Real Analysis - Math230A/B
Version: Fall 2020

András Domokos , PhD
California State University, Sacramento

February 17, 2021

Contents

Introduction

This textbook covers the material for the graduate Real Analysis classes Math 230A/B.

To preserve departmental traditions, the presentation of the material is mostly following the contents of Chapters 1-7 from Walter Rudin's Principles of Mathematical Analysis textbook. However, the logic and flow of the presentation is different in many places. Some new material is added, to make this course a better prerequisite for subsequent courses in Analysis, Measure Theory and Functional Analysis. The mathematical notations are updated to today's standards.

This textbook includes those exercises which are not present in Rudin's book, but which were usually given as extra worksheets for preparation for the Master's exam in Analysis. Also, the exercises are listed for each section instead of chapters, which allows a clearer correlation between exercises and the relevant theoretical knowledge.

One of the major goals is to provide a textbook that can be efficiently used by students with different mathematical backgrounds. Each section contains enough elementary information to give students the opportunity to refresh the prerequisite knowledge.

Throughout the textbook, elementary and advanced mathematical ideas are mixed in such a way to make the learning process more connected to undergraduate courses in calculus and analysis.

Chapter 1

Sets and Functions

1.1 Sets

The concept of set plays an important role in mathematics. We understand that a set is a collection of some objects, which are called elements of the set. Sets will be denoted by capital letters A, B, ... , X, Y, Z, and their elements by lower case letters, a, b , ... , x, y, z. We will write $a \in A$ for a being an element of A, and $a \notin A$ for a not being an element of A.

For defining a set, we will use one of the following methods:

1. - Listing the elements: $A = \{a, b, c\}$,
2. - Listing the properties of the elements: $A = \{a|$ properties of a$\}$.
3. - For the empty set we use the notation $\emptyset$.

We say that a set A is a subset of another set B, denoted by $A \subset B$, if every element of A belongs to B. We can also state this by the following implication: If $a \in A$, then $a \in B$.

We say that $A = B$ if $A \subset B$ and $B \subset A$.

The following set operations will be needed:

- Union:

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}.$$

- Intersection:

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$

- Difference:

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

- Symmetric difference:

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

- Complement: If $A \subset X$, then the complement of A within X is defined as:

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

- The Cartesian product:

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

We can extend the union and intersection to any number of sets. Let $\{A_i| \ i \in I\}$ be a collection of sets indexed by the elements of an index set I. Then,

$$\bigcup_{i \in I} A_i = \{x \mid \exists\, i_x \in I,\ x \in A_{i_x}\},$$

and

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

We say that to sets A and B are disjoint if $A \cap B = \emptyset$ and that a collection of sets $\{A_i| \ i \in I\}$ is pairwise disjoint if

$$A_i \cap A_j = \emptyset,\ \ \forall\, i, j \in I,\ i \neq j.$$

The following sets will be frequently used:

- The set of natural numbers:

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

- The set of integer numbers:

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}.$$

- The set of rational numbers:

$$\mathbb{Q} = \left\{ \frac{m}{n} \;\middle|\; m \in \mathbb{Z},\ n \in \mathbb{N} \right\}.$$

- The set of all real numbers: $\mathbb{R}$.

- The set of irrational numbers:

$$\mathbb{I} = \{x \in \mathbb{R} \mid x \notin \mathbb{Q}\}.$$

- Open intervals:

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

- Closed intervals:

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}.$$

- Half-closed intervals:

$$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\},$$
$$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}.$$

Exercises.

Exercise 1.1. Let $A, B, C \subset X$. Prove the following:

$$\begin{array}{ll}
(a) & A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \\
(b) & A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \\
(c) & A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C) \\
(d) & A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C) \\
(e) & (A \cup B)^c = A^c \cap B^c \\
(f) & (A \cap B)^c = A^c \cup B^c.
\end{array}$$

Exercise 1.2. Prove that:

$$(a) \qquad \bigcup_{n=1}^{\infty} \left[\frac{1}{n}, 1\right] = (0, 1]$$

$$(b) \qquad \bigcap_{n=1}^{\infty} \left(0, \frac{n+1}{n}\right) = (0, 1].$$

Exercise 1.3. Prove that $\sqrt{2} \notin \mathbb{Q}$.

1.2 Functions

Definition 1.1. Let X and Y be two sets. A function $f : X \to Y$ is a rule associating to each element of X a unique element of Y. The set X is called the domain of the function and the set Y is called the codomain of the function.

Notice that each function is defined by three things: domain, codomain and rule. Hence it is logical to require that two functions f and g are equal if they have the same domain and codomain, i.e. $f : X \to Y$, $g : X \to Y$, and moreover $f(x) = g(x)$ for each $x \in X$.

Definition 1.2. Let $f : X \to Y$, $A \subset X$ and $B \subset Y$. The image of A through the function f is defined by

$$f(A) = \{y \in Y \mid \exists\, x \in X,\ f(x) = y\},$$

and the pre-image of B through f is defined by

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

We call $f(X)$ the range of the function f.

Example 1.3. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Then,

$$\begin{aligned} f\big([0,2)\big) &= [0,4), \\ f\big((-1,3]\big) &= [0,9], \\ f^{-1}\big([0,4]\big) &= [-2,2], \\ f^{-1}\big(\{2\}\big) &= \{-\sqrt{2}, \sqrt{2}\}. \end{aligned}$$

Proposition 1.4. *Let $f : X \to Y$, $A_1, A_2 \subset X$ and $B_1, B_2 \subset Y$. The following relations hold:*

$$\begin{aligned} &(a) \qquad f(A_1 \cup A_2) = f(A_1) \cup f(A_2). \\ &(b) \qquad f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2). \\ &(c) \qquad f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2). \\ &(d) \qquad f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2). \\ &(e) \qquad f^{-1}(Y \setminus B_1) = X \setminus f^{-1}(B_1). \end{aligned}$$

Proof. (a) We will prove it by a sequence of equivalent statements.

$$\begin{aligned} & y \in f(A_1 \cup A_2) \\ \Leftrightarrow\ & \exists\, x \in A_1 \cup A_2 \text{ such that } f(x) = y \\ \Leftrightarrow\ & \exists\, x \in A_1 \text{ or } x \in A_2 \text{ such that } f(x) = y \\ \Leftrightarrow\ & y \in f(A_1) \text{ or } y \in f(A_2) \\ \Leftrightarrow\ & y \in f(A_1) \cup f(A_2). \end{aligned}$$

(b) We follow a similar proof as for (a), but we can claim only a one way implication.

$$\begin{aligned}
& y \in f(A_1 \cap A_2) \\
\Rightarrow\ & \exists\, x \in A_1 \cap A_2 \text{ such that } f(x) = y \\
\Rightarrow\ & \exists\, x \in A_1 \text{ and } x \in A_2 \text{ such that } f(x) = y \\
\Rightarrow\ & y \in f(A_1) \text{ and } y \in f(A_2) \\
\Rightarrow\ & y \in f(A_1) \cap f(A_2).
\end{aligned}$$

Which implication cannot be reversed?

(c) The proof is left as an exercise.

(d)

$$\begin{aligned}
& x \in f^{-1}(B_1 \cap B_2) \\
\Leftrightarrow\ & f(x) \in B_1 \cap B_2 \\
\Leftrightarrow\ & f(x) \in B_1 \text{ and } f(x) \in B_2 \\
\Leftrightarrow\ & x \in f^{-1}(B_1) \text{ and } x \in f^{-1}(B_2) \\
\Leftrightarrow\ & x \in f^{-1}(B_1) \cap f^{-1}(B_2).
\end{aligned}$$

(e) The proof is left as an exercise.

$\square$

Definition 1.5. Let $f : X \to Y$.
We say that f is a one-to-one function if for any $x_1, x_2 \in X$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$, or equivalently, if $f(x_1) = f(x_2)$ impies $x_1 = x_2$.
We say that f is an onto function if for every $y \in Y$ there exists $x \in X$ such that $f(x) = y$.

Example 1.6. The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ is not one-to-one and not onto. Indeed, $f(-1) = f(1) = 1$ and for $y = -4$ there is no $x \in \mathbb{R}$ such that $x^2 = -4$.

However, the function $f : [0, \infty) \to \mathbb{R}$, $f(x) = x^2$, is one-to-one, but not onto.

Moreover, the function $f : [0, \infty) \to [0, \infty)$, $f(x) = x^2$ is one-to-one and onto.

Proposition 1.7. *Let $f : X \to Y$ be a one-to-one function. Then there exist a function $g : Y \to X$ such that $g(f(x)) = x$, for every $x \in X$.*

Proof. Fix any $x_0 \in X$ and define $g : Y \to X$ as follows:

$$g(y) = \begin{cases} x & \text{if } y \in f(X) \text{ and } f(x) = y \\ x_0 & \text{if } y \notin f(X). \end{cases}$$

Notice that g is well defined, because by the one-to-one property of f, for each $y \in f(X)$ there is a unique $x \in X$ such that $f(x) = y$. The definition of g easily implies that $g(f(x)) = x$ for every $x \in X$.

$\square$

Proposition 1.8. *Let $f : X \to Y$ be an onto function. Then there exist a function $h : Y \to X$ such that $f(h(y)) = y$, for every $y \in Y$.*

Proof. The onto property of f implies that for each $y \in Y$ there exists $x \in X$ such that $f(x) = y$. Define $h : Y \to X$ by $h(y) = x$, where for each $y \in Y$ the $x \in X$ is chosen such that $f(x) = y$. In other words, for each $y \in Y$, we select an element $x \in f^{-1}(\{y\})$.

Therefore, $f(h(y)) = f(x) = y$ for every $y \in Y$.

$\square$

The next is the definition of the inverse function f^{-1}.

Definition 1.9. Let $f : X \to Y$ be a function. We say that f is invertible if there exist a function $f^{-1} : Y \to X$ such that $f^{-1}(f(x)) = x$, for every $x \in X$ and $f(f^{-1}(y) = y$ for every $y \in Y$.

It is easy to observe that if f is invertible, the inverse function is uniquely defined.

Proposition 1.10. *Let $f : X \to Y$ be a function. Then f is invertible if and only if f one-to-one and onto.*

Proof. The proofs of the previous two propositions show that if the function f is both one-to-one and onto, the functions g and h are equal, and this common function is the inverse function f^{-1}. The reverse implication is left as an exercise.

$\square$

Example 1.11. The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ is not invertible, so the notation $f^{-1}(4)$ doesn't make sense without extra explanations. However, $f^{-1}(\{4\})$, as the pre-image of the set $\{4\}$, does make sense.

The function $f : [0, \infty) \to \mathbb{R}$, $f(x) = x^2$ is one-to-one and hence the function g from Proposition 1.7 can be defined as $g : \mathbb{R} \to [0, \infty)$,

$$g(y) = \begin{cases} \sqrt{y} & \text{if } y \geq 0 \\ 5 & \text{if } y < 0. \end{cases}$$

The function $f : \mathbb{R} \to [0, \infty)$, $f(x) = x^2$ is onto, hence we can define the function h from Proposition 1.8 as $h : [0, \infty) \to \mathbb{R}$, $h(y) = \sqrt{y}$. Note that the choice of $h(y) = -\sqrt{y}$ is equally good.

