 5.12** Properties (a) and (b) of Exercise 5.10 mean that \( \preceq \) is a reflexive and transitive relation on the collection of all sets (see Remark 5.5). However \( \preceq \) is not an order relation, because it is not antisymmetric. This means that \( X \preceq Y \) and \( Y \preceq X \) does not imply that \( X = Y \). For example, \( \mathbb{N}^+ \sim \mathbb{Q} \), so \( \mathbb{N}^+ \preceq \mathbb{Q} \) and \( \mathbb{Q} \preceq \mathbb{N}^+ \), but \( \mathbb{N}^+ \neq \mathbb{Q} \).

**Theorem 5.13 (Cantor-Schröder-Bernstein theorem)** Let \( X \) and \( Y \) be sets. If \( X \preceq Y \) and \( Y \preceq X \) then \( X \sim Y \).

**Proof.** See Stewart and Tall, Theorem 6 on p.238.

The Cantor–Schröder–Bernstein theorem is extremely useful when it is difficult to find a bijection between two sets. Instead, it is enough to construct two injections.

**Example 5.14** \([0,1] \sim (0,1)\).

A simple proof via the Cantor–Schröder–Bernstein theorem. Define \( f : (0,1) \to [0,1] \) and \( g : [0,1] \to (0,1) \) by

\[
f(x) := x, \quad g(x) := \frac{1}{2}x + \frac{1}{4}.
\]

It is clear that both \( f \) and \( g \) are injections.
6 The power set and the hierarchy of cardinalities

Definition 6.1 Let $X$ be a set. The power set of $X$, $\mathcal{P}(X)$ or $2^X$, is the set of all subsets of $X$.

Example 6.2 Let $X = \{0, 1\}$. Then $2^X = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$. Notice that $2^\emptyset = \{\emptyset\}$, so $2^\emptyset$ has one element.

Exercise 6.3 Let $X = \{1, 2, \ldots, N\}$. Then the set $2^X$ has $2^N$ elements.

Theorem 6.4 Let $X$ be a set. Then $X \prec 2^X$.

Proof. It is clear that the function $f : X \rightarrow 2^X$ given by $f(x) := \{x\}$ is an injection, so $X \prec 2^X$. We now show $X \not\sim 2^X$.

Assume that $f : X \rightarrow 2^X$ is a bijection. Put

$$Z := \{x \in X | x \notin f(x)\}.$$ 

Then $Z \in 2^X$. As $f$ is a surjection, $f(z) = Z$ for some $z \in X$. Now,

$$z \in Z \Rightarrow z \notin f(z) \Rightarrow z \notin Z,$$

$$z \notin Z \Rightarrow z \notin f(z) \Rightarrow z \in Z,$$

a contradiction. So there is no bijection $f : X \rightarrow 2^X$. \qed

Applied to the set $\mathbb{N}^+$, the theorem says that $2^{\mathbb{N}^+}$ is uncountable. Actually it may be shown that $2^{\mathbb{N}^+}$ is a continuum.

Proposition 6.5 $2^{\mathbb{N}^+} \sim \mathbb{R}$.

Proof. See Stewart and Tall, pp.240-241. \qed

Exercise 6.6 Let $\mathcal{F} = \{f : [0, 1] \rightarrow \mathbb{R}\}$ be the set of all functions from the interval $[0, 1]$ to $\mathbb{R}$. Prove that $[0, 1] \prec \mathcal{F}$.

Hint. Consider the set $\mathcal{F}_0 = \{f_A : A \subset [0, 1]\}$ of all characteristic functions $f_A : [0, 1] \rightarrow \{0, 1\}$ of the form

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

Here, $A \in 2^{[0,1]}$ is a subset of $[0, 1]$. Establish a bijection between the set $\mathcal{F}_0$ and the set $2^{[0,1]}$ of all subsets of the interval $[0, 1]$.

Hierarchy of cardinalities. Theorem 6.4 leads us to a hierarchy of cardinalities. We begin with $\text{card}(\mathbb{N}^+)$. Then $\text{card}(2^{\mathbb{N}^+}) = \text{card}(\mathbb{R})$ is strictly bigger. Then follows $\text{card}(2^\mathbb{R})$, and so on:

$$\emptyset \prec \{1\} \prec \{1, 2\} \prec \cdots \prec \{1, 2, \ldots, N\} \prec \cdots \prec \mathbb{N}^+ \prec 2^{\mathbb{N}^+} \sim \mathbb{R} \prec 2^\mathbb{R} \prec 2^{2^\mathbb{R}} \prec \cdots$$
Continuum Hypothesis. In 1878, Cantor asked whether there exists a set $X$ such that $\mathbb{N}^+ \prec X \prec \mathbb{R}$. He conjectured that the answer was no, but was unable to prove it. This conjecture is known as the Continuum Hypothesis.

Continuum Hypothesis. Let $X$ be a set. If $\mathbb{N}^+ \prec X$ and $X \preceq \mathbb{R}$ then $X \sim \mathbb{R}$.

The Continuum Hypothesis was resolved by Gödel and Cohen in the middle of the 20th century. They proved that in the framework of Axiomatic Set Theory it is impossible to prove the continuum hypothesis and it is also impossible to disprove it: that is, the continuum hypothesis is independent of the axioms. This means that it is possible to construct axiomatic models of set theory in which this hypothesis is true, and models in which it is false. Its status therefore is similar to that of the parallel postulate in Euclidean geometry.

Paradoxes of Naive Set Theory. In 1899 Cantor discovered a paradox which arises if one considers the set of all sets. What is the cardinal number of the set of all sets? Clearly it must be the greatest possible cardinal, yet the cardinal of the set of all subsets of a set always has a greater cardinal than the set itself.

In 1902, Russell discovered another paradox now known as Russell’s paradox. Suppose that we assume the existence of the set $y$ consisting of all sets $x$ which do not belong to themselves, i.e.

$$y = \{x : x \notin x\}.$$  

If $y \in y$ then $y$ must satisfy the defining property of $y$, i.e. $y \notin y$. On the other hand, if $y \notin y$, then $y$ satisfies the defining property of $y$ and then $y \in y$. In either case, a contradiction is obtained.

A popular version of this paradox concerns a certain barber who shaves everyone in his town who does not shave himself. The question is: who shaves the barber? Each answer leads to a contradiction. The conclusion is that there is no such barber.

Similarly, it is meaningless to speak of the set all sets which do not belong to themselves. It is also meaningless to speak of the set all sets. There is no such set.

These paradoxes can be resolved in the framework of modern Axiomatic Set Theory.
Part III
Convergence and continuity

Recommended texts:


7 Subsequences and accumulation points

Definition 7.1 Let \((a_n)_{n \in \mathbb{N}^+}\) be a sequence. A subsequence of \((a_n)_{n \in \mathbb{N}^+}\) is a sequence
\[
(a_{m(k)})_{k \in \mathbb{N}^+} = (a_{m(1)}, a_{m(2)}, a_{m(3)}, \ldots, a_{m(k)}, \ldots),
\]
where \(m: \mathbb{N}^+ \to \mathbb{N}^+\) is a strictly increasing function, so that
\[1 \leq m(1) < m(2) < m(3) < \cdots < m(k) < \ldots.\]

Remark 7.2 (a) The sequence \((a_n)_{n \in \mathbb{N}^+}\) is a subsequence of itself (take \(m(k) = k\)).
(b) The monotonicity of the function \(m\) implies that \((\forall k \in \mathbb{N}^+)(m(k) \geq k)\). So,
\[m(k) \to \infty \quad \text{as} \quad k \to \infty.\]
(c) Any sequence has infinitely many subsequences. Take, for example,
\[m_1(k) = k, \quad m_2(k) = 2k, \quad m_3(k) = 3k, \quad \ldots\]
for \(k \in \mathbb{N}^+\).

Example 7.3 Consider the sequence \((a_n)_{n \in \mathbb{N}^+}\) defined by the formula
\[a_n = (-1)^n.\]
Then
\[
(a_{2k+1})_{k \in \mathbb{N}^+} = (-1, -1, -1, -1 \ldots),
(a_{2k})_{k \in \mathbb{N}^+} = (1, 1, 1, 1 \ldots),
(a_{3k})_{k \in \mathbb{N}^+} = (-1, -1, -1, 1 \ldots)
\]
are subsequences of \((a_n)_{n \in \mathbb{N}^+}\). Note that \(a_{2k+1} \to -1\) and \(a_{2k} \to 1\) as \(k \to \infty\), while \((a_{3k})_{k \in \mathbb{N}^+}\) diverges.

Example 7.4 Let \(X = \{x_1, x_2, \ldots, x_s\} \subseteq \mathbb{R}\) be a finite set of distinct real numbers. Consider the sequence
\[
(a_n)_{n \in \mathbb{N}^+} = (x_1, x_2, \ldots, x_s, x_1, x_2, \ldots, x_s, x_1, x_2, \ldots, x_s, \ldots).
\]
This sequence may be described by

\[ a_n = \begin{cases} 
  x_1 & \text{if } n \equiv 1 \pmod{s}, \\
  x_2 & \text{if } n \equiv 2 \pmod{s}, \\
  \vdots \ & \vdots \\
  x_s & \text{if } n \equiv 0 \pmod{s}.
\end{cases} \]

Then \( a_{sn+1} \to x_1, a_{sn+2} \to x_2, \ldots \), and \( a_{sn} \to x_s \) as \( n \to \infty \). In short, the sequence \( (a_n)_{n \in \mathbb{N}^+} \) possesses subsequences that tend to each element in \( X \).

**Example 7.5** Consider the sequence

\[ (a_n)_{n \in \mathbb{N}^+} = (1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, \ldots). \]

Then for any \( m \in \mathbb{N}^+ \) there is a subsequence

\[ (m, m, m, m, \ldots) \]

of \( (a_n)_{n \in \mathbb{N}^+} \) which tends to \( m \) as \( n \to \infty \).

**Lemma 7.6** Suppose that the sequence \( (a_n)_{n \in \mathbb{N}^+} \) converges to \( a \). Then any subsequence \( (a_{m(k)})_{k \in \mathbb{N}^+} \) also converges to \( a \).

**Proof.** Let \( \varepsilon > 0 \). As \( a_n \to a \) as \( n \to \infty \), we can find \( N \in \mathbb{N}^+ \) such that

\[ |a_n - a| < \varepsilon \quad \text{for } n > N. \]

Now, for \( k > N \), \( m(k) > m(N) \geq N \). Thus

\[ (\forall k \in \mathbb{N}^+) \left( (k > N) \Rightarrow (|a_{m(k)} - a| < \varepsilon) \right). \]

That is, \( a_{m(k)} \to a \) as \( k \to \infty \). \( \square \)

**Lemma 7.7** Let \( (a_n)_{n \in \mathbb{N}^+} \) be a bounded sequence. Then any subsequence is bounded.

**Proof.** Let \( (a_{m(k)})_{k \in \mathbb{N}^+} \) be a subsequence of \( (a_n)_{n \in \mathbb{N}^+} \). As \( (a_n)_{n \in \mathbb{N}^+} \) is bounded, there exists \( 0 \leq M < \infty \) such that

\[ |a_n| \leq M \quad \text{for } n \in \mathbb{N}^+. \]

But then

\[ (\forall k \in \mathbb{N}^+) \left( |a_{m(k)}| \leq M \right), \]

and so \( (a_{m(k)})_{k \in \mathbb{N}^+} \) is bounded. \( \square \)

**Definition 7.8** Let \( (a_n)_{n \in \mathbb{N}^+} \) be a sequence in \( \mathbb{R} \). We say that \( a \in \mathbb{R} \) is an **accumulation point** of \( (a_n)_{n \in \mathbb{N}^+} \) if there exists a subsequence \( (a_{m(k)})_{k \in \mathbb{N}^+} \) which converges to \( a \).

**Example 7.9** (a) Consider the sequence

\[ (-1, 1, -1, 1, -1, 1, -1, 1, \ldots). \]
from Example 7.3. Then the set of accumulation points of \((a_n)_{n \in \mathbb{N}^+}\) is \([-1, 1]\).

(b) The set of accumulation points of the sequence
\[(a_n)_{n \in \mathbb{N}^+} = (x_1, x_2, \ldots, x_s, x_1, x_2, \ldots, x_s, x_1, x_2, \ldots, x_s, \ldots)\]
from Example 7.4 is the set
\[X = \{x_1, x_2, \ldots, x_s\}.
\]

(c) Consider the sequence
\[(a_n)_{n \in \mathbb{N}^+} = (1, 1, 2, 3, 4, 1, 2, 3, 4, 5, \ldots)\]
from Example 7.5. The set of accumulation points of \((a_n)_{n \in \mathbb{N}^+}\) is the set of natural numbers
\[\mathbb{N}^+ = \{1, 2, 3, 4, 5, \ldots\}.
\]

Exercise 7.10 Let \((a_n)_{n \in \mathbb{N}^+}\) be a sequence that converges to \(a\). Then the set of accumulation points of \((a_n)_{n \in \mathbb{N}^+}\) comprises one element, namely \(\{a\}\).

Hint. Apply Lemma 7.6.

Exercise 7.11 Let \((a_n)_{n \in \mathbb{N}^+}\) be a bounded sequence. Then the set of accumulation points of \((a_n)_{n \in \mathbb{N}^+}\) is bounded.

Hint. Apply Lemma 7.7.

Characterisation of the set of all accumulation points of a sequence.

Proposition 7.12 A number \(a \in \mathbb{R}\) is an accumulation point of the sequence \((a_n)_{n \in \mathbb{N}^+}\) if and only if

(1) for any \(\varepsilon > 0\), the set \(\{n \in \mathbb{N}^+: a - \varepsilon < a_n < a + \varepsilon\}\) is infinite;

equivalently,

(2) \((\forall \varepsilon > 0)(\forall N \in \mathbb{N}^+)(\exists n \in \mathbb{N}^+)[(n \geq N) \land (|a_n - a| < \varepsilon)]\).

We also formulate the negation of Proposition 7.12.

Negation of 7.12 A number \(a \in \mathbb{R}\) is not an accumulation point of the sequence \((a_n)_{n \in \mathbb{N}^+}\) if and only if there is an \(\varepsilon > 0\) such that the set \(\{n \in \mathbb{N}^+: a - \varepsilon < a_n < a + \varepsilon\}\) is finite.

Proof. \((\Rightarrow)\) Assume that \(a \in \mathbb{R}\) is an accumulation point of \((a_n)_{n \in \mathbb{N}^+}\). We prove that (2) holds. Since \(a \in \mathbb{R}\) is an accumulation point, there is a subsequence \((a_{m(k)})_{k \in \mathbb{N}^+}\) such that \(\lim_{k \to \infty} a_{m(k)} = a\). Let \(\varepsilon > 0\) and \(N \in \mathbb{N}^+\) be given. By definition of the limit
\[\exists K > 0)(\forall k \in \mathbb{N}^+)[(k \geq K) \Rightarrow (|a_{m(k)} - a| < \varepsilon)]\]

Now choose \(k \in \mathbb{N}^+\) such that \(k \geq \max\{K, N\}\) and put \(n := m(k) (\geq k)\). Then
\[[(n \geq N) \land (|a_n - a| < \varepsilon)].\]
Now assume that (2) holds. Set $\varepsilon = 1$ and $N = 1$. By (2),
\[
(\exists n \in \mathbb{N}^+)[(n \geq N) \land (|a_n - a| < 1)].
\]
Put $m(1) = n$.

Set $\varepsilon = 1/2$ and $N = m(1) + 1$. By (2),
\[
(\exists n \in \mathbb{N}^+)[(n \geq N) \land (|a_n - a| < 1/2)].
\]
Put $m(2) = n$.

Set $\varepsilon = 1/3$ and $N = m(2) + 1$. By (2),
\[
(\exists n \in \mathbb{N}^+)[(n \geq N) \land (|a_n - a| < 1/3)].
\]
Put $m(3) = n$.

Continuing in this way we construct a subsequence $(a_{m(k)})_{k \in \mathbb{N}^+}$ such that
\[
(\forall k \in \mathbb{N}^+)[(a_{m(k)} - a| < 1/k).
\]
In particular, $\lim_{k \to \infty} a_{m(k)} = a$, and $a$ is an accumulation point of $(a_n)_{n \in \mathbb{N}^+}$.

**Example 7.13** The set of accumulation points of the sequence $(a_n)_{n \in \mathbb{N}^+}$ defined by
\[
a_n = (-1)^n + \frac{1}{n}
\]
comprises $\{-1, 1\}$.

**Proof.** As $a_{2k-1} \to -1$ and $a_{2k} \to 1$ as $k \to \infty$,

$-1$ and $1$ are accumulation points of $(a_n)_{n \in \mathbb{N}^+}$.

**Claim:** Each $b \in \mathbb{R} \setminus \{-1, 1\}$ is not an accumulation point of $(a_n)_{n \in \mathbb{N}^+}$.

That is,
\[
(\exists \varepsilon > 0)(\exists N \in \mathbb{N}^+)(\forall n \in \mathbb{N}^+)[(n \geq N) \Rightarrow (|a_n - b| > \varepsilon)].
\]

Put $\varepsilon = (1/2) \min\{|1 - b|, |1 + b|\}$. Then

\[
|a_n - b| = |(-1)^n + 1/n - b| \\
\geq |(-1)^n - b| - 1/n \\
\geq \text{min}\{|1 - b|, |1 + b|\} - 1/n \\
= 2\varepsilon - 1/n
\]

using the fact that $|x - y| \geq ||x| - |y||$ for any $x, y \in \mathbb{R}$.

Choose $N \in \mathbb{N}^+$ such that $N > 1/\varepsilon$ by the AP. Then for any $n \geq N$,
\[
|a_n - b| \geq 2\varepsilon - 1/n \geq 2\varepsilon - 1/N > 2\varepsilon - \varepsilon = \varepsilon.
\]

So $b$ is not an accumulation point. \qed
8 The Bolzano-Weierstrass theorem

Theorem 8.1 (Bolzano-Weierstrass theorem) Any bounded sequence has a convergent subsequence.

Proof. Let \((a_n)_{n \in \mathbb{N}^+}\) be a bounded sequence; say,

\[|a_n| \leq M \text{ for all } n \in \mathbb{N}^+,\]

where \(0 \leq M < \infty\).

Step 1 Set \([b_1, c_1] := [-M, M]\) \[c_1 - b_1 = 2M\]

Step 2 Divide \([b_1, c_1]\) into two equal intervals.
At least one interval contains infinitely many elements of \((a_n)_{n \in \mathbb{N}^+}\).
Denote this by \([b_2, c_2]\). \[b_2, c_2 \subset [b_1, c_1]\]
Denote this by \([b_2, c_2]\). \[c_2 - b_2 = M\]

Step 3 Divide \([b_2, c_2]\) into two equal intervals.
At least one interval contains infinitely many elements of \((a_n)_{n \in \mathbb{N}^+}\).
Denote this by \([b_3, c_3]\). \[b_3, c_3 \subset [b_2, c_2]\]
Denote this by \([b_3, c_3]\). \[c_3 - b_3 = 2^{-1}M\]

......

Step k Divide \([b_{k-1}, c_{k-1}]\) into two equal intervals.
At least one interval contains infinitely many elements of \((a_n)_{n \in \mathbb{N}^+}\).
Denote this by \([b_k, c_k]\). \[b_k, c_k \subset [b_{k-1}, c_{k-1}]\]
Denote this by \([b_k, c_k]\). \[c_k - b_k = 2^{1-k}M\]

......

In summary, we have constructed a sequence of intervals \([b_k, c_k]\) \(k \in \mathbb{N}^+\) such that:

(i) \((\forall k \in \mathbb{N}^+)\) \([b_{k+1}, c_{k+1}] \subset [b_k, c_k]\);

(ii) \(\lim_{k \to \infty} (c_k - b_k) = \lim_{k \to \infty} 2^{1-k}M = 0\).

By the nested intervals theorem (Theorem A 3.2.3), there exists a unique point \(\alpha \in \cap_{k \in \mathbb{N}^+} [b_k, c_k]\);

moreover,

\[(3) \quad \lim_{k \to \infty} b_k = \lim_{k \to \infty} c_k = \alpha.\]

We now construct a subsequence \((a_{m(k)})_{k \in \mathbb{N}^+}\) of the sequence \((a_n)_{n \in \mathbb{N}^+}\).

\(k = 1: \) Set \(a_{m(1)} = a_1\).

\(k = 2: \) Choose \(a_j \in [b_2, c_2]\) with \(j > m(1)\)
(possible, as \([b_2, c_2]\) contains \(\infty\)-ly many elements of \((a_n)_{n \in \mathbb{N}^+}\).
Set \(a_{m(2)} = a_j\).

\(k = 3: \) Choose \(a_j \in [b_3, c_3]\) with \(j > m(2)\)
(possible, as \([b_3, c_3]\) contains \(\infty\)-ly many elements of \((a_n)_{n \in \mathbb{N}^+}\).
Set \(a_{m(3)} = a_j\).

......

Note that

\((\forall k \in \mathbb{N}^+)\) \([b_k \leq a_{m(k)} \leq c_k]\).
Since \( \lim_{k \to \infty} b_k = \lim_{k \to \infty} c_k = \alpha \), by the sandwich rule, we conclude that

\[
\lim_{k \to \infty} a_{m(k)} = \alpha.
\]

Thus \( (a_{m(k)})_{k \in \mathbb{N}^+} \) is a convergent subsequence of \( (a_n)_{n \in \mathbb{N}^+} \). \( \square \)

**Remark 8.2** The following examples show that the Bolzano-Weierstrass theorem does not hold for unbounded sequences:
(a) the sequence \( a_n = n \) diverges to \(+\infty\) and has no convergent subsequences;
(b) the sequence \( a_n = (-n)^n \) diverges and has no convergent subsequences;
(c) the sequence \( a_n = n + (-1)^n n \) diverges but has a convergent subsequence \( a_{2k+1} = 0 \).

**Example 8.3** Consider the sequence \( (a_n)_{n \in \mathbb{N}^+} \) with \( a_n = \cos(n) \). This sequence is bounded, as \(| \cos(n) | \leq 1\) for all \( n \in \mathbb{N}^+ \). The Bolzano-Weierstrass theorem guarantees that \( (a_n)_{n \in \mathbb{N}^+} \) contains a convergent subsequence. It is by no means straightforward, however, to produce such a subsequence.