Finally, the function $f : [0, \infty) \to [0, \infty)$, $f(x) = x^2$, is both one-to-one and onto and has the inverse function $f^{-1} : [0, \infty) \to [0, \infty)$, $f^{-1}(y) = \sqrt{y}$.

Exercises.

Exercise 1.4. Let $f : X \to Y$ be a function. Show that f is one-to-one if and only if for every $y \in Y$, the pre-image $f^{-1}(\{y\})$ contains at most one element.

Exercise 1.5. Let $f : X \to Y$ be a function. Show that f is onto if and only if for every $y \in Y$, the pre-image $f^{-1}(\{y\}) \neq \emptyset$.

Exercise 1.6. Let $f : X \to Y$ be a function. Show that f is one-to-one if and only if $f(A_1 \cap A_2) = f(A_1) \cap f(A_2)$, for every pair of sets $A_1, A_2 \subset X$.

Exercise 1.7. Let $f : X \to Y$ be a function. Show that f is one-to-one if and only if $f^{-1}(f(A)) = A$, for every set $A \subset X$.

Exercise 1.8. Let $f : X \to Y$ be a function. Show that f is onto if and only if $f(f^{-1}(B)) = B$, for every set $B \subset Y$.

Exercise 1.9. Let $f : X \to Y$ be a function. Show that f is one-to-one and onto if and only if $f(A^c) = f(A)^c$, for every set $A \subset X$.

Exercise 1.10. Let $f : X \to Y$ and $g : Y \to Z$ be one-to-one and onto functions. Show that $g \circ f : X \to Z$ is invertible and

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1} .$$

Exercise 1.11. Let $f : X \to Y$ be a function. Show that f is one-to-one if and only if there exist a function $g : Y \to X$ such that $g(f(x)) = x$, for every $x \in X$.

Exercise 1.12. Let $f : X \to Y$ be a function. Show that f is onto if and only if there exist a function $h : Y \to X$ such that $f(h(y)) = y$, for every $y \in Y$.

1.3 Countable and uncountable sets

Definition 1.12. We say that two sets X and Y are equivalent , if there exists a one-to-one and onto function $f : X \to Y$. For this equivalence we will use the notation $X \sim Y$.

Example 1.13.

(1)
$$\{1, 2, 3\} \sim \{a, b, c\}.$$
The function $f : \{1, 2, 3\} \to \{a, b, c\}$, defined by $f(1) = a,\ f(2) = b$ and $f(3) = c$ is one-to-one and onto.

(2)
$$\{a, b, c\} \not\sim \{1, 2, 3, 4\}.$$
Indeed, if $f : \{a, b, c\} \to \{1, 2, 3, 4\}$ is any function, then $f(\{a, b, c\})$ cannot contain more then three elements, hence f cannot be onto.

(3)
$$\{2, 4, 6, ...\} \sim \mathbb{N}.$$
The function $f : \{2, 4, 6, ...\} \to \mathbb{N},\ f(n) = \frac{n}{2}$, is one-to-one and onto.

(4)
$$\mathbb{N} \sim \mathbb{Z}.$$
Define the function $f : \mathbb{N} \to \mathbb{Z}$ by $f(1) = 0,\ f(2n) = n$ and $f(2n+1) = -n$, for every $n \in \mathbb{N}$. It is evident that this is a one-to-one and onto function.

(5)
$$\mathbb{R} \sim (0, 1).$$
The function $f : \mathbb{R} \to (0, 1),\ f(x) = \frac{1}{\pi} \arctan x + \frac{1}{2}$ is one-to-one and onto.

The proof of the following proposition follows from the properties of one-to-one and onto functions, and it is left as an exercise.

Proposition 1.14. *Let X, Y, Z be nonempty sets. Then:*
(a) $X \sim X$.
(b) If $X \sim Y$, then $Y \sim X$.
(c) If $X \sim Y$ and $Y \sim Z$, then $X \sim Z$.

Lemma 1.15. *Let X and Y be two nonempty sets. If $Y \subset X$ and there exist a one-to-one function $f : X \to Y$, then $X \sim Y$.*

Proof. If $X = Y$, then evidently $X \sim Y$. If $X \neq Y$, then let $X_1 = X \setminus Y \neq \emptyset$. For every $n \in \mathbb{N}$ define $X_{n+1} = f(X_n)$ and let

$$\tilde{X} = \bigcup_{n=1}^{\infty} X_n \,.$$

We claim that $f(\tilde{X}) \subset \tilde{X}$. Indeed, if $x \in \tilde{X}$, than there exist $n \in \mathbb{N}$ such that $x \in X_n$. Then, by definition, $f(x) \subset X_{n+1} \subset \tilde{X}$.

Define the function $\tilde{f} : X \to Y$ by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in \tilde{X} \\ x & \text{if } x \notin \tilde{X}. \end{cases} \tag{1.1}$$

First, let's show that $\tilde{f}$ is one-to-one. Let $x_1, x_2 \in X$ such that $\tilde{f}(x_1) = \tilde{f}(x_2)$. If $x_1 \in \tilde{X}$ and $x_2 \notin \tilde{X}$, then $f(x_1) = x_2 \in \tilde{X}$, which is contradiction. Hence, both x_1 and x_2 must be either in $\tilde{X}$ or in its complement. Now $x_1 = x_2$ follows from the one-to-one property of f and of the identity function. To show the onto property, let $y \in Y$. If $y \in \tilde{X}$, then use the fact that

$$\tilde{X} \cap Y \subset \bigcup_{n=2}^{\infty} X_n$$

to obtain some $n \geq 2$ such that $y \in X_n$. Then, by definition, there exists $x \in X_{n-1}$ such that $\tilde{f}(x) = f(x) = y$. If $y \notin \tilde{X}$, then $y = \tilde{f}(y)$, so y is in the range of $\tilde{f}$. Therefore, we constructed a function $\tilde{f} : X \to Y$, which is one-to-one and onto, and hence $X \sim Y$.

□

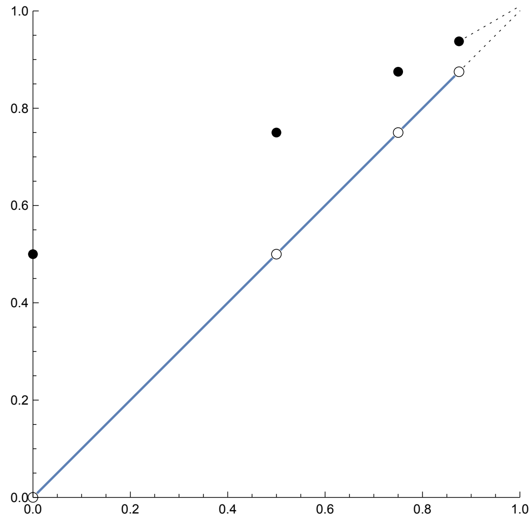
Example 1.16. Consider $X = [0,1)$ and $Y = (0,1)$. The function $f : [0,1) \to (0,1)$ defined by $f(x) = \frac{x+1}{2}$ is one-to-one. Let's construct the one-to-one and onto function $\tilde{f} : [0,1) \to (0,1)$.

Start with $X_1 = X \setminus Y = \{0\}$. Then

$$\begin{aligned}
&X_2 = \{f(0)\} = \Big\{\frac{1}{2}\Big\} \\
&X_3 = \Big\{f\Big(\frac{1}{2}\Big)\Big\} = \Big\{\frac{3}{4}\Big\} \\
&\ldots\ldots \\
&X_{n+1} = \Big\{f\Big(\frac{2^{n-1}-1}{2^{n-1}}\Big)\Big\} = \Big\{\frac{2^n-1}{2^n}\Big\} \\
&\ldots\ldots
\end{aligned}$$

Hence,

$$\tilde{f}(x) = \begin{cases} \frac{2^n-1}{2^n} & \text{if } x = \frac{2^{n-1}-1}{2^{n-1}}, \; n \in \mathbb{N} \\ \\ x & \text{otherwise} \end{cases} .$$



Theorem 1.17. *(The Cantor-Bernstein theorem)*
Let X and Y be two nonempty sets. If there exist two one-to-one functions $g : X \to Y$ and $h : Y \to X$, then $X \sim Y$.

Proof. The function $f = h \circ g : X \to h(Y) \subset X$ is one-to-one as a composition of two one-to-one functions. We can use the proof of Lemma 1.15 to obtain a function $\tilde{f} : X \to h(Y)$, which is one-to-one and onto. By restricting the codomain of h from X to $h(Y)$, we obtain a function $\tilde{h} : Y \to h(Y)$, which is one-to-one and onto and hence has an inverse function $(\tilde{h})^{-1} : h(Y) \to Y$. Define the function $F : X \to Y$, by $F = (\tilde{h})^{-1} \circ \tilde{f}$, which is one-to-one and onto as a composition of two functions with these properties.

□

Definition 1.18.
(a) We say that a set X is finite if there exists $n \in \mathbb{N}$ such that $X \sim \{1, 2, ..., n\}$.
(b) We say that a set X is infinite if it is not finite.
(c) We say that an infinite set X is countable if $X \sim \mathbb{N}$.
(d) We say that a set X is at most countable if it is finite or countable.
(e) We say that an infinite set X is uncountable if it is not countable.

For the most common infinite sets we have the following characterizations.

Proposition 1.19.
(a) $\mathbb{N}$, $\mathbb{Z}$ *and* $\mathbb{Q}$ *are countable.*
(b) $[0, 1]$ *is uncountable.*
(c) $[0, 1] \sim \mathbb{R}$.
(d) $\mathbb{R}$ *is uncountable.*
(e) $\mathbb{I}$ *is uncountable.*

Proof. (a) It is evident that $\mathbb{N} \sim \mathbb{N}$ and in Example 1.3.1 we have shown that $\mathbb{N} \sim \mathbb{Z}$. Hence, $\mathbb{N}$ and $\mathbb{Z}$ are countable. To prove that $\mathbb{Q}$ is countable, let us write

$$\mathbb{Q} = \bigcup_{n=1}^{\infty} X_n,$$

where the sets $X_n = \{x_{nk}\}$ are defined in the following way.

Let $X_1 = \mathbb{Z} = \{0, 1, -1, 2, -2, ...\} = \{x_{1k}\}$, with the understanding that x_{1k} denotes the element in the k^{th} position.

Let $X_2 = \{\frac{1}{2}, -\frac{1}{2}, \frac{2}{2}, -\frac{2}{2}, \frac{3}{2}, -\frac{3}{2}, ...\} \setminus X_1 = \{x_{2k}\}$, which means that from the set of all fractions with denominator 2 we delete the numbers which are already present in X_1.

In a similar way, $X_n = \{x_{nk}\}$ will be set of all fractions with denominator n, from which we delete the numbers which are already present in $X_1 \cup ... \cup X_{n-1}$.