**Exercise 8.4** Prove that any unbounded sequence has a subsequence that diverges to \(+\infty\) or to \(-\infty\).

*Hint.* Imitate the proof of the Bolzano-Weierstrass theorem by selecting a sequence of unbounded intervals of the form \((-\infty, n]\) or \([n, +\infty)\) that contain infinitely many members of the sequence.
9 Limit superior and limit inferior

**Definition 9.1** Let \((a_n)_{n \in \mathbb{N}^+}\) be a bounded sequence. Let \(\mathcal{A}\) denote the set of accumulation points of \((a_n)_{n \in \mathbb{N}^+}\). The **limit superior** \(\limsup_{n \to \infty} a_n\) of \((a_n)_{n \in \mathbb{N}^+}\) is defined by
\[
\limsup_{n \to \infty} a_n := \sup \mathcal{A}.
\]
The **limit inferior** \(\liminf_{n \to \infty} a_n\) of \((a_n)_{n \in \mathbb{N}^+}\) is defined by
\[
\liminf_{n \to \infty} a_n := \inf \mathcal{A}.
\]

**Remark 9.2** The limit superior and the limit inferior of a bounded sequence always exist. Indeed, let \(\mathcal{A}\) be the set of accumulation points of a bounded sequence \((a_n)_{n \in \mathbb{N}^+}\). By the Bolzano-Weierstrass theorem the set \(\mathcal{A}\) is nonempty, and by Exercise 7.11 the set \(\mathcal{A}\) is bounded. Therefore from the Completeness Axiom and Theorem A 2.4.1 we conclude that both \(\sup \mathcal{A}\) and \(\inf \mathcal{A}\) exist.

**Example 9.3**

a) Let \(a_n = \frac{1}{n}\). This sequence converges to zero. Hence the set of all accumulation points is \(\mathcal{A} = \{0\}\). Therefore
\[
\limsup_{n \to \infty} a_n = \liminf_{n \to \infty} a_n = 0.
\]
b) Let \(a_n = (-1)^n\). The set of all accumulation points of this sequence is \(\mathcal{A} = \{-1, 1\}\). Hence
\[
\limsup_{n \to \infty} a_n = 1, \quad \liminf_{n \to \infty} a_n = -1.
\]
c) Let \(a_n = (-1)^n + \frac{1}{n}\). The set of all accumulation points of this sequence is \(\mathcal{A} = \{-1, 1\}\) (see Example 7.13). Hence
\[
\limsup_{n \to \infty} a_n = 1, \quad \liminf_{n \to \infty} a_n = -1.
\]
d) Let \(a_n = n\). This sequence is unbounded. The limit superior and the limit inferior is undefined for an unbounded sequence.

**Exercise 9.4** Let \((a_n)_{n \in \mathbb{N}^+}\) be a convergent sequence and \(\lim_{n \to \infty} a_n = a\). Prove that
\[
\limsup_{n \to \infty} a_n = \liminf_{n \to \infty} a_n = a.
\]
**Hint.** Use Exercise 7.10.

**Theorem 9.5** Let \((a_n)_{n \in \mathbb{N}^+}\) be a bounded sequence and set
\[
\alpha := \limsup_{n \to \infty} a_n \quad \text{and} \quad \beta := \liminf_{n \to \infty} a_n.
\]
Then \(\alpha\) and \(\beta\) are accumulation points of \((a_n)_{n \in \mathbb{N}^+}\).
Proof. Let $\mathcal{A}$ denote the set of all accumulation points of $(a_n)_{n \in \mathbb{N}^+}$; then $\alpha = \sup \mathcal{A}$. We prove that $\alpha \in \mathcal{A}$. Fix $\varepsilon > 0$ and $N \in \mathbb{N}^+$. Since $\alpha = \sup \mathcal{A}$,

$$\exists a \in \mathcal{A} \forall a > \alpha - \frac{\varepsilon}{2}.$$ 

As $a \in \mathcal{A}$, using Proposition 7.12,

$$\exists k \in \mathbb{N}^+: (k \geq N) \land (|a_k - a| < \frac{\varepsilon}{2}).$$

By the triangle inequality,

$$|\alpha - a_k| = |(\alpha - a) + (a - a_k)| \leq (\alpha - a) + |a - a_k| < \varepsilon.$$

Proposition 7.12 now implies that $\alpha$ is an accumulation point of $(a_n)_{n \in \mathbb{N}^+}$.

Characterisation of the limit superior and limit inferior

Proposition 9.6 Let $(a_n)_{n \in \mathbb{N}^+}$ be a bounded sequence. Then $\alpha = \limsup_{n \to \infty} a_n$ if and only if for any $\varepsilon > 0$

(i) the set $\{n \in \mathbb{N}^+: a_n > \alpha + \varepsilon\}$ is finite, and

(ii) the set $\{n \in \mathbb{N}^+: a_n > \alpha - \varepsilon\}$ is infinite.

Likewise, $\beta = \liminf_{n \to \infty} a_n$ if and only if for any $\varepsilon > 0$

(iii) the set $\{n \in \mathbb{N}^+: a_n < \beta - \varepsilon\}$ is finite, and

(iv) the set $\{n \in \mathbb{N}^+: a_n < \beta + \varepsilon\}$ is infinite.

Proof. We show equivalence with (i) and (ii).

($\Rightarrow$) Let $\mathcal{A}$ be the set of accumulation points of $(a_n)_{n \in \mathbb{N}^+}$ so that $\alpha = \limsup_{n \to \infty} a_n = \sup \mathcal{A}$. Fix $\varepsilon > 0$. Now $\alpha \in \mathcal{A}$ by Theorem 9.5. So there exists a subsequence $(a_{m(k)})_{k \in \mathbb{N}^+}$ such that $\lim_{k \to \infty} a_{m(k)} = \alpha$. By definition of the limit this means that

$$\exists K \in \mathbb{N}^+ \forall k \in \mathbb{N}^+ (k > K \Rightarrow (|a_{m(k)} - \alpha| < \varepsilon).$$

In particular,

$$\forall k > K (a_{m(k)} > \alpha - \varepsilon).$$

So the set $\{n \in \mathbb{N}^+: a_n > \alpha - \varepsilon\}$ is infinite and condition (ii) holds.

Assume, for a contradiction, that (i) does not hold. This means that there exists $\varepsilon > 0$ and a subsequence $(a_{m(k)})_{k \in \mathbb{N}^+}$ such that

$$\forall k \in \mathbb{N}^+ (a_{m(k)} > \alpha + \varepsilon).$$

Since $(a_{m(k)})_{k \in \mathbb{N}^+}$ is bounded, by the Bolzano-Weierstrass theorem and Theorem 7.12 we conclude that $(a_{m(k)})_{k \in \mathbb{N}^+}$ has an accumulation point $\gamma \geq \alpha + \varepsilon$. This contradicts the definition of $\alpha$. 

27
Now let $\alpha \in \mathbb{R}$ be such that for any $\varepsilon > 0$ properties (i) and (ii) hold. We prove that $\alpha = \limsup_{n \to \infty} a_n$. First, we show that $\alpha \in \mathcal{A}$. Indeed, from (i) and (ii), it follows that for any $\varepsilon > 0$ the set $\{n \in \mathbb{N}^+: \alpha - \varepsilon < a_n < \alpha + \varepsilon\}$ is infinite. By Theorem 7.12 we conclude that $\alpha \in \mathcal{A}$.

We now prove that $\alpha = \sup \mathcal{A}$. Assume for a contradiction that there exists $\gamma \in \mathcal{A}$ such that $\gamma > \alpha$. Set $\varepsilon := (1/2)(\gamma - \alpha) > 0$. As $\gamma \in \mathcal{A}$, by Proposition 7.12, the set $\{n \in \mathbb{N}^+: \gamma - \varepsilon < a_n < \gamma + \varepsilon\}$ is infinite. But

$$\gamma - \varepsilon = \gamma - (1/2)(\gamma - \alpha) = \alpha + (1/2)(\gamma - \alpha) = \alpha + \varepsilon.$$

This contradicts (i). We conclude that $\alpha = \sup \mathcal{A}$. \qed

**Theorem 9.7 (Formulae for limit superior and limit inferior)** Let $(a_n)_{n \in \mathbb{N}^+}$ be a bounded sequence. Then

$$\limsup_{n \to \infty} a_n = \lim_{n \to \infty} \left( \sup_{k \geq n} a_k \right),$$

$$\liminf_{n \to \infty} a_n = \lim_{n \to \infty} \left( \inf_{k \geq n} a_k \right).$$

**Remark 9.8** For each $n \in \mathbb{N}^+$, put $c_n := \sup_{k \geq n} a_k$. Then the sequence $(c_n)_{n \in \mathbb{N}^+}$ is decreasing. In fact, for any $n \in \mathbb{N}^+$,

$$c_n = \sup A_n \text{ where } A_n := \{a_k : k \geq n\}.$$

Now, $A_{n+1} \subseteq A_n$. Therefore,

$$c_{n+1} = \sup A_{n+1} \leq \sup A_n = c_n.$$

The sequence $(c_n)_{n \in \mathbb{N}^+}$ is also bounded below. For, as $(a_n)_{n \in \mathbb{N}^+}$ is bounded,

$$-M \leq a_n \leq M \text{ for all } n \in \mathbb{N}^+$$

for some $0 \leq M < \infty$. So

$$c_n = \sup \{a_k : k \geq n\} \geq a_n \geq -M \text{ for any } n \in \mathbb{N}^+.$$

Therefore, by Theorem A 3.2.2, the limit $\lim_{n \to \infty} c_n = \lim_{n \to \infty} \left( \sup_{k \geq n} a_k \right)$ exists in $\mathbb{R}$.

Likewise, the sequence $(b_n)_{n \in \mathbb{N}^+}$ with $b_n := \inf_{k \geq n} a_k$ is increasing and bounded above. By Theorem A 3.2.1, the limit $\lim_{n \to \infty} b_n = \lim_{n \to \infty} \left( \inf_{k \geq n} a_k \right)$ exists in $\mathbb{R}$.

**Proof.** Put $\alpha := \lim_{n \to \infty} \left( \sup_{k \geq n} a_k \right)$. This means that

$$\forall \varepsilon > 0)(\exists N \in \mathbb{N}^+)(\forall n \in \mathbb{N}^+)\left(\{n \geq N\} \Rightarrow (|\alpha - \sup_{k \geq n} a_k| < \varepsilon)\right),$$

or, equivalently,

$$\forall \varepsilon > 0)(\exists N \in \mathbb{N}^+)(\forall n \in \mathbb{N}^+)\left(\{n \geq N\} \Rightarrow (\alpha - \varepsilon < \sup_{k \geq n} a_k < \alpha + \varepsilon)\right).$$

28
Fix \( \varepsilon > 0 \). Choose \( N \in \mathbb{N}^+ \) as above. Then \( \sup_{k \geq N} a_k < \alpha + \varepsilon \). Therefore,

\[
\{ n \geq N : a_n > \alpha + \varepsilon \} = \emptyset.
\]

Consequently, the set

\[
\{ n \in \mathbb{N}^+ : a_n > \alpha + \varepsilon \}
\]

is finite, containing at most \( N - 1 \) elements. On the other hand, in view of (7),

\[
(\forall n \geq N)(\exists k \geq n)[a_k \geq \alpha - \varepsilon].
\]

So the set

\[
\{ n \in \mathbb{N}^+ : a_n > \alpha - \varepsilon \}
\]

is infinite. By Proposition 9.6 we conclude that \( \alpha = \limsup_{n \to \infty} a_n \). The argument in the case of the limit inferior is similar.

**Example 9.9** Let \( a_n = (-1)^n + \frac{1}{n} \). Find \( \limsup_{n \to \infty} a_n \) and \( \liminf_{n \to \infty} a_n \) by using (4) and (5).

**Solution.** Set \( A_n := \sup_{k \geq n} a_k \) and \( B_n := \inf_{k \geq n} a_k \). Then we obtain

\[
A_n = \sup\{a_n, a_{n+1}, a_{n+2}, a_{n+3}, \ldots\} = \begin{cases} 1 + \frac{1}{n} & \text{if } n \text{ is even}, \\ 1 + \frac{1}{n+1} & \text{if } n \text{ is odd}. \end{cases}
\]

So \( \limsup_{n \to \infty} a_n = \lim_{n \to \infty} A_n = 1 \). Similarly,

\[
B_n = \inf\{a_n, a_{n+1}, a_{n+2}, a_{n+3}, \ldots\} = -1.
\]

So \( \liminf_{n \to \infty} a_n = \lim_{n \to \infty} B_n = -1 \).