Define a function $f : \mathbb{Q} \to \mathbb{N}$, by $f(x_{nk}) = 2^n 3^k$. By the uniqueness of the prime factorization of the natural numbers, the function f is one-to-one and hence, by Lemma 1.15, we obtain that $\mathbb{Q} \sim \mathbb{N}$.

(b) Assume that $[0, 1]$ is countable. Then there exists a function $f : \mathbb{N} \to [0, 1]$, which is one-to-one and onto. Therefore, we can write $[0, 1] = \{x_1, x_2, ..., x_n, ...\}$, where $x_n = f(n)$. Using the decimal number system, we can write each number $x \in [0, 1]$ as a sequence formed by the numbers $0, 1, \cdots, 9$. Hence,

$$\begin{aligned} x_1 &= 0.x_{11}x_{12}...x_{1k}... \\ x_2 &= 0.x_{21}x_{22}...x_{2k}... \\ &...... \\ x_n &= 0.x_{n1}x_{n2}...x_{nk}... \\ &......, \end{aligned} \tag{1.2}$$

where each $x_{nk} \in \{0, 1, , \ldots, 9\}$. Define a number $c = 0.c_1c_2...c_n...$ such that for each $n \in \mathbb{N}$ we have

$$c_n = \begin{cases} 3 & \text{if } x_{nn} = 5 \\ 5 & \text{if } x_{nn} \neq 5. \end{cases}$$

By definition, $c \in [0, 1]$, but also it is different than any x_n, which contradicts the fact the the full interval $[0, 1]$ was listed as the sequence $\{x_1, x_2, ...\}$.

Note that two different decimal forms in (1.2) might give the same number only if one has a terminating decimal form and the other has only numbers 9 after the first position where the decimal numbers are different.

(c) The function $f : [0, 1] \to (0, 1)$, defined by $f(x) = \frac{1}{2}x + \frac{1}{4}$ is one-to-one, hence by Lemma 1.15, $[0, 1] \sim (0, 1)$. By Example 1.3.1, we know $(0, 1) \sim \mathbb{R}$, so Proposition 1.3.1 gives $[0, 1] \sim \mathbb{R}$.

(d) $\mathbb{R}$ is uncountable, as a consequence of (b) and (c).

(e) Assume that $\mathbb{I}$ is countable. Then, as a union of two countable sets, $\mathbb{R} = \mathbb{Q} \cup \mathbb{I}$ is countable, which contradicts (d). Therefore, $\mathbb{I}$ is uncountable.
Note: To prove that the union of two countable sets is countable is left as an exercise.

$\square$

Definition 1.20. Let X be a set. We define the power set of X (denoted by 2^X), as the collection of all subsets of X, including the empty set.

In general, we could define Y^X as the set of all functions $f : X \to Y$. In this sens, 2^X is the the set of all functions from X into a set of 2 elements. Therefore, we can also define

$$2^X = \{\text{all functions } \; f : X \to \{0, 1\}\}.$$

Example 1.21.
(a) Let $X = \emptyset$. Then $2^X = \{\emptyset\}$.
(b) Let $X = \{1\}$. Then $2^X = \{\emptyset, \{1\}\}$
(c) Let $X = \{1, 2\}$. Then $2^X = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

Theorem 1.22. *(Cantor's Theorem)*
For any set X, the power set 2^X is not equivalent to X.

Proof. If $X = \emptyset$, then 2^X, which is a set of one element, is not equivalent to the empty set.

If $X \neq \emptyset$, then we can define a one-to-one function $f : X \to 2^X$, by $f(x) = \{x\}$, for every $x \in X$. Therefore, we have to show that there is no onto function from X to 2^X.

Let us assume, that there is an onto function $g : X \to 2^X$. Define

$$X_0 = \{x \in X \mid x \notin g(x)\}\,.$$

Because of g is onto, there exists $x_0 \in X$, such that $g(x_0) = X_0$. Regarding the x_0, we have only two options.

If $x_0 \in X_0$, then by definition of X_0, we have $x_0 \notin g(x_0) = X_0$, which leads to a contradiction $x_0 \in X_0$.

If $x_0 \notin X_0$, the again by definition of X_0, we have $x_0 \in g(x_0) = X_0$, which leads again to a contradiction.

As a conclusion, we obtain that the assumption that g is onto, cannot be true.

$\square$

Remark 1.23. Cantor's Theorem shows that the power set 2^X is "larger" than X. The meaning of "larger" needs explanation, because $\mathbb{Z}$ is not "larger" that $\mathbb{N}$, even though $\mathbb{Z}$ has seemingly more elements than $\mathbb{N}$.

Definition 1.24. Let X and Y be two sets.
(a) We say that X and Y have the same cardinality, $\text{card}(X) = \text{card}(Y)$, if $X \sim Y$.
(b) We say that that $\text{card}(X) \leq \text{card}(Y)$, if there exist an one-to-one function $f : X \to Y$.
(c) We say that $\text{card}(X) < \text{card}(Y)$, if $\text{card}(X) \leq \text{card}(Y)$ and $\text{card}(X) \neq \text{card}(Y)$.

Based on the definition of cardinality , we adopt the following notations:

(i) $\text{card}(X) = n$, if $X \sim \{1, 2, ..., n\}$.

(ii) $\text{card}(\mathbb{N}) = \aleph_0$.

(iii) $\text{card}(\mathbb{R}) = \mathfrak{c}$.

Therefore, we can write:

$$\text{card}(\{1, 2, ..., n\}) < \text{card}(\mathbb{N}) < \text{card}(\mathbb{R})\,.$$

1.3.1 Some naïve remarks on set theory

Without proper settings, set theory is known to contain paradoxes. Usually in Analysis we will use an informal and seemingly logical language regarding sets and therefore, we don't need to present an axiomatic set theory based on the Zermelo-Fraenkel axioms. However, three things are worth mentioning. The first two will only show what type of paradoxes we could run into, but we will use later the third one.

First, based on Theorem 1.22, it is unlikely to have a universal set, that contains everything. The power set of this universal set, should be something that it is even larger than the universal set.

Second, sets should not contain themselves as elements. If we accept that such sets exist, then consider $X = \{\text{sets } A \mid A \notin A\}$. Then for X we have $X \in X$ or $X \notin X$. In both cases we are lead to contradiction.

Third is the Axiom of Choice .

Axiom 1. (The Axiom of Choice) Given any collection of sets, it is possible to select one element of each set.

The Axiom of Choice is widely accepted for use in Analysis, however, occasionally it leads to some strange results. For example, it implies that any set can be well-ordered, including an uncountable set like $\mathbb{R}$. Or, a solid three dimensional ball can be decomposed in finitely many disjoint (not "measurable") pieces and rearranged into a ball twice the radius.

We will also assume the Continuum Hypothesis.

Axiom 2. (The Continuum Hypothesis) There is no set X such that $\aleph_0 < \text{card}(X) < \mathfrak{c}$.

Exercises

Exercise 1.13. Prove Proposition 1.14.

Exercise 1.14. Prove that the union of two countable sets is countable.

Exercise 1.15. Prove that union of countably many countable sets is countable.

Exercise 1.16. Let A be the collection of all sequences of the digits 0 and 1, for which the number of the digits 1 is finite. Show that A is countable.

Exercise 1.17. Let X be the set of all numbers from $[0, 1]$, whose decimal expansion contains only the digits 3 and 5. Is X countable or uncountable?

Exercise 1.18. Construct a one-to-one and onto function $f : [0, 1] \to (0, 1)$. Graph this function.

Exercise 1.19. Show that if $\text{card}(X) = n$, then $\text{card}(2^X) = 2^n$.

Exercise 1.20. Show that $2^{\mathbb{N}} \sim \mathbb{R}$, which can be also written as $2^{\aleph_0} = \mathfrak{c}$.

Exercise 1.21. Show that $\text{card}(\mathbb{I}) = \mathfrak{c}$.

1.4 Axioms and ordering in the set of the real numbers

In this section we list the fundamental properties of real numbers. Since we do not want to engage in a lengthy discussion regarding the construction of real numbers, we will present the fundamental properties as axioms.

We accept that the set of real numbers $\mathbb{R}$ exists and that for any $x, y, z \in \mathbb{R}$ the following properties of addition, multiplication and order hold.

Axiom 3. (The Field Axioms)

(a1) $x + y = y + x$. (Commutative property for addition)

(a2) $x + (y + z) = (x + y) + z$. (Associative property for addition)

(a3) $0 + x = x$. (Identity element for addition)

(a4) $x + (-x) = 0$. (Inverse element for addition)

(m1) $xy = yx$. (Commutative property for multiplication)

(m2) $x(yz) = (xy)z$. (Associative property for multiplication).

(m3) $1x = x$. (Identity element for multiplication)

(m4) If $x \neq 0$, then there exists $x_{-1} = \frac{1}{x}$ such that $x\,\frac{1}{x} = 1$. (Inverse element for multiplication.)

(d1) $x(y + z) = xy + xz$. (Distributive property of multiplication with respect to the addition).

Axiom 4. (The Order Axioms)

(o1) $x < y$ if and only if $y - x > 0$.

(o2) If $x < y$, then $x + z < y + z$.

(o3) If $x < y$ and $z > 0$, then $xz < yz$.

With the Field and Order axioms, $\mathbb{R}$ can be regarded as an ordered field.

We continue with the definitions of lower and upper bounds.

Definition 1.25. Let $\emptyset \neq A \subset \mathbb{R}$.

(1) We say that $\alpha \in \mathbb{R}$ is a lower bound for A, if $\alpha \leq a$, for any $a \in A$.

(2) We say that A is bounded from below if there exists a lower bound for A.

(3) We say that $\beta \in \mathbb{R}$ is an upper bound for A, if $a \leq \beta$, for any $a \in A$.

(4) We say that A is bounded from above if there exists an upper bound for A.

(5) We say that A is bounded if there exist a lower bound and an upper bound for A.

Definition 1.26. Let $A \subset \mathbb{R}$ be bounded from below. We define $\inf A$ to be a real number with the following properties:

(1) $\inf A$ is a lower bound for A.

(2) For any lower bound α for A, we have $\alpha \leq \inf A$.

The $\inf A$ is also called the greatest lower bound.

The next proposition gives a characterization of $\inf A$, and its proof is left as an exercise.

Proposition 1.27. *Let $A \subset \mathbb{R}$ be bounded from below. Then $\alpha^* = \inf A$ if and only if the following conditions hold:*

(i) α^* is a lower bound for A.

(ii) For any $\varepsilon > 0$ there exists $a_\varepsilon \in A$ such that $\alpha^* \leq a_\varepsilon < \alpha^* + \varepsilon$.

Definition 1.28. Let $A \subset \mathbb{R}$ be bounded from above. We define $\sup A$ to be a real number with the following properties:

(1) $\sup A$ is an upper bound for A.

(2) For any upper bound β for A, we have $\sup A \leq \beta$.

The $\sup A$ is also called the least upper bound.

The next proposition gives a characterization of $\sup A$, and its proof is left as an exercise.

Proposition 1.29. *Let $A \subset X$ be bounded from above. Then $\beta^* = \sup A$ if and only if the following conditions hold:*

(i) β^* is an upper bound for A.

(ii) For any $\varepsilon > 0$ there exists $b_\varepsilon \in A$ such that $\beta^* - \varepsilon < b_\varepsilon \leq \beta^*$.

The Field and Order axioms don't imply the existence of supremum and infimum of sets. We need the following axiom for this purpose.

Axiom 5. (The Dedekind-completeness of $\mathbb{R}$.)
For every nonempty $A \subset \mathbb{R}$, which is bounded from above, $\sup A$ exists and for every nonempty $B \subset \mathbb{R}$, which is bounded from below, $\inf B$ exists.

The Dedekind completeness is also called the least upper bound property or greatest lower bound property. For simplicity, we chose to include in the axiom the existence of both supremum and infimum, however assuming only the existence of one of them implies the existence of the other.

Theorem 1.30. (*The well-ordering property of natural numbers.*)
Any nonempty set of $\mathbb{N}$ *has a smallest element.*

Proof. Let $\emptyset \neq A \subset \mathbb{N}$. It follows that, as for a subset of $\mathbb{R}$, $0 \in \mathbb{R}$ is a lower bound for A. Hence, by the Dedekind completeness, $\inf A$ exists. Evidently, $\inf A \in A$, because otherwise the interval $(\inf A, \inf A + 1)$ must contain infinitely elements of A (see Exercise 1.24), which is impossible for a subset of $\mathbb{N}$.

$\square$

Theorem 1.31. (*The Archimedean property of* $\mathbb{R}$.)
If $x, y \in \mathbb{R}$ *and* $x > 0$*, then there exists* $n \in \mathbb{N}$ *such that* $xn > y$.

Proof. If $y \leq 0$, then $n = 1$ satisfies the required inequality.

If $0 < y \leq x$, then $n = 2$ satisfies the required inequality.

If $x < y$, let's begin a proof by contradiction, assuming that for all $n \in \mathbb{N}$ we have $xn \leq y$. Define $A = \{xn \mid n \in \mathbb{N}\}$. The set A is bounded above because y is an upper bound. By the Dedekind completeness, $\sup A$ exists. Since $x > 0$, the number $\sup A - x$ cannot be an upper bound, so there exists $n_0 \in \mathbb{N}$ such that $\sup A - x < n_0 x$. From here it follows that $\sup A < (n_0 + 1)x \in A$, which is a contradiction with the fact that $\sup A$ is an upper bound for A. Therefore, the assumption that $xn \leq y$ for all $n \in \mathbb{N}$ is false and hence there exists $n \in \mathbb{N}$ such that $xn > y$.

$\square$

If we use the Archimedean property for $x = \varepsilon$ and $y = 1$, we get the following result, which will be important later for convergence of sequences.

Corollary 1.32. *For every* $\varepsilon > 0$ *there exists* $n \in \mathbb{N}$ *such that* $\frac{1}{n} < \varepsilon$.

Theorem 1.33. (*The density property of rational numbers.*)

If $x, y \in \mathbb{R}$ and $x < y$, then there exists $q \in \mathbb{Q}$ such that $x < q < y$.

Proof. It is enough to show the proof for $0 < x < y$. Since $y - x > 0$, we can use the corollary from above to obtain $n_1 \in \mathbb{N}$ such that $\frac{1}{n_1} < y - x$. Consider the set $A = \{m \in \mathbb{N} \mid \frac{m}{n_1} \geq y\}$. By the Archimedean property, the set A is nonempty. Hence by the well-ordering of $\mathbb{N}$, there exists $m_1 \in A$ such that $m_1 = \inf A$. Note that $m_1 \geq 2$.

Then, $m_1 - 1 \notin A$, so

$$y > \frac{m_1 - 1}{n_1} = \frac{m_1}{n_1} - \frac{1}{n_1} > y - (y - x) = x.$$

Therefore, we can choose $q = \frac{m_1 - 1}{n_1}$.

□

Exercises

Exercise 1.22. Prove Proposition 1.27.

Exercise 1.23. Prove Proposition 1.29.

Exercise 1.24. (i) Let $A \subset \mathbb{R}$ be a nonempty set, which is bounded from above. Show that if $\sup A \notin A$, then for all $\varepsilon > 0$ the open interval $(\sup A - \varepsilon, \sup A)$ contains infinitely many terms of A.
(ii) Let $B \subset \mathbb{R}$ be a nonempty set, which is bounded from below. Show that if $\inf B \notin B$, then for all $\varepsilon > 0$ the open interval $(\inf B, \inf B + \varepsilon)$ contains infinitely many terms of B.

Exercise 1.25. Let $A, B \subset \mathbb{R}$ be nonempty, bounded sets and let $c \in \mathbb{R}$. Define the following sets:

$$\begin{aligned} A + B &= \{a + b \mid a \in A,\ b \in B\} \\ A - B &= \{a - b \mid a \in A,\ b \in B\} \\ A \cdot B &= \{ab \mid a \in A,\ b \in B\} \\ cA &= \{ca \mid a \in A\}. \end{aligned}$$

Prove that:

(1) $\inf(A + B) = \inf(A) + \inf(B)$.
(2) $\sup(A + B) = \sup(A) + \sup(B)$.
(3) $\sup(-A) = -\inf(A)$.
(4) $\inf(-A) = -\sup(A)$.
(5) $\sup(A - B) = \sup(A) - \inf(B)$.
(6) $\inf(A - B) = \inf(A) - \sup(B)$.
(7) $\sup(cA) = c \sup(A)$ if $c > 0$.
(8) $\inf(cA) = c \inf(A)$ if $c > 0$.
(9) $\sup(cA) = c \inf(A)$ if $c < 0$.
(10) $\inf(cA) = c \sup(A)$ if $c < 0$.
(11) Is it true that $\sup(A \cdot B) = \sup(A) \cdot \sup(B)$?

Exercise 1.26. State and prove the density property of the irrational numbers.

Chapter 2

Metric and measure

2.1 Metric Spaces

Definition 2.1. A metric space is defined by a pair (X, d), where X is a nonempty set and $d : X \times X \to [0, \infty)$ is a function which statisfies the following properties for every $x, y, z \in X$:

(a) $d(x, y) = 0$ if and only if $x = y$.

(b) $d(x, y) = d(y, x)$.

(c) $d(x, z) \leq d(x, y) + d(y, z)$.

The property (b) is called symmetry and (c) is called the triangle inequality. We can have multiple metrics for the same space X. If in a certain context there is only one metric used, and there is no risk of confusion, we can refer to the metric space as only X.

For the following examples of metric spaces, we leave the proofs, that the metrics satisfy the three properties, as exercises.

Example 2.2.

(1) On any nonempty set X we can define the discrete metric:

$$d_0(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y\,. \end{cases}$$

(2) On $\mathbb{R}$ we define the metric $d(x, y) = |x - y|$.

(3) On $\mathbb{R}^m$ we use the notation $x = (x_1, ..., x_m)$, $y = (y_1, ..., y_m)$ and we define the metric

$$d(x, y) = \sqrt{\sum_{i=1}^{m} (x_i - y_i)^2}.$$

(4) Let $p \geq 1$ and on $\mathbb{R}^m$ we define the metric

$$d_p(x, y) = \left(\sum_{i=1}^{m} |x_i - y_i|^p \right)^{\frac{1}{p}} .$$

(5) On $\mathbb{R}^m$ we define the metric

$$d_\infty(x, y) = \max_{1 \leq i \leq m} |x_i - y_i|.$$

The metrics defined in (2) and (3) are called Euclidean metrics and by comparing them to (4), we see that $d = d_2$.

We can also define the Euclidean norm of $x \in \mathbb{R}^m$, which is

$$||x|| = d(x, 0) = \sqrt{\sum_{i=1}^{m} (x_i)^2} .$$

We can see that $d(x, y) = ||x - y||$.

More examples of metric spaces will be presented in the chapters covering sequences of functions.

Exercises

Exercise 2.1. Show that for all $x, y \in \mathbb{R}$ we have

$$\Big| |x| - |y| \Big| \leq |x - y| .$$

Exercise 2.2. (Young's inequality) Let $a, b \geq 0$, $p, q > 1$ and assume that $\frac{1}{p} + \frac{1}{q} = 1$. Prove that the following inequality holds:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} .$$

Exercise 2.3. (Hölder's inequality) Let $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Prove that for any $x, y \in \mathbb{R}^m$ the following inequality holds:

$$\sum_{i=1}^{m} |x_i y_i| \leq \left(\sum_{i=1}^{m} |x_i|^p \right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{m} |y_i|^q \right)^{\frac{1}{q}} .$$

Exercise 2.4. (Minkowski's inequality) Let $p \geq 1$. Prove that for any $x, y \in \mathbb{R}^m$ the following inequality holds:

$$\left(\sum_{i=1}^{m} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{m} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{m} |y_i|^p\right)^{\frac{1}{p}}.$$

Exercise 2.5. Prove that the spaces presented in Example 2.2 are metric spaces.

2.2 Open and closed sets

Definition 2.3. Let X be a metric space, $x_0 \in X$ and $r > 0$.
We define the open ball centered at x_0 and radius r by

$$B_r(x_0) = \{x \in X \mid d(x_0, x) < r\}.$$

We define the closed ball centered at x_0 and radius r by

$$B_r[x_0] = \{x \in X \mid d(x_0, x) \leq r\}.$$

With the help of the open balls we can state the following definitions.

Definition 2.4. Let X be a metric space.

(a) A set $A \subset X$ is called an open set if for every $x \in A$ there exists $r > 0$ such that $B_r(x) \subset X$.

(b) A set $B \subset X$ is called a closed set if B^c is open.

(c) A set A is called a neighborhood of the point $x \in A$ if there exists $r > 0$ such that $B_r(x) \subset A$.

(d) A point $x \in A$ is called an interior point of A if there exists $r > 0$ such that $B_r(x) \subset A$. The set of all interior points of A is called the interior of A.

(e) If $A \subset B \subset X$, we say that A is open relative to B, if there exists an open set $C \subset X$ such that $A = B \cap C$.

From the fact that $(A^c)^c = A$, we can deduce the following proposition.

Proposition 2.5. *Let X be a metric space. A set $A \subset X$ is open if and only if A^c is closed.*

The next proposition shows that an open ball is, indeed, an open set.

Proposition 2.6. *Let X be a metric space, $x_0 \in X$ and $r > 0$. Then the open ball $B_r(x_0)$ is, indeed, an open set.*

Proof. Let us choose any $x \in B_r(x_0)$. Then, $d(x, x_0) < r$. Choose $r_x = r - d(x, x_0)$. Then, $B_{r_x}(x) \subset B_r(x_0)$, because for any $y \in B_{r_x}(x)$ we have $d(y, x) < r_x$ and hence

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < r - d(x, x_0) + d(x, x_0) = r.$$

□

The proof of the next proposition is left as an exercise.

Proposition 2.7. *Let X be a metric space, $x_0 \in X$ and $r > 0$. Then the closed ball $B_r[x_0]$ is, indeed, a closed set.*

Example 2.8.