**Exercise 9.10** (a) Give an example of bounded sequences \( (a_n)_{n \in \mathbb{N}^+} \) and \( (b_n)_{n \in \mathbb{N}^+} \) such that

\[
\limsup_{n \to \infty} (a_n + b_n) \neq \limsup_{n \to \infty} a_n + \limsup_{n \to \infty} b_n.
\]

(b) Prove that for any bounded sequences \( (a_n)_{n \in \mathbb{N}^+} \) and \( (b_n)_{n \in \mathbb{N}^+} \) one has

\[
\limsup_{n \to \infty} (a_n + b_n) \leq \limsup_{n \to \infty} a_n + \limsup_{n \to \infty} b_n,
\]

\[
\liminf_{n \to \infty} (a_n + b_n) \geq \liminf_{n \to \infty} a_n + \liminf_{n \to \infty} b_n.
\]

## 10 Cauchy sequences

**Definition 10.1** A sequence \( (a_n)_{n \in \mathbb{N}^+} \) is called a **Cauchy sequence** if

\[
(\forall \varepsilon > 0)(\exists N \in \mathbb{N}^+)(\forall m, n \in \mathbb{N}^+)[(m \geq N) \land (n \geq N) \Rightarrow (|a_m - a_n| < \varepsilon)].
\]

**Example 10.2** The sequence \( (a_n)_{n \in \mathbb{N}^+} \) with \( a_n = \frac{1}{n} \) is a Cauchy sequence.
Proof. Fix $\varepsilon > 0$ and take $N \in \mathbb{N}^+$ such that $N > \frac{1}{\varepsilon}$. Let $m, n \geq N$. Then

$$|a_m - a_n| = |\frac{1}{m} - \frac{1}{n}| \leq \max\left\{\frac{1}{m}, \frac{1}{n}\right\} \leq \frac{1}{N} < \varepsilon.$$ 

So $(a_n)_{n \in \mathbb{N}^+}$ is a Cauchy sequence.

Example 10.3 The sequence $(a_n)_{n \in \mathbb{N}^+}$ with $a_n = \frac{n^2 + 1}{n^2}$ is a Cauchy sequence.

Proof. Fix $\varepsilon > 0$ and take $N \in \mathbb{N}^+$ such that $N > \frac{1}{\sqrt{\varepsilon}}$. Let $m, n \geq N$. Then

$$|a_m - a_n| = \left|\frac{m^2 + 1}{m^2} - \frac{n^2 + 1}{n^2}\right| = \left|\frac{1}{m^2} - \frac{1}{n^2}\right| \leq \max\left\{\frac{1}{m^2}, \frac{1}{n^2}\right\} \leq \frac{1}{N^2} < \varepsilon.$$ 

This means that $(a_n)_{n \in \mathbb{N}^+}$ is a Cauchy sequence.

Theorem 10.4 Let $(a_n)_{n \in \mathbb{N}^+}$ be a convergent sequence. Then $(a_n)_{n \in \mathbb{N}^+}$ is a Cauchy sequence.

Proof. Let $(a_n)_{n \in \mathbb{N}^+}$ be a convergent sequence and $a = \lim_{n \to \infty} a_n$. Fix $\varepsilon > 0$. Then

$$(\exists N \in \mathbb{N}^+)(\forall n \in \mathbb{N}^+)((n \geq N) \Rightarrow (|a_n - a| < \frac{\varepsilon}{2})).$$

Let $m, n \geq N$. Then by the triangle inequality,

$$|a_m - a_n| \leq |a_m - a| + |a - a_n| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$ 

That is, $(a_n)_{n \in \mathbb{N}^+}$ is a Cauchy sequence.

In fact, the converse to Theorem 10.4 is also true.

Theorem 10.5 (Cauchy’s theorem: general principle of convergence) Let $(a_n)_{n \in \mathbb{N}^+}$ be a Cauchy sequence. Then $(a_n)_{n \in \mathbb{N}^+}$ converges.

Proof. Let $(a_n)_{n \in \mathbb{N}^+}$ be a Cauchy sequence.

Claim. $(a_n)_{n \in \mathbb{N}^+}$ is bounded.

Fix $\varepsilon = 1$. Then

$$(\exists N \in \mathbb{N}^+)(\forall m, n \in \mathbb{N}^+)((m \geq N) \land (n \geq N) \Rightarrow (|a_m - a_n| < 1)).$$

Let $N^+ \ni n \geq N$. By the triangle inequality,

$$|a_n| \leq |a_n - a_N| + |a_N| \leq 1 + |a_N|.$$ 

Put $M := \max\{|a_1|, |a_2|, \ldots, |a_N|\}$. Then

$$|a_n| \leq M + 1 \text{ for all } n \in \mathbb{N}^+.$$ 

Therefore, $(a_n)_{n \in \mathbb{N}^+}$ is bounded.

Claim. $(a_n)_{n \in \mathbb{N}^+}$ converges.
Since \((a_n)_{n \in \mathbb{N}^+}\) is bounded, by the Bolzano-Weierstrass theorem, \((a_n)_{n \in \mathbb{N}^+}\) has a convergent subsequence \((a_m(k))_{k \in \mathbb{N}^+}\) with limit \(a\), say. We prove that \((a_n)_{n \in \mathbb{N}^+}\) also converges to \(a\). Fix \(\varepsilon > 0\). Then

\[
(\exists N \in \mathbb{N}^+)(\forall m, n \in \mathbb{N}^+)[(m \geq N) \land (n \geq N) \Rightarrow (|a_m - a_n| < \frac{\varepsilon}{2})],
\]

and

\[
(\exists K \in \mathbb{N}^+)(\forall k \in \mathbb{N}^+)[(k \geq K) \Rightarrow (|a_m(k) - a| < \frac{\varepsilon}{2})].
\]

Choose \(l \in \mathbb{N}^+\) such that \(m(l) \geq l \geq \max\{N, K\}\). Let \(\mathbb{N}^+ \ni n \geq N\). Then

\[
|a_n - a| \leq |a_n - a_m(l)| + |a_m(l) - a| < \varepsilon.
\]

In short, \(\lim_{n \to \infty} a_n = a\).

\[\square\]

**Remark 10.6** Combining Theorems 10.4 and 10.5 we see that

a sequence \((a_n)_{n \in \mathbb{N}^+}\) converges if and only if \((a_n)_{n \in \mathbb{N}^+}\) is a Cauchy sequence.

This convergence criterion is intrinsic in the sense that it characterises convergence in terms of a property of the elements of the sequence.

**Remark 10.7** Theorem 10.5 may be useful when we need to show that a sequence diverges. Indeed, combining Theorems 10.4 and 10.5 we see that

a sequence \((a_n)_{n \in \mathbb{N}^+}\) diverges if and only if \((a_n)_{n \in \mathbb{N}^+}\) is not a Cauchy sequence.

**Example 10.8** Prove that the sequence \(a_n = (-1)^n\) diverges.

**Solution.** Let \(N \in \mathbb{N}^+\). Then for \(\mathbb{N}^+ \ni m \geq N\) we have

\[
|a_m - a_{m+1}| = 2.
\]

So, if we fix \(\varepsilon := 1\) (say), and choose \(n := m + 1\), then

\[
|a_m - a_n| > \varepsilon.
\]

So \((a_n)_{n \in \mathbb{N}^+}\) is not a Cauchy sequence and therefore \((a_n)_{n \in \mathbb{N}^+}\) diverges.

\[\square\]

## 11 Uniformly continuous functions

**Definition 11.1** Let \(A \subseteq \mathbb{R}\) and \(f : A \to \mathbb{R}\) a function. Then \(f\) is said to be **uniformly continuous** on \(A\) if

\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x, y \in A)[(|x - y| < \delta) \Rightarrow (|f(x) - f(y)| < \varepsilon)].
\]

In the above definition, \(\delta = \delta(\varepsilon)\) depends on \(\varepsilon\) but not on \(x\), nor \(y\).
Lemma 11.2 Let \( f : A \to \mathbb{R} \) be a uniformly continuous function on \( A \subseteq \mathbb{R} \). Then \( f \) is continuous on \( A \).

**Proof.** We need to prove that \( f \) is continuous at every point in \( A \). Fix \( a \in A \). From (8),
\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x, y \in A)[(|x - y| < \delta) \Rightarrow (|f(x) - f(y)| < \varepsilon)].
\]
In particular,
\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall y \in A)[(|x - a| < \delta) \Rightarrow (|f(x) - f(a)| < \varepsilon)].
\]
This means that \( f \) is continuous at \( a \). As this argument works for any \( a \in A \), \( f \) is continuous on \( A \).

Example 11.3 Let \(-\infty < a < b < +\infty\). Then the function \( f(x) = x \) is uniformly continuous on \([a, b] \).

**Solution.** Let \( \varepsilon > 0 \). Choose \( \delta = \varepsilon \). Then for any \( x, y \in [a, b] \),
\[
(|x - y| < \delta) \Rightarrow (|f(x) - f(y)| = |x - y| < \delta = \varepsilon).
\]
This means that \( f \) is uniformly continuous on \([0, 1]\).

Example 11.4 Let \(-\infty < a < b < +\infty\). Then the function \( f(x) = x^2 \) is uniformly continuous on \([a, b] \).

**Solution.** Set \( c := \max\{|a|, |b|\} \). Note that
\[
|x + y| \leq |x| + |y| \leq 2c
\]
for any \( x, y \in [a, b] \). Hence,
\[
|x^2 - y^2| = |x + y||x - y| \leq 2c|x - y|
\]
for any \( x, y \in [a, b] \). Let \( \varepsilon > 0 \). Choose \( \delta = \varepsilon/2c \). Then, for any \( x, y \in [a, b] \),
\[
(|x - y| < \delta) \Rightarrow (|f(x) - f(y)| = |x^2 - y^2| = |x + y||x - y| < 2c \delta = \varepsilon).
\]
This means that \( f \) is uniformly continuous on \([a, b]\).

From (11.1), the function \( f : A \to \mathbb{R} \) is **not** uniformly continuous on \( A \) if
\[
(\exists \varepsilon > 0)(\forall \delta > 0)(\exists x, y \in A)[(|x - y| < \delta) \wedge (|f(x) - f(y)| \geq \varepsilon)].
\]
Alternatively,

**Negation of Definition 11.1.** A function \( f : A \to \mathbb{R} \) is **not** uniformly continuous on \( A \) if
\[
(\exists \varepsilon > 0)(\forall n \in \mathbb{N}^+)(\exists x_n, y_n \in A)[(|x_n - y_n| < 1/n) \wedge (|f(x_n) - f(y_n)| \geq \varepsilon)].
\]

**Exercise 11.5** Prove that (10) is equivalent to (9).
Example 11.6 The function \( f(x) = 1/x \) is continuous on \((0, 1)\) but is not uniformly continuous on \((0, 1)\).

Solution. The fact that \( f(x) = 1/x \) is continuous on \((0, 1)\) follows from Theorem A 4.2.1. We prove that \( f \) is not uniformly continuous on \((0, 1)\).

Fix \( \varepsilon = 1. \) For each \( n \in \mathbb{N}^+ \), choose \( x_n = 1/2n \) and \( y_n = 1/4n \). Then, for any \( n \in \mathbb{N}^+ \),

\[
|x_n - y_n| = 1/4n < 1/n,
\]

and

\[
|f(x_n) - f(y_n)| = \left| \frac{1}{x_n} - \frac{1}{y_n} \right| = |2n - 4n| = 2n > \varepsilon = 1.
\]

By (10), \( f \) is not uniformly continuous on \((0, 1)\).