1. Let $X = \mathbb{R}$, $x_0 \in \mathbb{R}$ and $r > 0$. The open ball $B_r(x_0)$ is the open interval $(x_0 - r, x_0 + r)$. The closed ball $B_r[x_0]$ is the closed interval $[x_0 - r, x_0 + r]$.

2. Any open interval $(a, b) \subset \mathbb{R}$ is an open set. Indeed, for any $x \in (a, b)$ define $r = \max\{x - a, b - x\}$. Then, we obtain $(x - r, x + r) \subset (a, b)$. Notice that r depends on x.

3. The intervals $(-\infty, a)$ and (a, ∞) are open, too. Similarly, to the previous example, for each x from these intervals we can define $r = |x - a|$.

4. Any closed interval $[a, b]$ is a closed set, because

$$[a, b]^c = (-\infty, a) \bigcup (b, \infty),$$

which is a union of two open sets. In the next theorem we will prove that such a union is open.

5. Also, the intervals $(-\infty, a]$ and $[a, \infty)$ are closed, because their complements are open.

6. Any subset of $\mathbb{R}$ made of finitely many points is closed, because its complement is a union of finitely many open intervals.

7. $\mathbb{N}$ and $\mathbb{Z}$, as subsets of $\mathbb{R}$, are closed, because their complements are unions of open sets.

8. The density properties of $\mathbb{Q}$ and $\mathbb{I}$ show that they cannot be open. Also, as they are complements of each other, they are not closed either.

9. The half-closed intervals $(a, b]$ or $[a, b)$ are not open, nor closed in $X = \mathbb{R}$. However, if we consider $X = [0, 1]$ and any $0 < r < 1$, then $B_r(0) = [0, r)$ is an open ball.
Also, if we return to $X = \mathbb{R}$, we can say that $[0, r)$ is open relative of $[0, 1]$.

The next two theorems give the properties of open and closed sets.

Theorem 2.9. *Let X be a metric space. Then the following properties hold:*
(i) $\emptyset$ *and* X *are open sets.*
(ii) *If* $\{A_i\}_{i \in I}$ *is any collection of open sets, then* $\bigcup_i A_i$ *is open.*
(iii) *If* $A_1, ..., A_n$ *is a finite collection of open sets, then* $\bigcap_{i=1}^n A_i$ *is open.*

Proof. (i) The empty set is open because the definition is vacuously true, while the space X is open because the whole space contains any ball $B_r(x)$.

(ii) Let $x \in \bigcup_i A_i$. Then there exists $i_x \in I$ such that $x \in A_{i_x}$. The set A_{i_x} is open, so there exists $r_x > 0$ such that $B_{r_x}(x) \subset A_{i_x}$. Hence,

$$B_{r_x}(x) \subset A_{i_x} \subset \bigcup_i A_i \,,$$

and therefore $\bigcup_i A_i$ is open.

(iii) Let $x \in \bigcap_{i=1}^n A_i$. Then, $x \in A_i$ for every $1 \leq i \leq n$. Therefore, for each i there exist $r_i > 0$ such that $B_{r_i}(x) \subset A_i$. Let $r = \min\{r_1, ..., r_n\}$. As the smallest of finitely many positive numbers, we have $r > 0$, and hence for each i holds

$$B_r(x) \subset B_{r_i}(x) \subset A_i \,.$$

It follows that,

$$B_r(x) \subset \bigcap_{i=1}^n A_i \,,$$

and therefore, $\bigcap_{i=1}^n A_i$ is open.

□

Note that (iii) is not true for infinitely many sets. Let $A_i = (0, 1 + \frac{1}{i})$ for all $i \in \mathbb{N}$. Then, while each A_i is open,

$$\bigcap_{i=1}^{\infty} \left(0, 1 + \frac{1}{i}\right) = (0, 1] \,,$$

which is not open.

Theorem 2.10. *Let X be a metric space. Then the following properties hold:*
(i) $\emptyset$ *and X are closed sets.*
(ii) *If $\{B_i\}$ is any collection of closed sets, then $\bigcap_i B_i$ is closed.*
(iii) *If $B_1, ..., B_n$ is a finite collection of closed sets, then $\bigcup_{i=1}^n B_i$ is closed.*

Proof. The proof follows from the properties of complement sets.

(i) The empty set and X are closed as complements of open sets.

(ii) If B_i is closed, then $(B_i)^c$ is open and by the De Morgan law

$$\left(\bigcap_i B_i\right)^c = \bigcup_i (B_i)^c \,,$$

is open as union of open sets. Therefore, $\bigcap_i B_i$ is closed.

(iii) The proof, which is similar to (ii), uses again the De Morgan law

$$\left(\bigcup_{i=1}^n B_i\right)^c = \bigcap_{i=1}^n (B_i)^c \,.$$

□

Definition 2.11. Let X be a metric space and $A \subset X$.

(1) The closure of A, denoted by $\overline{A}$, is defined as the intersection of all closed sets containing A.

(2) A point $x \in X$ is a limit point point of A, if for every $r > 0$ we have

$$A \bigcap \Big(B_r(x) \setminus \{x\} \Big) \neq \emptyset.$$

The set of limit points of A will be denoted by A'.

(3) If $x \in A$ is not a limit point of A, then it is called an isolated point of A.

Proposition 2.12. *Let X be a metric space and $A \subset X$ be a closed set. If $x \in X$ is a limit point of A, then $x \in A$.*

Proof. Let us assume that, by contradiction, $x \notin A$. Then $x \in A^c$ and A^c is open, so there exists $r > 0$ such that $B_r(x) \subset A^c$. Therefore, $B_r(x) \bigcap A = \emptyset$, which contradicts the fact that x is a limit point of A. $\square$

Example 2.13. (a) If $X = \mathbb{R}$ and $A = (-1, 1)$, then $\overline{A} = [-1, 1]$. Also, $A' = [-1, 1]$.

(b) If $X = \mathbb{R}$ and $A = \{-1, 1\}$, then $\overline{A} = \{-1, 1\}$. Also, $A' = \emptyset$.

(c) If $X = [0, 1]$ and $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$, then $\overline{A} = A \bigcup \{0\}$ and $A' = \{0\}$.

(d) (c) If $X = (0, 1]$ and $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$, then $\overline{A} = A$ and $A' = \emptyset$.

Theorem 2.14. *Let X be a metric space and $A \subset X$. Then A is closed if and only if A contains all its limit points.*

Proof. First, let us assume that A is closed and prove that A contains all its limit points. This is exactly Proposition 2.12.

Second, we want to prove that if A contains all its limit points, then A is closed. Let us assume, by contradiction, that A is not closed. Then A^c is not open, so there exists $x \in A^c$ such that for all $r > 0$,

$$B_r(x) \not\subset A^c\,.$$

Therefore,

$$A \bigcap B_r(x) \neq \emptyset\,,$$

and because of $x \notin A$, it follows that

$$A \bigcap (B_r(x) \setminus \{x\}) \neq \emptyset.$$

Hence, x is a limit point of A, but not in A, which contradicts the assumption that A contains all its limit points.

$\square$

Definition 2.15. Let X be a metric space.

(1) A set $A \subset X$ is called a dense subset of X, if $\overline{A} = X$.

(2) The metric space X is called a separable metric space, if there exists a countable and dense subset.

Proposition 2.16. *$\mathbb{Q}$ is a dense subset of $\mathbb{R}$.*

Proof. We want to show that every $x \in \mathbb{R}$ is a limit point of $\mathbb{Q}$. Let $x \in \mathbb{R}$ and $r > 0$ be arbitrary numbers. If we use Theorem 1.33 for x and $y = x + r$, we find a rational number q such that $x < q < x + r$. This proves that x is a limit point of $\mathbb{Q}$.

□

Exercises

Exercise 2.6. Let X be a metric space and $A \subset B \subset X$. Show that $\overline{A} \subset \overline{B}$.

Exercise 2.7. Let X be a metric space and $A \subset X$. Show that $\overline{A} = A \cup A'$.

Exercise 2.8. Let X be a metric space and $A_i \subset X$, for all $i \in \mathbb{N}$.

(1) Prove that

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i} \ .$$

(2) Is it true that

$$\overline{\bigcup_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \overline{A_i} \ ?$$

Exercise 2.9. Let X be a metric space and $A \subset X$. Prove that A' is a closed set.

Exercise 2.10. Show that every open open subset of $\mathbb{R}$ is a finite or countable union of open intervals.

Exercise 2.11. Let $\emptyset \neq A \subset \mathbb{R}$. Assume that all the points of A are isolated. Show that A is at most countable.

Exercise 2.12. Let $A \subset \mathbb{R}$ be an uncountable set. Show that A has limit points. Is this statement true if we assume that A is countable?

Exercise 2.13. Show that $\overline{\mathbb{I}} = \mathbb{R}$.

Exercise 2.14. Is it true in any metric space that $\overline{B_r(x)} = B_r[x]$?

2.3 Sequences

Definition 2.17. Let X be a set. By a sequence in X we understand a function $f : \mathbb{N} \to X$. Using the notation $x_n = f(n)$, we will describe a sequence by its elements in one of the following ways:

$$\{x_n\}_{n\in\mathbb{N}} \text{ or } \{x_n\}.$$

We call the range of $\{x_n\}$ the set which is the range of the function f.

Example 2.18.

- Consider $f : \mathbb{N} \to \mathbb{R}$, defined by $f(n) = \frac{1}{n}$. Then $x_n = \frac{1}{n}$ for all $n \in \mathbb{N}$, and we have the sequence $\{\frac{1}{n}\}$ with terms $1, \frac{1}{2}, \frac{1}{3}, \dots$. The range of the sequence is the set $\{1, \frac{1}{2}, \frac{1}{3}, \dots\}$.
- Consider $f : \mathbb{N} \to \mathbb{R}$, defined by $f(n) = (-1)^n$. Then $x_n = (-1)^n$ for all $n \in \mathbb{N}$, and we have the sequence $\{(-1)^n\}$ with terms $-1, 1, -1, 1, \dots$. The range of the sequence is the set $\{-1, 1\}$.
- Consider $f : \mathbb{N} \to \mathbb{R}^2$, defined by $f(n) = \left(\frac{1}{n}, n\right)$. Then $x_n = (\frac{1}{n}, n)$ for all $n \in \mathbb{N}$, and we have the sequence $\{(\frac{1}{n}, n)\}$ with terms $(1,1), (\frac{1}{2}, 2), (\frac{1}{3}, 3), \dots$. The range of the sequence is the set $\{(1,1), (\frac{1}{2}, 2), (\frac{1}{3}, 3), \dots\}$.

Definition 2.19. Let $\{x_n\}$ be a sequence in a metric space X.

(i) We say that $\{x_n\}$ is convergent if there exist a point $x \in X$ with the property that for all $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that

$$d(x_n, x) < \varepsilon, \ \forall\, n > N.$$

(ii) For convergence we use one of the following notations:

$$\lim_{n\to\infty} x_n = x \quad \text{or} \quad \{x_n\} \to x.$$

(iii) We say that the sequence $\{x_n\}$ is divergent if it is not convergent.

(iv) We say that the sequence $\{x_n\}$ is bounded if there exist a point $x \in X$ and a real number $R > 0$ such that

$$d(x_n, x) \le R, \ \forall \ n \in \mathbb{N}.$$

Note that, if the sequence $\{x_n\}$ is from $\mathbb{R}$ then the convergence can be formulated in the following way: There exists $x \in \mathbb{R}$ such that for all $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that

$$|x_n - x| < \varepsilon, \ \ \forall \ n > N.$$

This inequality means that

$$x - \varepsilon < x_n < x + \varepsilon, \ \ \forall \ n > N.$$

Example 2.20.

(a) Prove that $\{\frac{1}{n}\} \to 0$. Consider any $\varepsilon > 0$. By Corollary 1.32, there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $\frac{1}{N} < \varepsilon$. Hence, for all $n > N$ we have

$$-\varepsilon < 0 < \frac{1}{n} - 0 < \frac{1}{N} < \varepsilon,$$

and this proves that $\{\frac{1}{n}\} \to 0$.

(b) Prove that $\left\{\frac{2n}{n+3}\right\} \to 2$. Consider any $\varepsilon > 0$. Observe that, it is enough to consider $0 < \varepsilon < 1$. Therefore,

$$\begin{aligned} &\left|\frac{2n}{n+3} - 2\right| < \varepsilon \\ \Leftrightarrow &\left|\frac{-6}{n+3}\right| < \varepsilon \\ \Leftrightarrow &\frac{6}{n+3} < \varepsilon \\ \Leftrightarrow &\, n > \frac{6}{\varepsilon} - 3 > 0. \end{aligned}$$

Hence, if we define $N = N(\varepsilon) = \frac{6}{\varepsilon} - 3$, then for any $n > N$ we have

$$\left|\frac{2n}{n+3} - 2\right| < \varepsilon,$$

and this proves that $\{\frac{1}{n}\} \to 0$.

(c) Prove that, if $|a| < 1$ then $\{a^n\} \to 0$. Let $\varepsilon > 0$. Without loss of generality we can consider $0 < \varepsilon < 1$. Then,

$$\begin{aligned} &|a^n| < \varepsilon \\ \Leftrightarrow &|a|^n < \varepsilon \\ \Leftrightarrow &\, n > \frac{\ln \varepsilon}{\ln |a|}, \end{aligned}$$

we can define $N = N(\varepsilon) = \frac{\ln \varepsilon}{\ln |a|} > 0$ and this proves the claim.

Proposition 2.21. *Let $\{x_n\}$ be a sequence from a metric space X. If $\{x_n\}$ is convergent, then its limit is unique.*

Proof. Assume that we have $x, y \in X$ such that $\{x_n\} \to x$ and $\{x_n\} \to y$. Let $\varepsilon > 0$. Then there exist $N_1 = N_1(\varepsilon) > 0$ and $N_2 = N_2(\varepsilon) > 0$ such that

$$d(x_n, x) < \frac{\varepsilon}{2}, \quad \forall \; n > N_1,$$

and

$$d(x_n, y) < \frac{\varepsilon}{2}, \quad \forall \; n > N_2.$$

Hence, for any $n > N = \max\{N_1, N_2\}$ we have

$$d(x, y) < d(x, x_n) + d(x_n, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore, we obtained that for any $\varepsilon > 0$ we have $d(x, y) < \varepsilon$, which implies that $d(x, y) = 0$, and hence $x = y$.

□

Proposition 2.22. *Let $\{x_n\}$ be a sequence from a metric space X. If $\{x_n\}$ is convergent, then it is bounded.*

Proof. Let us consider $\{x_n\} \to x$. For $\varepsilon = 1$ there exists $N > 0$ such that $d(x_n, x) < 1$ for all $n > N$. Let n_1 the smallest integer greater than N. Define

$$r = \max\{1, d(x_1, x), ..., d(x_{n_1 - 1}, x)\}.$$

We have shown that for all $n \in \mathbb{N}$ we have $d(x_n, x) \leq r$, which implies that the sequence is bounded.

□

Remark 2.23. Note that the reverse implication in Proposition 2.22 is not true. Indeed, the sequence $\{(-1)^n\}$ is bounded, but not convergent.

Theorem 2.24. *Let $\{x_n\}$ and $\{y_n\}$ be two sequences in $\mathbb{R}$. Suppose that $\{x_n\} \to x$ and $\{y_n\} \to y$. Then:*

(a)
$$\lim_{n\to\infty} (x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n = x + y\,.$$

(b) *For any $c \in \mathbb{R}$ we have*
$$\lim_{n\to\infty} (c x_n) = c \lim_{n\to\infty} x_n = cx\,.$$

(c)
$$\lim_{n\to\infty} (x_n \cdot y_n) = \lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n = xy\,.$$

(d) *If* $x_n \neq 0$ *for all* $n \in \mathbb{N}$ *and* $x \neq 0$, *then*

$$\lim_{n\to\infty} \frac{1}{x_n} = \frac{1}{\lim_{n\to\infty} x_n} = \frac{1}{x}\,.$$

Proof. (a) For any $\varepsilon > 0$ there exists $N_1 = N_1(\varepsilon) > 0$ and $N_2 = N_2(\varepsilon) > 0$ such that

$$|x_n - x| < \frac{\varepsilon}{2}, \;\; \forall \;\; n > N_1$$

and

$$|y_n - y| < \frac{\varepsilon}{2}, \;\; \forall \;\; n > N_2.$$

Let $N_3 = N_3(\varepsilon) = \max\{N_1, N_2\}$. Then, for all $n > N_3$ we have

$$|(x_n + y_n) - (x + y)| \leq |x_n - x| + |y_n - y| < \varepsilon\,,$$

which proves that $\{x_n + y_n\} \to x + y$.

(b) For any $\varepsilon > 0$ there exists $N_4 = N_4(\varepsilon) > 0$ such that

$$|x_n - x| < \frac{\varepsilon}{|c| + 1}, \;\; \forall \;\; n > N_4.$$

Then, for all $n > N_4$ we have

$$|c\,x_n - c\,x| = |c|\,|x_n - x| < |c|\,\frac{\varepsilon}{|c| + 1} < \varepsilon\,,$$

which proves that $\{c\,x_n\} \to c\,x$.

(c) Let $\varepsilon > 0$ be an arbitrary, but fixed number. Convergent sequences are bounded, hence there exists $M \geq 0$ such that $|y_n| \leq M$ for all $n \in \mathbb{N}$. Also, there exists $N_5 = N_5(\varepsilon) > 0$ and $N_6 = N_6(\varepsilon) > 0$ such that

$$|x_n - x| < \frac{\varepsilon}{2(M + 1)}, \;\; \forall \;\; n > N_5$$

and

$$|y_n - y| < \frac{\varepsilon}{2(|x| + 1)}, \;\; \forall \;\; n > N_6.$$

Let $N_7 = N_7(\varepsilon) = \max\{N_5, N_6\}$. Then, for all $n > N_7$ we have

$$\begin{aligned} &|(x_n y_n) - (xy)| = |x_n y_n - x y_n + x y_n - xy| \\ &\leq |y_n|\,|x_n - x| + |x|\,|y_n - y| \\ &< M \frac{\varepsilon}{2(M + 1)} + |x| \frac{\varepsilon}{2(|x| + 1)} < \varepsilon\,, \end{aligned}$$

which proves that $\{x_n y_n\} \to xy$.

(d) Choose $N_8 > 0$ such that $|x_n - x| < \frac{|x|}{2}$ for all $n > N_8$. Hence, by the properties of the absolute value,

$$|x| - |x_n| \leq |x - x_n| = |x_n - x| < \frac{|x|}{2}\,,$$

and therefore

$$|x_n| \geq \frac{|x|}{2}\,,\ \ \forall\ \ n > N_8\,.$$

For any $\varepsilon > 0$ choose $N_9 = N_9(\varepsilon) > 0$ such that

$$|x_n - x| < \frac{\varepsilon\,|x|^2}{2}\,,\ \ \forall\ \ n > N_9\,.$$

We can conclude that for all $n > N_9$ we have

$$\left|\frac{1}{x_n} - \frac{1}{x}\right| = \frac{|x_n - x|}{|x_n|\,|x|} < \frac{2}{|x|^2} \cdot \frac{\varepsilon|x|^2}{2} = \varepsilon\,.$$

This proves that $\frac{1}{x_n} \to \frac{1}{x}$.

$\square$

To generalize the above results to sequences in $\mathbb{R}^m$, let us introduce the following notations:

$$x = (x_1, ..., x_m) \ \text{ and } \ x_n = (x_{n1}, ..., x_{nm})\,.$$

Also, if not otherwise stated, in $\mathbb{R}^m$ we use the Euclidean distance

$$d(x, y) = \left(\sum_{k=1}^{m} (x_k - y_k)^2\right)^{\frac{1}{2}}\,.$$

To see more similarities between $\mathbb{R}$ and $\mathbb{R}^m$ we can use the norm of points, which is the distance from the origin:

$$||x|| = \left(\sum_{k=1}^{m} (x_k)^2\right)^{\frac{1}{2}}\,.$$

With this notation, let us remind that

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

Theorem 2.25. *Let $\{x_n\}$ be a sequence in $\mathbb{R}^m$. Then, $\{x_n\} \to x$ if and only $\{x_{nk}\} \to x_k$ for any integer $1 \leq k \leq m$.*

Proof. Assume that $\{x_n\} \to x$. For any $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that $d(x_n, x) < \varepsilon$ for any $n > N$.

Fix any integer $1 \le k \le m$. Then, from the inequality

$$|x_{nk} - x_k| \le \left(\sum_{k=1}^{m} (x_{nk} - x_k)^2 \right)^{\frac{1}{2}} < \varepsilon$$

it follows that $\{x_{nk}\} \to x_k$.

Assume now that for any integer $1 \le k \le m$ we have $\{x_{nk}\} \to x_k$. Then for any $\varepsilon > 0$ there exists $N_k = N_k(\varepsilon) > 0$ such that

$$|x_{nk} - x_k| < \frac{\varepsilon}{\sqrt{m}}, \ \ \forall \ n > N_k .$$

Therefore, for $n > N = N(\varepsilon) = \max\{N_1, ..., N_m\}$ we have

$$d(x_n, x) = \left(\sum_{k=1}^{m} (x_{nk} - x_k)^2 \right)^{\frac{1}{2}} < \left(m \frac{\varepsilon^2}{m} \right)^{\frac{1}{2}} = \varepsilon.$$

In this way we proved that $\{x_n\} \to x$.

□

We introduce now the monotone sequences.

Definition 2.26. Consider a sequence $\{x_n\}$ from $\mathbb{R}$.

(a) We say that $\{x_n\}$ is monotonically increasing if $x_n \le x_{n+1}$ for all $n \in \mathbb{N}$.

(b) We say that $\{x_n\}$ is monotonically decreasing if $x_n \ge x_{n+1}$ for all $n \in \mathbb{N}$.