Example 11.7 The function \( f(x) = x^2 \) is not uniformly continuous on \((0, \infty)\).

Solution. Fix \( \varepsilon = 1. \) Then for each \( n \in \mathbb{N}^+ \), choose \( x_n = n \) and \( y_n = n + 1/2n \). Then

\[
|x_n - y_n| = 1/2n < 1/n
\]

and

\[
|f(x_n) - f(y_n)| = \left| n^2 - \left( n + \frac{1}{2n} \right)^2 \right| = 1 + \frac{1}{4n^2} > 1 = \varepsilon.
\]

By (10), \( f \) is not uniformly continuous on \((0, \infty)\).

The following is a fundamental theorem in Analysis.

**Theorem 11.8 (Cantor’s theorem on uniform continuity)** Let \(-\infty < a < b < +\infty\). Let \( f : [a,b] \to \mathbb{R} \) be a continuous function on \([a,b]\). Then \( f \) is uniformly continuous on \([a,b]\).

**Proof.** Assume for a contradiction that \( f \) is not uniformly continuous on \([a,b]\). Then, according to (10),

\[
(\exists \varepsilon > 0)(\forall n \in \mathbb{N}^+)(\exists x_n, y_n \in [a,b])[(|x_n - y_n| < 1/n) \land (|f(x_n) - f(y_n)| \geq \varepsilon)].
\]

Observe that the sequence \((x_n)_{n \in \mathbb{N}^+} \subseteq [a,b] \) is bounded. By the Bolzano-Weierstrass theorem, we may extract a convergent subsequence \((x_{m(k)})_{k \in \mathbb{N}^+} \) with limit \( x_0 \) (say) in \([a,b]\). Now,

\[
|x_{m(k)} - y_{m(k)}| < 1/m(k) \to 0 \quad \text{as} \quad k \to \infty.
\]

Also, by the triangle inequality,

\[
|y_{m(k)} - x_0| \leq |y_{m(k)} - x_{m(k)}| + |x_{m(k)} - x_0| \to 0 \quad \text{as} \quad k \to \infty.
\]

In short,

\[
y_{m(k)} \to x_0 \quad \text{as} \quad k \to \infty.
\]

Now, by hypothesis, \( f \) is continuous at \( x_0 \). Consequently,

\[
f(x_{m(k)}) \to f(x_0) \quad \text{and} \quad f(y_{m(k)}) \to f(x_0) \quad \text{as} \quad k \to \infty.
\]
Moreover, by the triangle inequality,
\[ |f(x_{m(k)}) - f(y_{m(k)})| \leq |f(x_{m(k)}) - f(x_0)| + |f(x_0) - f(y_{m(k)})| \rightarrow 0 \quad \text{as} \quad k \rightarrow \infty. \]

On the other hand, (11) tells us that
\[ |f(x_{m(k)}) - f(y_{m(k)})| \geq \varepsilon \]
for each \( k \in \mathbb{N}^+ \). In this way, we arrive at a contradiction. We conclude that \( f \) is uniformly continuous on \([a, b]\). \qed

**Example 11.9** The function \( f(x) = \sqrt{x} \) is uniformly continuous on \([0, 1]\).

**Solution.** \( f \) is continuous on \([0, 1]\). Now use Theorem 11.8. \qed
12 Pointwise and uniform convergence

**Definition 12.1** Let \( A \subseteq \mathbb{R} \). Let \( f \) and \((f_n)_{n \in \mathbb{N}^+}\) be \( \mathbb{R} \)-valued functions on \( A \). We say that \((f_n)_{n \in \mathbb{N}^+}\) **converges pointwise** to \( f \) on \( A \) if

\[
(\forall x \in A)(\forall \varepsilon > 0)(\exists N \in \mathbb{N}^+)(\forall n \in \mathbb{N}^+)(n \geq N) \Rightarrow (|f_n(x) - f(x)| < \varepsilon);
\]

in other words,

\[
(\forall x \in A)[\lim_{n \to \infty} f_n(x) = f(x)].
\]

**Remark 12.2** In this definition, \( N = N(\varepsilon, x) \) depends on both \( \varepsilon \) and \( x \).

**Definition 12.3** Let \( A \subseteq \mathbb{R} \). Let \( f \) and \((f_n)_{n \in \mathbb{N}^+}\) be \( \mathbb{R} \)-valued functions on \( A \). We say that \((f_n)_{n \in \mathbb{N}^+}\) **converges uniformly** to \( f \) on \( A \) if

\[
(\forall \varepsilon > 0)(\exists N \in \mathbb{N}^+)(\forall x \in A)(\forall n \in \mathbb{N}^+)(n \geq N) \Rightarrow (|f_n(x) - f(x)| < \varepsilon);
\]

in other words,

\[
(\forall x \in A)[\lim_{n \to \infty} f_n(x) = f(x)]
\]

**Remark 12.4** In this definition, \( N \) depends on \( \varepsilon \) only, and not on \( x \).

**Remark 12.5** If \((f_n)_{n \in \mathbb{N}^+}\) converges uniformly to \( f \) on \( A \) then \((f_n)_{n \in \mathbb{N}^+}\) converges pointwise to \( f \) on \( A \). But the converse is not true in general (see Example 12.6 below).

**Negation of Definition 12.3** A sequence of functions \((f_n)_{n \in \mathbb{N}^+}\) does **not** converge uniformly to \( f \) on \( A \) if

\[
(\exists \varepsilon > 0)(\forall N \in \mathbb{N}^+)(\exists x \in A)(\exists n \in \mathbb{N}^+)(n \geq N) \land (|f_n(x) - f(x)| \geq \varepsilon).
\]

**Example 12.6** Let

\[
f_n : [0, 1] \to \mathbb{R}; x \to x^n \quad (n \in \mathbb{N}^+).
\]

Then \((f_n)_{n \in \mathbb{N}^+}\) converges pointwise on \([0, 1]\) to \( f : [0, 1] \to \mathbb{R} \) given by

\[
f(x) = \begin{cases} 
0 & \text{if } x \in [0, 1), \\
1 & \text{if } x = 1.
\end{cases}
\]

Note that \( f \) is discontinuous. Moreover, for any \( a \in (0, 1) \), \((f_n)_{n \in \mathbb{N}^+}\) converges uniformly to \( f \) on \([0, a] \). But \((f_n)_{n \in \mathbb{N}^+}\) does not converge uniformly to \( f \) on \([0, 1]\).

**Solution.** First, note that

\[
\lim_{n \to \infty} x^n = \begin{cases} 
0 & \text{if } x \in [0, 1), \\
1 & \text{if } x = 1.
\end{cases}
\]

This means that \((f_n)_{n \in \mathbb{N}^+}\) converges pointwise to \( f \) on \([0, 1]\).

Let \( a \in (0, 1) \). We now show that \((f_n)_{n \in \mathbb{N}^+}\) converges uniformly to \( f \) on \([0, a] \). Fix \( \varepsilon > 0 \). Choose \( N \in \mathbb{N}^+ \) such that \( a^N < \varepsilon \). Then, for any \( x \in [0, a] \) and \( n \geq N \) we have

\[
|f_n(x) - f(x)| = |x^n - 0| \leq a^n \leq a^N < \varepsilon.
\]

This means that the sequence \((f_n)_{n \in \mathbb{N}^+}\) converges uniformly to \( f \) on \([0, a] \).

Lastly, we show that \((f_n)_{n \in \mathbb{N}^+}\) does not converge uniformly to \( f \) on \([0, 1]\). According to (14), we need to verify that

\[
(\exists \varepsilon > 0)(\forall N \in \mathbb{N}^+)(\exists x \in [0, 1])(\exists n \in \mathbb{N}^+)(n \geq N) \land (|f_n(x) - f(x)| \geq \varepsilon).
\]
12  POINTWISE AND UNIFORM CONVERGENCE

Let \( \varepsilon = 1/2 \). Fix \( N \in \mathbb{N}^+ \). Set
\[
x := 2^{-1/N}
\]
and \( \mathbb{N}^+ \ni n := N \). Note that \( x \in [0, 1] \). Then \( n \geq N \) and
\[
|f_n(x) - f(x)| = |x^n - 0| = 1/2 = \varepsilon \geq \varepsilon.
\]
Thus \( (f_n)_{n \in \mathbb{N}^+} \) does not converges uniformly to \( f \) on \([0, 1] \).
\( \square \)

**Theorem 12.7 (Weierstrass’s theorem on uniform convergence)** Let \( A \subseteq \mathbb{R} \). Let \( (f_n)_{n \in \mathbb{N}^+} \) be a sequence of continuous \( \mathbb{R} \)-valued functions on \( A \). Let \( f \) be an \( \mathbb{R} \)-valued function on \( A \). Assume that \( (f_n)_{n \in \mathbb{N}^+} \) converges uniformly to \( f \) on \( A \). Then \( f \) is continuous on \( A \).

**Proof.** Fix \( a \in A \). We show that \( f \) is continuous at \( a \). Let \( \varepsilon > 0 \). From the definition of uniform convergence, there exists \( N \in \mathbb{N}^+ \) such that
\[
(15) \quad (\forall x \in A)(\forall n \in \mathbb{N}^+)[(n \geq N) \Rightarrow (|f_n(x) - f(x)| < \varepsilon/3)].
\]
Now, \( f_N \) is continuous at \( a \). So there exists \( \delta > 0 \) with the property that
\[
(16) \quad (\forall x \in A)[(|x - a| < \delta) \Rightarrow (|f_N(x) - f_N(a)| < \varepsilon/3)].
\]
Suppose that \( x \in A \) with \( |x - a| < \delta \). By the triangle inequality,
\[
|f(x) - f(a)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)| < \varepsilon.
\]
This shows that \( f \) is continuous at \( a \). \( \square \)

**Example 12.8 (Example 12.6 revisited)** Let
\[
f_n : [0, 1] \to \mathbb{R}; \quad x \to x^n \quad (n \in \mathbb{N}^+).
\]
Then \( (f_n)_{n \in \mathbb{N}^+} \) does not converge uniformly to \( f \) on \([0, 1] \) (with \( f \) as before).

**Solution.** \( (f_n)_{n \in \mathbb{N}^+} \) constitutes a sequence of continuous \( \mathbb{R} \)-valued functions on \([0, 1] \). Assume that \( (f_n)_{n \in \mathbb{N}^+} \) converges uniformly to \( f \) on \([0, 1] \). By Theorem 12.7, the limit function \( f \) must be continuous. But \( f \) is not continuous. It follows that \( (f_n)_{n \in \mathbb{N}^+} \) cannot converge uniformly to \( f \) on \([0, 1] \). \( \square \)

**Example 12.9** Let
\[
f_n : [0, +\infty) \to \mathbb{R}; \quad x \to \frac{2nx}{1 + n^2x^2} \quad (n \in \mathbb{N}^+).
\]
Then \( (f_n)_{n \in \mathbb{N}^+} \) converges pointwise to 0 on \([0, +\infty) \) but does not converge uniformly to 0 there.

**Solution.** If \( x = 0 \) then \( f_n(x) = 0 \) for each \( n \in \mathbb{N}^+ \). For \( x \in (0, \infty) \),
\[
f_n(x) = \frac{2x/n}{1/n^2 + x^2} \quad (n \in \mathbb{N}^+).
\]
Therefore, \( f_n(x) \to 0 \) as \( n \to \infty \). In sum, \( (f_n)_{n \in \mathbb{N}^+} \) converges pointwise to 0 on \([0, +\infty)\).

Now,

\[
f'_n(x) = \frac{2n(1 - n^2x^2)}{(1 + n^2x^2)^2} \quad (x \in [0, \infty)) \quad (n \in \mathbb{N}^+).
\]

This indicates that \( f_n \) has a maximum at \( x = 1/n \). In fact,

\[
f_n(1/n) = 1 \quad (n \in \mathbb{N}^+).
\]

This means that the convergence cannot be uniform. For, in (14), choose

\[
\epsilon := 1/2, \quad x := 1/N \quad \text{and} \quad n := N.
\]

\[\square\]

**Series of functions.** Let \( A \subseteq \mathbb{R} \) and \( (f_n)_{n \in \mathbb{N}^+} \) a sequence of \( \mathbb{R} \)-valued functions on \( A \). Consider the partial sum

\[
s_n(x) := \sum_{k=0}^{n} f_k(x) \quad (x \in A) \quad (n \in \mathbb{N}^+).
\]

We say that the series of functions

\[
\sum_{k=0}^{\infty} f_k(x)
\]

is **pointwise convergent** on \( A \) if \( (s_n)_{n \in \mathbb{N}^+} \) converges pointwise on \( A \) to a function \( s : A \to \mathbb{R} \). The function \( s \) is called the **sum** of the series. The series is said to be **uniformly convergent** if the sequence of partial sums \( (s_n)_{n \in \mathbb{N}^+} \) converges uniformly on \( A \).

**Example 12.10** The **geometric series**

\[
\sum_{k=0}^{\infty} x^k
\]

converges pointwise to

\[
s : (-1, 1) \to \mathbb{R} ; \quad x \to \frac{1}{1-x}
\]

on \((-1, 1)\). For any \( r \in (0,1) \), the series converges uniformly to \( s \) on \([-r,r]\).

**Example 12.11** The **exponential series**

\[
\sum_{k=0}^{\infty} \frac{x^k}{k!}
\]

converges pointwise on \( \mathbb{R} \). Its sum \( s \) is called the **exponential function** and denoted by \( \exp(x) \) or \( e^x \). For any \( r > 0 \), the series converges uniformly on \([-r,r]\).
Part IV
The Riemann integral

Recommended texts:


13 Definition of the integral

Definition 13.1 Let \(-\infty < a < b < \infty\). A partition \(P\) of \([a, b]\) is a finite collection of points

\[ P = \{x_0, x_1, \ldots, x_n\} \subseteq [a, b] \]

such that

\[ a = x_0 < x_1 < \cdots < x_n = b. \]

The length of the longest subinterval in \(P\) is called the norm of \(P\) and denoted \(\|P\|\). Thus,

\[ \|P\| = \max_{1 \leq i \leq n} (x_i - x_{i-1}). \]

Example 13.2 Let \(n \in \mathbb{N}^+\). Define a partition \(P_n\) of \([0, 1]\) via

\[ P_n := \{0, 1/n, 2/n, \ldots, (n - 1)/n, 1\}. \]

\(P_n\) is said to be a uniform partition of \([0, 1]\). This is because all subintervals of the partition have equal length. Note that \(\|P_n\| = 1/n\).

Definition 13.3 Let \(f : [a, b] \to \mathbb{R}\) be a bounded function. Let \(P = \{x_0, \ldots, x_n\}\) be a partition of \([a, b]\). Set

\[ m_i = \inf\{f(x) : x_{i-1} \leq x \leq x_i\}, \quad M_i = \sup\{f(x) : x_{i-1} \leq x \leq x_i\}. \]

The lower sum \(L(f, P)\) of \(f\) with respect to the partition \(P\) is defined by

\[ L(f, P) := \sum_{i=1}^{n} m_i(x_i - x_{i-1}). \]

The upper sum \(U(f, P)\) of \(f\) with respect to the partition \(P\) is defined by

\[ U(f, P) := \sum_{i=1}^{n} M_i(x_i - x_{i-1}). \]
Remark 13.4 Put
\[ m := \inf_{x \in [a, b]} f(x), \quad M := \sup_{x \in [a, b]} f(x). \]
Then \( m \leq m_i \leq M_i \leq M \) for each \( i \in \{1, 2, \ldots, n\} \). Therefore,
\[
m(b - a) = \sum_{i=1}^{n} m(x_i - x_{i-1}) \leq \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \leq \sum_{i=1}^{n} M_i(x_i - x_{i-1}) = M(b - a).
\]
This implies that for any partition \( P \) of \([a, b]\) the following inequality holds:
\[
m(b - a) \leq L(f, P) \leq U(f, P) \leq M(b - a).
\]
In other words, the sets
\[
\{ L(f, P) : P \text{ is a partition of } [a, b] \},
\]
\[
\{ U(f, P) : P \text{ is a partition of } [a, b] \}
\]
in \( \mathbb{R} \) are bounded (and non-empty). This means that their infimum and supremum exist.

Definition 13.5 The upper integral of \( f \) over \([a, b]\) is defined by
\[
J := \inf_P U(f, P),
\]
and the lower integral of \( f \) over \([a, b]\) is defined by
\[
j := \sup_P L(f, P),
\]
where the infimum and supremum are taken over all partitions \( P \) of the interval \([a, b]\).

Definition 13.6 A function \( f : [a, b] \to \mathbb{R} \) is said to be Riemann integrable over \([a, b]\) if
\[
J = j.
\]
In this case, the common value is called the Riemann integral of \( f \) over \([a, b]\) and denoted by \( \int_a^b f(x) \, dx \). Thus,
\[
\int_a^b f(x) \, dx = J = j.
\]

Example 13.7 Let \( f : [a, b] \to \mathbb{R} ; \ x \mapsto c \) be a constant function \((c \in \mathbb{R})\). For any partition \( P \) of \([a, b]\) we have
\[
L(f, P) = U(f, P) = c(b - a).
\]
Thus, \( J = j = c(b - a) \). We conclude that \( f \) is Riemann integrable over \([a, b]\) and
\[
\int_a^b f(x) \, dx = c(b - a).
\]

Remark 13.8 If \( j \neq J \) we say that \( f \) is not Riemann integrable.
Example 13.9 (Example of a function which is not Riemann integrable) The Dirichlet function $D : [0, 1] \to \mathbb{R}$ is defined by

$$D(x) = \begin{cases} 1 & \text{if } x \text{ is rational,} \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

It is everywhere discontinuous. For any partition $P$ of $[0, 1]$, we have

$$L(D, P) = 0 \quad \text{and} \quad U(D, P) = 1.$$ 

Thus $j = 0$ and $J = 1$. Hence the Dirichlet function $D$ is not Riemann integrable.

Example 13.10 The function $f(x) = x^2$ is Riemann integrable over $[0, 1]$ and

$$\int_0^1 f(x) \, dx = 1/3.$$ 

Solution. Let $P_n$ be the uniform partition of $[0, 1]$,

$$0 < \frac{1}{n} < \frac{2}{n} < \cdots < \frac{n-1}{n} < 1.$$ 

Now $f(x) = x^2$ is a strictly increasing function on $[0, 1]$. So,

$$m_i = \inf_{[x_{i-1}, x_i]} f(x) = f(x_{i-1}) = \frac{(i-1)^2}{n^2}$$

and

$$M_i = \sup_{[x_{i-1}, x_i]} f(x) = f(x_i) = \frac{i^2}{n^2}.$$ 

Therefore,

$$L(f, P_n) = \sum_{i=1}^{n} \frac{(i-1)^2}{n^2} \frac{1}{n} = 1 \frac{n-1}{n^3} \sum_{i=0}^{n-1} i^2$$

and

$$U(f, P_n) = \sum_{i=1}^{n} \frac{i^2}{n^2} \frac{1}{n} = \frac{1}{n^3} \sum_{i=1}^{n} i^2.$$ 

Recall that

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$ 

(This may be proved by induction.) We then calculate

$$L(f, P_n) = \frac{n(n-1)(2n-1)}{6n^3} = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2},$$ 

$$U(f, P_n) = \frac{n(n+1)(2n+1)}{6n^3} = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}.$$ 

Therefore

$$\lim_{n \to \infty} L(f, P_n) = \frac{1}{3}, \quad \lim_{n \to \infty} U(f, P_n) = \frac{1}{3}.$$
13 DEFINITION OF THE INTEGRAL

and

(18) \[ \lim_{n \to \infty} [U(f, P_n) - L(f, P_n)] = 0. \]

Note. The above computations take into account only uniform partitions \( P_n \) of \([0, 1]\). The definition of the Riemann integral, however, involves all partitions of \([0, 1]\). But, by Corollary 14.2 below, (18) and (17) together imply that the function \( f(x) = x^2 \) is Riemann integrable over \([0, 1]\) with

\[ j = \lim_{n \to \infty} L(f, P_n) = \lim_{n \to \infty} U(f, P_n) = J, \]

and

\[ \int_0^1 x^2 \, dx = j = J = \frac{1}{3}. \]

This completes the argument. \( \square \)

Properties of upper and lower sums.

Proposition 13.11 Let \( f : [a, b] \to \mathbb{R} \) be a bounded function and \( P, Q \) partitions of \([a, b]\) such that \( P \subseteq Q \). Then

\[ L(f, P) \leq L(f, Q), \quad U(f, P) \geq U(f, Q). \]

Remark 13.12 If \( P \subseteq Q \), we say that \( Q \) is a refinement of \( P \).

Proof. First, let \( P' \) be a partition formed from \( P \) by adding one extra point, say \( c \in [x_{k-1}, x_k] \). Let

\[ m'_k = \inf_{x \in [x_{k-1}, c]} f(x), \quad m''_k = \inf_{x \in [c, x_k]} f(x). \]

Then \( m'_k \geq m_k, \quad m''_k \geq m_k \), and

\[ L(f, P') = \sum_{i=1}^{k-1} m_i(x_i - x_{i-1}) + m'_k(c - x_{k-1}) + m''_k(x_k - c) + \sum_{i=k+1}^{n} m_i(x_i - x_{i-1}) \]

\[ \geq \sum_{i=1}^{k-1} m_i(x_i - x_{i-1}) + m_k(x_k - x_{k-1}) + \sum_{i=k+1}^{n} m_i(x_i - x_{i-1}) = L(f, P). \]

Similarly one obtains

\[ U(f, P') \leq U(f, P). \]

To prove the assertion, add a finite number of points to \( P \) to form \( Q \), and apply the above result repeatedly. \( \square \)

Proposition 13.13 Let \( f : [a, b] \to \mathbb{R} \) be a bounded function, and let \( P \) and \( Q \) be arbitrary partitions of \([a, b]\). Then

\[ L(f, P) \leq U(f, Q). \]

Proof. Consider the partition \( P \cup Q \). By Proposition 13.11 we obtain

\[ L(f, P) \leq L(f, P \cup Q) \leq U(f, P \cup Q) \leq U(f, Q). \]

\( \square \)
Proposition 13.14 Let $f : [a, b] \to \mathbb{R}$ be a bounded function and let $J$ and $j$ be the upper and lower integrals of $f$ over $[a, b]$. Then
\[ j \leq J. \]

Proof. Fix a partition $Q$. By Proposition 13.13, for any partition $P$ of $[a, b]$ we have
\[ L(f, P) \leq U(f, Q). \]
Therefore
\[ j = \sup_P L(f, P) \leq U(f, Q). \]
In other words, for any partition $P$,
\[ j \leq U(f, P). \]
Hence
\[ j \leq \inf_P U(f, P) = J. \]

14 Criterion of integrability

Theorem 14.1 (Criterion of integrability) A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if
\[ \forall \varepsilon > 0 \ \exists \ \text{a partition } P \text{ of } [a, b] \ [U(f, P) - L(f, P) < \varepsilon]. \]

Proof. First, assume that $f$ is integrable; that is, $J = j$. By definition of the supremum, there exists a partition $P_1$ such that
\[ L(f, P_1) > j - \varepsilon/2. \]
Also, by definition of the infimum, there exists a partition $P_2$ such that
\[ U(f, P_2) < J + \varepsilon/2. \]
Set $Q = P_1 \cup P_2$. Then $Q$ is a refinement of both $P_1$ and $P_2$. By Propositions 13.11 and 13.13 we obtain
\[ j - \varepsilon/2 < L(f, P_1) \leq L(f, Q) \leq U(f, Q) \leq U(f, P_2) < J + \varepsilon/2. \]
Since $J = j$ by assumption, we conclude that
\[ U(f, Q) - L(f, Q) < \varepsilon. \]
That is, (19) holds.