Theorem 2.27. *Let $\{x_n\}$ be a sequence of real numbers which is monotonically increasing. Then $\{x_n\}$ is convergent if and only if it is bounded from above. Similarly, a monotonically decreasing sequence is convergent if and only if it is bounded from below.*

Proof. Assume that $\{x_n\}$ is monotonically increasing. If $\{x_n\}$ is convergent, then it is bounded by Proposition 2.22. To prove the reverse implication, let us assume that $\{x_n\}$ is bounded. Let $X \subset \mathbb{R}$ be the range of $\{x_n\}$. The fact that the sequence is bounded from above implies that X is bounded from above. Therefore, by the Dedeckind-completeness of $\mathbb{R}$, we get that $x = \sup X$ exists. Then, by the definition of $\sup X$, for any $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that

$$x - \varepsilon < x_{N(\varepsilon)} \le x .$$

Therefore, by the facts that the sequence is monotonically increasing and x is an upper bound for X, we get that

$$x - \varepsilon < x_{N(\varepsilon)} \le x_n \le x < x + \varepsilon, \ \ \forall \ n > N(\varepsilon) .$$

This proves that $\{x_n\} \to x$. □

We will use Theorem 2.27 to prove the following theorem.

Theorem 2.28. *(Bolzano-Weierstrass Theorem in $\mathbb{R}$)*
Let $A \subset \mathbb{R}$ be a bounded set, which contains infinitely many elements. Then there exists at least one limit point of A.

Proof. Since A is bounded, we can choose a number $M > 0$ such that $A \subset [-M, M]$. Split the interval into two halves: $[-M, M] = [-M, 0] \cup [0, M]$. By the fact that A has infinitely many elements, at least one of the halves must contains infinitely many elements of A. Denote the chosen half interval by $[a_1, b_1]$. Again, split this interval in two halves, choose one of them which contains infinitely many elements of A and denote it by $[a_2, b_2]$. Continue this process indefinitely. In this way we construct a sequence of intervals

$$[a_1, b_1] \supset [a_2, b_2] \supset \cdots \supset [a_n, b_n] \supset \dots ,$$

such that $b_n - a_n = \frac{M}{2^n}$ for all $n \in \mathbb{N}$. The sequence $\{a_n\}$ is monotonically increasing and bounded above by M, therefore it is convergent. Similarly, the sequence $\{b_n\}$ is monotonically decreasing and bounded below by $-M$, and hence it is convergent. The two sequences have the same limit, because $\{b_n - a_n\} \to 0$. Denote this common limit by a.

To show that a is a limit point of A, consider any $r > 0$. Choose $n \in \mathbb{N}$ such that $\frac{M}{2^n} \leq \frac{r}{3}$. Then,

$$a \in [a_n, b_n] \subset (a - r, a + r)\,,$$

and by the fact that $[a_n, b_n]$ contains infinitely many elements of A we get that

$$\emptyset \neq A \bigcap \Big([a_n, b_n] \setminus \{a\} \Big) \subset A \bigcap \Big((a - r, a + r) \setminus \{a\} \Big)\,.$$

Therefore, a is a limit point of A.

□

Theorem 2.29. *(Bolzano-Weierstrass Theorem in $\mathbb{R}^m$)*
Let $A \subset \mathbb{R}^m$ be a bounded set, which contains infinitely many elements. Then there exists at least one limit point of A.

Proof. The proof is almost identical to the proof in $\mathbb{R}$. The only difference is that, at each step, we have to divide closed rectangular boxes (see Definition 2.53) into 2^m equal smaller boxes and choose one in which A has infinitely many elements. See also the proof of Proposition 2.55.

□

Definition 2.30. Let $\{x_n\}$ be any sequence. Consider $n_1 < n_2 < \dots$, a sequence from $\mathbb{N}$, and denote it by $\{n_k\}_{k \in \mathbb{N}}$. Then the sequence $\{x_{n_k}\}_{k \in \mathbb{N}}$ is called a subsequence of $\{x_n\}$.

Example 2.31. Consider $x_n = (-1)^n$, for all $n \in \mathbb{N}$. Let $n_k = 2k$, for all $k \in \mathbb{N}$, which is the sequence of even numbers. Then the subsequence $\{x_{n_k}\}$ has the terms

$$x_{n_k} = x_{2k} = (-1)^{2k} = 1\,.$$

Definition 2.32. Consider a sequence $\{x_n\}$ for a metric space X. We call $x \in X$ a subsequential limit point of $\{x_n\}$, if there is a subsequence $\{x_{n_k}\}$ such that $\{x_{n_k}\} \to x$.

Example 2.33. Consider $x_n = \sin(n\frac{\pi}{2})$ for all $n \in \mathbb{N}$. The set of subsequential limit points is $\{-1, 0, 1\}$. Indeed,

$$x_{4k} = x_{4k+2} = 0, \ \ \forall \ k \in \mathbb{N},$$

$$x_{4k+1} = 1, \ \ \forall \ k \in \mathbb{N},$$

$$x_{4k+3} = -1, \ \ \forall \ k \in \mathbb{N},$$

and there are no other options.

Proposition 2.34. *Let $\{x_n\}$ be a sequence in a metric space X. Then, $\{x_n\} \to x \in X$ if and only if every subsequence is convergent to x.*

Proof. Assume that $\{x_n\}$ is convergent. Then, for any $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that

$$d(x_n, x) < \varepsilon, \ \ \forall \ n > N\,.$$

Consider any subsequence $\{x_{n_k}\}$. As $\{n_k\}$ is a strictly increasing sequence of natural numbers, there exists $K = K(\varepsilon) > 0$ such that $n_k > N$ for all $k > K$. In conclusion, for any $\varepsilon > 0$ we can choose $K = K(\varepsilon) > 0$ such that

$$d(x_{n_k}, x) < \varepsilon, \ \ \forall \ k > K,$$

and this proves that $\{x_{n_k}\} \to x$.

Assume now that every subsequence is converging to the same $x \in X$. Then the sequence, which is a subsequence of itself, must be convergent to x.

□

Lemma 2.35. *Let X be a metric space and $A \subset X$. If $x \in X$ is a limit point of A, then there exists a sequence $\{a_n\}$, with terms from A, such that $\{a_n\} \to x$*

Proof. In the definition of a limit point, let us use $r = \frac{1}{n}$ for each $n \in \mathbb{N}$. Then for each n there exists

$$a_n \in A \bigcap \Big(B_{\frac{1}{n}}(x) \setminus \{x\} \Big)\,.$$

Hence, $d(a_n, x) < \frac{1}{n}$, for all $n \in \mathbb{N}$ which proves that $\{a_n\} \to x$.

□

The next theorem is a consequence of the Bolzano-Weierstrass theorem in $\mathbb{R}$.

Theorem 2.36. *Let $\{x_n\}$ be a bounded sequence in $\mathbb{R}$. Then it has a convergent subsequence.*

Proof. Let us denote by S the range of the sequence $\{x_n\}$. Since the sequence is bounded, it follows that its range is a bounded set. If S contains only finitely many elements, then one of them, let's say $x \in \mathbb{R}$ must show up infinitely many times in the sequence. Hence we can construct a subsequence $\{x_{n_k}\}$, such that $x_{n_k} = x$ for all $k \in \mathbb{N}$. Therefore, $\{x_{n_k}\} \to x$.

If S contains infinitely many elements, then by the Bolzano-Weiertrass theorem it has a limit point $x \in \mathbb{R}$. We can use now Lemma 2.35 to obtain a sequence from S converging to x. This sequence is from the range of $\{x_n\}$, which means that it can be organized as a subsequence of $\{x_n\}$. □

For the proof of the next theorem we could use Theorem 2.29 in a similar way to the above, but let's try something different.

Theorem 2.37. *Let $\{x_n\}$ be a bounded sequence in $\mathbb{R}^m$. Then it has a convergent subsequence.*

Proof. Consider first the sequence made by the first coordinates $\{x_{n1}\}$. This is a bounded sequence of real numbers, so by Theorem 2.36, we can find a convergent subsequence $\{x_{n_k1}\}$. Use the indices of this subsequence for the second coordinates, $\{x_{n_k2}\}$. This is a bounded sequence, so it has a convergent subsequence $\{x_{n_l2}\}$. Note that $\{x_{n_l1}\}$ is a subsequence of $\{x_{n_k1}\}$, hence it is also convergent. Continue this process until finding a subsequence $\{x_{n_p}\}$ that converges for each coordinate. This is the convergent subsequence we were looking for. □

Definition 2.38. Let $\{x_n\}$ be a sequence of real numbers.
We say that $\{x_n\}$ is divergent to $+\infty$, and write it as $\{x_n\} \to +\infty$, if for all $M > 0$ there exists $N = N(M) > 0$ such that

$$x_n > M, \ \forall \ n > N\,.$$

Similarly, we say that $\{x_n\}$ is divergent to $-\infty$, and write it as $\{x_n\} \to -\infty$, if for all $M > 0$ there exists $N = N(M) > 0$ such that

$$x_n < -M, \ \forall \ n > N\,.$$

Example 2.39. Consider $x_n = n^2$ for all $n \in \mathbb{N}$. If $M > 0$, then

$$n^2 > M \Leftrightarrow n > \sqrt{M}.$$

This shows that we can choose $N(M) = \sqrt{M}$, and by this we conclude that $\{n^2\} \to +\infty$.

Definition 2.40. The extended real number system, denoted by $\overline{\mathbb{R}}$, consists of all real numbers and two symbols, $-\infty$ and $+\infty$.

Definition 2.41. Let $\{x_n\}$ be a sequence of real numbers. Let

$$S = \{x \in \overline{\mathbb{R}} \mid \exists \{x_{n_k}\} \to x\}.$$

We define

$$\limsup x_n = \sup S \quad \text{and} \quad \liminf x_n = \inf S.$$

Note that, by definition, $\liminf x_n \le \limsup x_n$.

Remark 2.42. Another way to define $\limsup x_n$ is

$$\limsup x_n = \inf_{n \ge 1} \sup\{x_n, x_{n+1}, \dots\}.$$

Similarly,

$$\liminf x_n = \sup_{n \ge 1} \inf\{x_n, x_{n+1}, \dots\}.$$

We leave as exercises to show that the above definitions for lim inf and lim sup are equivalent.

Example 2.43. (a) $\liminf(-n) = \limsup(-n) = -\infty$.

(b) $\liminf n = \limsup n = +\infty$.

(c) Let $x_n = n(1 + (-1)^n)$ for all $n \in \mathbb{N}$. Then

$$\liminf x_n = 0 \quad \text{and} \quad \limsup x_n = +\infty.$$

(d) Let $x_n = 1 + (-1)^n$ for all $n \in \mathbb{N}$. Then

$$\liminf x_n = 0 \quad \text{and} \quad \limsup x_n = 2.$$

(e) Let $x_n = (-0.5)^n$ for all $n \in \mathbb{N}$. Then

$$\liminf x_n = 0 \quad \text{and} \quad \limsup x_n = 0.$$

(f) List all rational numbers as a sequence $\{x_n\}$. This is possible because $\mathbb{Q}$ is countable. Then, the set S from Definition 2.41 is, $S = \overline{\mathbb{R}}$. Therefore,

$$\liminf x_n = -\infty, \quad \text{and} \quad \limsup x_n = +\infty.$$

The following theorem follows from Definition 2.41 and the fact that for a convergent sequence any subsequence converges to the unique limit of the sequence.