Conversely, assume that (19) holds. Fix $\varepsilon > 0$. Let $P$ be a partition such that
\[ U(f, P) - L(f, P) < \varepsilon. \]
Note that $J \leq U(f, P)$ and $j \geq L(f, P)$ by definition of $j$ and $J$. So
\[ J - j \leq U(f, P) - L(f, P) < \varepsilon. \]
It follows that 

$$ \forall \varepsilon > 0 \ [J - j < \varepsilon]. $$

Since $\varepsilon > 0$ can be chosen arbitrarily small, we see that $J = j$, and hence that $f$ is integrable. \qed

The following corollary is a reformulation of the above criterion of integrability. It is convenient in computations (and was used in Example 13.10).

**Corollary 14.2** Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Assume that there exists a sequence of partitions $(P_n)_{n \in \mathbb{N}^+}$ of $[a, b]$ such that

$$ \lim_{n \to \infty} (U(f, P_n) - L(f, P_n)) = 0.$$

Then $f$ is Riemann integrable over $[a, b]$, and

$$ \int_a^b f(x) \, dx = \lim_{n \to \infty} U(f, P_n) = \lim_{n \to \infty} L(f, P_n). $$

**Proof.** Exercise. \qed

### 15 Classes of integrable functions

**Theorem 15.1** Let $f : [a, b] \to \mathbb{R}$ be monotone (that is, either increasing or decreasing). Then $f$ is Riemann integrable over $[a, b]$.

**Proof.** Assume that $f$ is increasing. If $f$ is constant, then $f$ is Riemann integrable by Example 13.7. If $f$ is not constant, then $f(a) < f(b)$. Fix $\varepsilon > 0$. Consider a partition $P$ of $[a, b]$ with the property that

$$ \|P\| < \delta := \frac{\varepsilon}{f(b) - f(a)}. $$

For this partition we obtain

$$ U(f, P) - L(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))(x_i - x_{i-1}) $$

$$ < \delta \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) = \delta (f(b) - f(a)) = \varepsilon. $$

By the Criterion of Integrability, we conclude that $f$ is integrable. \qed

**Theorem 15.2** Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then $f$ is Riemann integrable over $[a, b]$.

**Proof.** Fix $\varepsilon > 0$. By Cantor’s Theorem on uniform continuity, $f$ is uniformly continuous on $[a, b]$. Set $\eta := \varepsilon/(b - a)$. There exists $\delta > 0$ such that

$$ \forall x, y \in [a, b] \ [(|x - y| < \delta) \Rightarrow (|f(x) - f(y)| < \eta). $$
Let \( P = \{x_0, \ldots, x_n\} \) be a partition with norm \( \|P\| < \delta \). Recall that a continuous function on a closed bounded interval attains its extrema. Therefore, for each \( i = 1, \ldots, n \),

\[
M_i = \sup_{x \in [x_{i-1}, x_i]} f(x) = f(y_i)
\]

and

\[
m_i = \inf_{x \in [x_{i-1}, x_i]} f(x) = f(z_i)
\]

for some \( y_i, z_i \in [x_{i-1}, x_i] \). So

\[
M_i - m_i = f(y_i) - f(z_i) < \eta
\]
as \( |y_i - z_i| < \delta \). We then have

\[
U(f, P) - L(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \eta \sum_{i=1}^{n} (x_i - x_{i-1}) = \eta (b - a) = \varepsilon.
\]

By the Criterion of Integrability, we conclude that \( f \) is integrable. \( \square \)

There are, however, Riemann integrable functions that are neither monotone nor continuous.

**Example 15.3** Define \( f : [0, 1] \to \mathbb{R} \) by

\[
f(x) = \begin{cases} 
0 & \text{if } x \notin \mathbb{Q}, \\
1/q & \text{if } x \in \mathbb{Q} \text{ and } x = p/q \text{ in lowest terms.}
\end{cases}
\]

Note that \( f \) is continuous on \([0, 1] \setminus \mathbb{Q}\) and discontinuous on \( \mathbb{Q} \cap [0, 1] \) and is not monotone. Fix \( \varepsilon > 0 \). Put

\[
S := \{ x \in [0, 1] : f(x) > \varepsilon/2 \}.
\]

Then \( S \) contains only finitely many points. Say

\[
S = \{y_1, \ldots, y_m\}.
\]

Choose a partition \( P = \{x_0, \ldots, x_n\} \) of \([0, 1]\) with \( \|P\| < \eta := \varepsilon/4m \). If the interval \([x_{i-1}, x_i]\) contains no points in \( S \), then \( f(x) \leq \varepsilon \) for all \( x \in [x_{i-1}, x_i] \). Also, at most \( 2m \) intervals \([x_{i-1}, x_i]\) contain points in \( S \). And, in addition, \( f(x) \leq 1 \) for any \( x \in [0, 1] \). Therefore,

\[
U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) < \varepsilon/2 + 2m \eta = \varepsilon.
\]

It is also the case that \( L(f, P) = 0 \). So \( f \) is Riemann integrable and

\[
\int_{0}^{1} f(x) \, dx = 0.
\]
16 Inequalities and the mean-value property of the integral

The following inequality is useful for estimating integrals.

**Theorem 16.1** Let \( f, g \) be bounded and integrable on \([a, b]\). Suppose that

\[
(\forall x \in [a, b]) \; (f(x) \leq g(x)).
\]

Then

\[
\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx.
\]

**Proof.** Let \( P \) be a partition of \([a, b]\). Then

\[
L(f, P) \leq L(g, P) \leq \int_a^b g(x) \, dx.
\]

Taking the supremum over all partitions \( P \) of \([a, b]\) yields the result. \( \square \)

**Exercise 16.2** Suppose that \( f \) is bounded and integrable on \([a, b]\). Assume that

\[
(\forall x \in [a, b]) \; (m \leq f(x) \leq M)
\]

for some \( m, M \in \mathbb{R} \). Then

\[
m(b - a) \leq \int_a^b f(x) \, dx \leq M(b - a).
\]

**Hint.** Apply Theorem 16.1. \( \square \)

**Corollary 16.3** (Mean-value property of the Riemann integral) Let \( f \) be continuous on \([a, b]\). Then there exists \( \theta \in [a, b] \) such that

\[
\int_a^b f(x) \, dx = f(\theta)(b - a).
\]

**Proof.** Set

\[
m := \min_{[a, b]} f(x), \quad M := \max_{[a, b]} f(x).
\]

By Exercise 16.2,

\[
m \leq \frac{1}{b - a} \int_a^b f(x) \, dx \leq M,
\]

Applying the intermediate value theorem (Theorem A 4.3.1) to the function \( f \) we see that there exists \( \theta \in [a, b] \) such that

\[
f(\theta) = \frac{1}{b - a} \int_a^b f(x) \, dx.
\]

This completes the proof. \( \square \)
17 Further properties of the integral

In this section we present some elementary properties of integral which are familiar from Calculus.

**Theorem 17.1** Let \( a < c < b \). Let \( f \) be integrable on \([a,b]\). Then \( f \) is integrable on \([a,c]\) and on \([c,b]\) and
\[
\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.
\]
Conversely, if \( f \) is integrable on \([a,c]\) and on \([c,b]\) then it is integrable on \([a,b]\).

**Proof.** Suppose that \( f \) is integrable on \([a,b]\). Fix \( \varepsilon > 0 \). Then there exists a partition \( P = \{x_0, \ldots, x_n\} \) of \([a,b]\) such that
\[
U(f,P) - L(f,P) < \varepsilon.
\]
We can assume that \( c \in P \) so that \( c = x_j \) for some \( j \in \{0, 1, \ldots, n\} \) (otherwise consider the refinement of \( P \) adding the point \( c \)). Then \( P_1 = \{x_0, \ldots, x_j\} \) is a partition of \([a,c]\) and \( P_2 = \{x_j, \ldots, x_n\} \) is a partition of \([c,b]\). Moreover,
\[
L(f,P) = L(f,P_1) + L(f,P_2), \quad U(f,P) = U(f,P_1) + U(f,P_2).
\]
Therefore we have
\[
[U(f,P_1) - L(f,P_1)] + [U(f,P_2) - L(f,P_2)] = U(f,P) - L(f,P) < \varepsilon.
\]
Since each of the terms on the left hand side is non-negative, each one is less than \( \varepsilon \), which proves that \( f \) is integrable on \([a,c]\) and on \([c,b]\). Note also that
\[
L(f,P_1) \leq \int_a^c f(x)dx \leq U(f,P_1), \quad L(f,P_2) \leq \int_c^b f(x)dx \leq U(f,P_2),
\]
so that
\[
L(f,P) \leq \int_a^c f(x)dx + \int_c^b f(x)dx \leq U(f,P).
\]
This is true for any partition of \([a,b]\). Therefore
\[
\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.
\]
Now assume that \( f \) is integrable on \([a,c]\) and \([c,b]\). Fix \( \varepsilon > 0 \). Then there exists a partition \( P_1 \) of \([a,c]\) and a partition \( P_2 \) of \([c,b]\) such that
\[
U(f,P_1) - L(f,P_1) < \varepsilon/2, \quad U(f,P_2) - L(f,P_2) < \varepsilon/2.
\]
Let \( P = P_1 \cup P_2 \). Then
\[
U(f,P) - L(f,P) = [U(f,P_1) - L(f,P_1)] + [U(f,P_2) - L(f,P_2)] < \varepsilon.
\]
By the Criterion of Integrability we conclude that \( f \) is integrable on \([a,b]\). \( \square \)
Remark 17.2 The integral $\int_a^b f(x)dx$ was defined only for $a < b$. We add by definition that
\[
\int_a^a f(x)dx = 0 \quad \text{and} \quad \int_b^a f(x)dx = -\int_a^b f(x)dx \quad \text{if } a > b.
\]
With this convention we always have that
\[
\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.
\]

Theorem 17.3 Let $f$ and $g$ be integrable on $[a, b]$. Then $f + g$ is also integrable on $[a, b]$ and
\[
\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx.
\]
Proof. Let $P = \{x_0, \ldots, x_n\}$ be a partition of $[a, b]$. Let
\[
m_i' = \inf\{f(x) : x_{i-1} \leq x \leq x_i\},
m_i'' = \inf\{g(x) : x_{i-1} \leq x \leq x_i\},
m_i = \inf\{f(x) + g(x) : x_{i-1} \leq x \leq x_i\}.
\]
Define $M_i, M_i', M_i''$ similarly. The following inequalities hold
\[
m_i \geq m_i' + m_i'', \quad M_i \leq M_i' + M_i''.
\]
Therefore we have
\[
L(f, P) + L(g, P) \leq L(f + g, P), \quad U(f + g, P) \leq U(f, P) + U(g, P).
\]
Hence for any partition $P$
\[
L(f, P) + L(g, P) \leq L(f + g, P) \leq U(f + g, P) \leq U(f, P) + U(g, P),
\]
or otherwise
\[
U(f + g, P) - L(f + g, P) \leq [U(f, P) - L(f, P)] + [U(g, P) - L(g, P)].
\]
Fix $\varepsilon > 0$. Since $f$ and $g$ are integrable there are partitions $P_1$ and $P_2$ such that
\[
U(f, P_1) - L(f, P_1) < \varepsilon/2, \quad U(g, P_2) - L(g, P_2) < \varepsilon/2.
\]
Thus for the partition $P = P_1 \cup P_2$ we obtain that
\[
U(f + g, P) - L(f + g, P) < \varepsilon.
\]
This proves that $f + g$ is integrable on $[a, b]$. Moreover,
\[
L(f, P) + L(g, P) \leq L(f + g, P) \leq \int_a^b [f(x) + g(x)]dx \leq U(f + g, P) \leq U(f, P) + U(g, P), \quad \text{and}
\]
\[
L(f, P) + L(g, P) \leq \int_a^b f(x)dx + \int_a^b g(x)dx \leq U(f, P) + U(g, P).
\]
Therefore it follows that
\[
\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx.
\]
The proof is complete. \qed
Theorem 17.4 Let $f$ be integrable on $[a, b]$. Then, for any $c \in \mathbb{R}$, $cf$ is also integrable on $[a, b]$ and
\[
\int_a^b cf(x)dx = c \int_a^b f(x)dx.
\]
Proof. The proof is left as an exercise. Consider separately two cases: $c \geq 0$ and $c \leq 0$. \qed

Theorem 17.5 Let $f$ be integrable on $[a, b]$. Then $|f|$ is integrable on $[a, b]$ and
\[
|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx.
\]
Proof. Note that for any interval $[\alpha, \beta]$ we have
\[
\sup_{[\alpha, \beta]} |f(x)| - \inf_{[\alpha, \beta]} |f(x)| \leq \sup_{[\alpha, \beta]} f(x) - \inf_{[\alpha, \beta]} f(x).
\]
Indeed,
\[
(\forall x, y \in [\alpha, \beta]) \left( f(x) - f(y) \leq \sup_{[\alpha, \beta]} f(x) - \inf_{[\alpha, \beta]} f(x) \right),
\]
so that
\[
(\forall x, y \in [\alpha, \beta]) \left( |f(x)| - |f(y)| \leq \sup_{[\alpha, \beta]} f(x) - \inf_{[\alpha, \beta]} f(x) \right),
\]
which proves (20) by passing to the supremum in $x$ and $y$. It follows from (20) that for any partition of $[a, b]$
\[
U(|f|, P) - L(|f|, P) \leq U(f, P) - L(f, P),
\]
which proves the integrability of $|f|$ by the Criterion of Integrability. The last assertion follows from Theorem 16.1. \qed

Theorem 17.6 Let $f : [a, b] \rightarrow \mathbb{R}$ be an integrable function such that
\[
(\forall x \in [a, b])(m \leq f(x) \leq M).
\]
Let $g : [m, M] \rightarrow \mathbb{R}$ be a continuous function. Then the composite function $h : [a, b] \rightarrow \mathbb{R}$
defined by $h(x) = g(f(x))$ is integrable.
Proof. Fix $\varepsilon > 0$. Since $g$ is uniformly continuous on $[m, M]$, there exists $\delta > 0$ such that
\[
(\forall t, s \in [m, M]) [(|t - s| < \delta) \Rightarrow (|g(t) - g(s)| < \varepsilon)].
\]
By integrability of $f$ there exists a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ such that
\[
U(f, P) - L(f, P) < \delta^2.
\]
Let $m_i = \inf_{[x_{i-1}, x_i]} f(x)$, $M_i = \sup_{[x_{i-1}, x_i]} f(x)$ and $m_i^* = \inf_{[x_{i-1}, x_i]} h(x)$, $M_i^* = \sup_{[x_{i-1}, x_i]} h(x)$. Decompose the set \{1, \ldots, n\} into two subset: $(i \in A) \Leftrightarrow (M_i - m_i < \delta)$ and $(i \in B) \Leftrightarrow (M_i - m_i \geq \delta)$.
For $i \in A$ by the choice of $\delta$ we have that $M_i^* - m_i^* < \varepsilon$.
For $i \in B$ we have that $M_i^* - m_i^* \leq 2K$ where $K = \sup_{t \in [m,M]} |g(t)|$. By (21) we have
\[
\delta \sum_{i \in A} (x_i - x_{i-1}) \leq \sum_{i \in B} (M_i - m_i)(x_i - x_{i-1}) < \delta^2,
\]
48
so that \( \sum_{i \in B}(x_i - x_{i-1}) < \delta \). Therefore
\[
U(h, P) - L(h, P) = \sum_{i \in A}(M_i^* - m_i^*)(x_i - x_{i-1}) + \sum_{i \in B}(M_i^* - m_i^*)(x_i - x_{i-1}) < \varepsilon(b - a) + 2K\delta < \varepsilon[(b - a) + 2K],
\]
which proves the assertion since \( \varepsilon > 0 \) is arbitrary. \( \square \)

**Corollary 17.7** Let \( f, g \) be integrable on \([a, b]\). Then the product \( fg \) is integrable on \([a, b]\).

**Proof.** Since \( f + g \) and \( f - g \) are integrable on \([a, b]\), \((f + g)^2\) and \((f - g)^2\) are integrable on \([a, b]\) by the previous theorem. Therefore
\[
fg = \frac{1}{4}[(f + g)^2 - (f - g)^2]
\]
is integrable on \([a, b]\). \( \square \)

18 Integration as the inverse to differentiation

**Definition 18.1** Let \( f: [a, b] \to \mathbb{R} \) be a given function. Suppose that \( F: [a, b] \to \mathbb{R} \) is continuous and differentiable on \((a, b)\) with the property that
\[
F'(x) = f(x) \quad \text{for all } x \in (a, b).
\]
Then \( F \) is said to be a **primitive** or **antiderivative** of \( f \) on \([a, b]\).

**Remark 18.2** Primitives are not uniquely defined. If \( F \) is a primitive of \( f \) then so is \( F + C \) where \( C \) is an arbitrary constant. Also, with \( C := -F(a) \), the primitive \( G(x) := F(x) + C \) satisfies \( G(a) = 0 \).

**Example 18.3** Let \( n \in \mathbb{N}^+ \). For each \( C \in \mathbb{R} \), the function \( F(x) = \frac{x^{n+1}}{n+1} + C \) is a primitive of \( f(x) = x^n \).

Let \( f \) be a bounded integrable function on some interval \([a, b]\). We now use the Riemann integral to construct a primitive of \( f \) on \([a, b]\). Firstly,

**Theorem 18.4** Let \( f \) be bounded and integrable on \([a, b]\). Define \( F: [a, b] \to \mathbb{R} \) by
\[
F(x) := \int_a^x f(t) \, dt \quad (x \in [a, b]).
\]
Then \( F \) is continuous on \([a, b]\).

**Proof.** Put \( M := \sup_{[a,b]} |f(x)| \). Let \( x, y \in [a, b] \) with \( x < y \). Making use of Theorem 17.1, we obtain
\[
|F(y) - F(x)| = \left| \int_x^y f(t) \, dt \right| \leq M |x - y|.
\]
It follows from this estimate that \( F \) is continuous on \([a, b]\) (in fact, it is uniformly continuous there [cf. FTA]). \( \square \)
Theorem 18.5 (Existence of primitive) Let $f$ be continuous on $[a, b]$. Define $F : [a, b] \to \mathbb{R}$ as in (22). Then $F$ is differentiable on $(a, b)$ and

$$F'(x) = f(x) \quad (x \in (a, b)).$$

That is, $F$ is a primitive of $f$ on $[a, b]$.

Proof. By Theorem 18.4, $F$ is continuous on $[a, b]$. Let $x \in (a, b)$. Let $h > 0$. Again with the help of Theorem 17.1,

$$\frac{F(x + h) - F(x)}{h} = \frac{1}{h} \int_{x}^{x+h} f(t) \, dt.$$

By the mean value property of the integral (Corollary 16.3), there exists $\theta \in [x, x + h]$ such that

$$\int_{x}^{x+h} f(t) \, dt = f(\theta) \, h.$$

Thus,

$$\frac{F(x + h) - F(x)}{h} = f(\theta).$$

As $h \to 0$, $\theta \to x$. By continuity of $f$, therefore, $\lim_{h \to 0} f(\theta) = f(x)$. This shows that

$$F'_+(x) = f(x).$$

A similar argument yields $F'_-(x) = f(x)$. So $F$ is differentiable at $x$ and $F'(x) = f(x)$.

We now arrive at the two central theorems of this section.

Theorem 18.6 (Fundamental theorem of calculus) Let $f$ be continuous on $[a, b]$ and $G$ a primitive of $f$ on $[a, b]$. Then, for any $x \in [a, b]$,

$$\int_{a}^{x} f(t) \, dt = G(x) - G(a).$$

In particular,

$$\int_{a}^{b} f(t) \, dt = G(b) - G(a).$$

Proof. Define $F : [a, b] \to \mathbb{R}$ as in (22). Define $H := F - G$ on $[a, b]$. Then $H$ is continuous on $[a, b]$ (Theorem 18.4) and differentiable on $(a, b)$ (Theorem 18.5) with the property that

$$H' = F' - G' = f - f = 0$$

on $(a, b)$. By Corollary A 5.2.2, there exists $C \in \mathbb{R}$ such that

$$H = C, \text{ or } F = G + C$$

on $[a, b]$. Now $F(a) = 0$, so $C = -G(a)$. Thus, for $x \in [a, b]$,

$$F(x) = \int_{a}^{x} f(t) \, dt = G(x) + C = G(x) - G(a),$$

as required.

The above statement also holds for bounded Riemann integrable functions in place of continuous functions.
Theorem 18.7 (fundamental theorem of calculus for integrable functions) Let \( f \) be bounded and integrable on \([a,b]\) and \( G \) a primitive of \( f \) on \([a,b]\). Then
\[
\int_a^b f(x) \, dx = G(b) - G(a).
\]

Proof. Let \( P = \{x_0, \ldots, x_n\} \) be a partition of \([a,b]\). Note that \( G \) is continuous on \([x_{i-1}, x_i]\) and differentiable on \((x_{i-1}, x_i)\) by definition of the primitive. By the mean value theorem, there exists \( t_i \in (x_{i-1}, x_i) \) such that
\[
G(x_i) - G(x_{i-1}) = G'(t_i)(x_i - x_{i-1}) = f(t_i)(x_i - x_{i-1}).
\]

Put
\[
m_i := \inf_{[x_{i-1}, x_i]} f(x), \quad M_i := \sup_{[x_{i-1}, x_i]} f(x).
\]

Then
\[
m_i (x_i - x_{i-1}) \leq f(t_i)(x_i - x_{i-1}) \leq M_i (x_i - x_{i-1}),
\]
in other words,
\[
m_i (x_i - x_{i-1}) \leq G(x_i) - G(x_{i-1}) \leq M_i (x_i - x_{i-1}).
\]

Adding these inequalities for \( i = 1, \ldots, n \) we obtain
\[
\sum_{i=1}^n m_i (x_i - x_{i-1}) \leq G(b) - G(a) \leq \sum_{i=1}^n M_i (x_i - x_{i-1}).
\]

So, for any partition \( P \), we have
\[
L(f, P) \leq G(b) - G(a) \leq U(f, P).
\]

As \( f \) is integrable, this entails that
\[
G(b) - G(a) = \int_a^b f(x) \, dx.
\]