Theorem 2.44. *Let $\{x_n\}$ be a sequence of real numbers. Then $\{x_n\}$ is convergent if and only if*

$$-\infty < \liminf x_n = \limsup x_n < +\infty\,.$$

The next theorem gives a characterization of lim inf.

Theorem 2.45. *Let $\{x_n\}$ be a bounded sequence in $\mathbb{R}$. Then $\alpha = \liminf x_n$ if and only if the following two conditions hold:*

(i) *α is a subsequential limit point of $\{x_n\}$.*

(ii) *For any $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $\alpha - \varepsilon < x_n$ for all $n > N$.*

Proof. First, let us assume that $\alpha = \liminf x_n$. Then, from the fact that α is the infimum of the set of subsequential limit points, it follows that for all $k \in \mathbb{N}$ there exists a subsequential limit point s_k and also a term x_{n_k} such that

$$\alpha \le s_k < \alpha + \frac{1}{2k} \quad \text{and} \quad |x_{n_k} - s_k| < \frac{1}{2k}\,.$$

Therefore, for all $k \in \mathbb{N}$ we have

$$|x_{n_k} - \alpha| \le |x_{n_k} - s_k| + |s_k - \alpha| < \frac{1}{k}\,,$$

which shows that $\{x_{n_k}\} \to \alpha$. This proved (i).

To prove (ii), let us assume, by contradiction, that there exists $\varepsilon > 0$ such that for all $k \in \mathbb{N}$ there exist $n_k > k$ such that $x_{n_k} \le \alpha - \varepsilon$. The subsequence $\{x_{n_k}\}$ is bounded, hence by Theorem 2.36, it has a convergent subsequence. This convergent subsequence is still a subsequence of $\{x_n\}$, so we will use the same notation $\{x_{n_k}\}$. Therefore, $\{x_{n_k}\} \to x \le \alpha - \varepsilon$, which contradicts that $\alpha = \liminf x_n$.

Second, let us assume that (i) and (ii) are true. By (i) we get that $\liminf x_n \le \alpha$, while (ii) implies that for any $\varepsilon > 0$ we have $\alpha - \varepsilon \le \liminf x_n$. Therefore,

$$\alpha - \varepsilon \le \liminf x_n \le \alpha, \quad \forall\ \varepsilon > 0,$$

which proves that $\alpha = \liminf x_n$.

□

A similar theorem holds for $\limsup x_n$ and its proof is left as an exercise.

Theorem 2.46. *Let $\{x_n\}$ be a bounded sequence in $\mathbb{R}$. Then $\alpha = \limsup x_n$ if and only if the following two conditions hold:*

(i) *α is a subsequential limit point of $\{x_n\}$.*

(ii) *For any $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $x_n < \alpha + \varepsilon$ for all $n > N$.*

Exercises

Exercise 2.15. Let $\{x_n\}$ be a sequence in $\mathbb{R}$. Show that

$$\lim_{n\to\infty} x_n = x \implies \lim_{n\to\infty} |x_n| = |x|.$$

Is the converse true?

Exercise 2.16. Let $\{x_n\}$ a sequence in $\mathbb{R}$. Show that

$$\lim_{n\to\infty} x_n = 0 \iff \lim_{n\to\infty} |x_n| = 0.$$

Exercise 2.17. Let $\{x_n\}$ be a sequence in a metric space X. Prove that

$$\lim_{n\to\infty} x_n = x \iff \lim_{n\to\infty} d(x_n, x) = 0.$$

Exercise 2.18. Let $\{x_n\}$ and $\{y_n\}$ be sequences in a metric space X. Show that if $\{x_n\} \to x$ and $\{d(x_n, y_n)\} \to 0$, then $\{y_n\} \to x$.

Exercise 2.19. Let $\{x_n\}$ and $\{y_n\}$ be sequences in $\mathbb{R}$. Suppose that $\{x_n\} \to x \neq 0$ and $\{x_n \cdot y_n\}$ is also convergent. Prove that $\{y_n\}$ is convergent.

What can be said about the case $x = 0$?

Exercise 2.20. Let $\{x_n\}$, $\{y_n\}$ and $\{z_n\}$ be sequences of real numbers. Suppose that

$$x_n \le y_n \le z_n \ \ \forall\ \ n \in \mathbb{N}$$

and there exists $a \in \mathbb{R}$ such that

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n = a\,.$$

Show that $\lim_{n\to\infty} y_n = a$.

Exercise 2.21. Let $\{x_n\}$ be a sequence in $\mathbb{R}$ with $x_n > 0$ for all $n \in \mathbb{N}$. Suppose that

$$\lim_{n\to\infty} \frac{x_{n+1}}{x_n} = L.$$

(a) Show that if $L < 1$, then $\{x_n\} \to 0$.
(b) Show that if $L > 1$, then $\{x_n\} \to \infty$.
(c) What can we say if $L = 1$?

Exercise 2.22. Determine whether the following sequences converge or diverge. Find their limits in case of convergence.

$$
\begin{aligned}
&(a)\ \ x_n = \frac{2n^2-1}{3n^2+n+5}\\
&(b)\ \ x_n = \sqrt{n^2+n}-\sqrt{n^2-n}\\
&(c)\ \ x_n = \frac{n!}{n^n}\\
&(d)\ \ x_n = \frac{1\cdot 3\cdot 5\cdot\dots\cdot(2n-1)}{n!}\\
&(e)\ \ x_n = \frac{3^n+n^2}{4^n+n}\\
&(f)\ \ x_n = \sqrt[n]{n}
\end{aligned}
$$

Exercise 2.23. Let $k \in \mathbb{N}$ and the sequence $\{x_n\}$ defined by:

$$x_1 = k,\ \ x_{n+1} = \frac{x_n}{2} + \frac{k}{2x_n}\,,\ \ \forall\ n \in \mathbb{N}.$$

Prove that $\{x_n\}$ is convergent and find its limit.

Exercise 2.24. Let the sequence $\{x_n\}$ be defined by:

$$x_1 = 3,\ \ x_{n+1} = 1 + \frac{1}{x_n}\,,\ \ \forall\ n \in \mathbb{N}.$$

Prove that $\{x_n\}$ is convergent and find its limit.

Exercise 2.25. Let the sequence $\{x_n\}$ be defined by:

$$x_1 = 2,\ \ x_{n+1} = \sqrt{2+\sqrt{x_n}}\,,\ \ \forall\ n \in \mathbb{N}.$$

Prove that $\{x_n\}$ is convergent. What can we say about its limit?

Exercise 2.26. Let

$$x_n = \sum_{k=1}^{n} \frac{1}{\sqrt{n^2+k}}\,,\ \ \forall\ n \in \mathbb{N}.$$

Show that $\{x_n\}$ is convergent and find its limit.

Exercise 2.27. Let

$$x_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}\,, \quad \forall\ n \in \mathbb{N}.$$

Show that $\{x_n\}$ is convergent. What can we say about its limit?

Exercise 2.28. Let $\{x_n\}$ and $\{y_n\}$ be sequences of real numbers. Prove that, if we exclude the $+\infty - \infty$ and $0 \cdot \infty$ cases, the following statements hold:

(a) $\liminf x_n + \liminf y_n \leq \liminf(x_n + y_n) \leq \limsup(x_n + y_n) \leq \limsup x_n + \limsup y_n\,.$

(b) If $\{x_n\} \to x$, then $\liminf(x_n+y_n) = x+\liminf y_n$ and $\limsup(x_n+y_n) = x+\limsup y_n$.

(c) If $x_n \geq 0$, $y_n \geq 0$ for all $n \in \mathbb{N}$, then

$$\liminf x_n \cdot \liminf y_n \leq \liminf(x_n\, y_n) \leq \limsup(x_n\, y_n) \leq \limsup x_n \cdot \limsup y_n\,.$$

(d) If $x_n \geq 0$, $y_n \geq 0$ for all $n \in \mathbb{N}$, and $\{x_n\} \to x > 0$, then

$$x \cdot \liminf y_n = \liminf(x_n\, y_n) \leq \limsup(x_n\, y_n) = x \cdot \limsup y_n\,.$$

What is happening if $x = 0$?

Exercise 2.29. Let $\{x_n\}$ be a sequence from $\mathbb{R}$ and

$$s_n = \frac{x_1 + \cdots + x_n}{n}\,, \quad \forall\ n \in \mathbb{N}\,.$$

(a) Prove that

$$\liminf x_n \leq \liminf s_n \leq \limsup s_n \leq \limsup x_n\,.$$

(b) Prove that if $\{x_n\} \to x$, then $\{s_n\} \to x$.

(c) Give an example of a divergent sequence $\{x_n\}$ such that $\{s_n\}$ is convergent.

2.4 Compact sets

Definition 2.47. Let X be a metric space and $A \subset X$. We say that a collection of open sets $\{B_i\}_{i \in I}$ is an open cover of A if $A \subset \bigcup_{i \in I} B_i$.

Definition 2.48. Let X be a metric space and $K \subset X$. We say that K is a compact set if from every open cover of K we can select a finite subcover.

Example 2.49. Consider the interval $A = (0,1)$ and the open sets $B_n = (\frac{1}{n}, 1)$ for all $n \in \mathbb{N}$. Then the collection $\{B_n\}_{n \in \mathbb{N}}$ is an open cover of $(0,1)$, but we cannot choose a finite subcover. Indeed, if we were able to choose a finite subcover, then there would be a largest set B_{n_0}, which contains all the other sets in the subcover. However, there are real numbers $0 < x < \frac{1}{n_0}$ which are not covered by the finite subcover, and this leads to a contradiction. In conclusion, $A = (0,1)$ is not a compact set.

Proposition 2.50. *Let X be a metric space and $K \subset X$ be a finite set. Then K is compact.*

Proof. Let $K = \{x_1, ..., x_n\}$ and $\{B_i\}_{i \in I}$ be an open cover of K. For every element x_k of K there exist a member B_{i_k} of the open cover such that $x_k \in B_{i_k}$. Therefore, the collection $\{B_{i_1}, ..., B_{i_n}\}$ is a finite subcover of K. □

Proposition 2.51. *Let X be a metric space, $\{x_n\}$ be a sequence from X and S be the range of $\{x_n\}$. If $\{x_n\} \to x$ and $K = S \cup \{x\}$, then K is compact.*

Proof. Let $\{B_i\}_{i \in I}$ be an open cover of K. Select B_{i_0} such that $x \in B_{i_0}$. The set B_{i_0} is open, hence there exists $r > 0$ such that $B_r(x) \subset B_{i_0}$. By the fact that $\{x_n\} \to x$, we can select $N \in \mathbb{N}$ such that $x_n \subset B_r(x)$ for all $n > N$. From the open cover select B_{i_1}, B_{i_2}, ... , B_{i_N} which will cover x_1, x_2, ... , x_N. Therefore, B_{i_0}, B_{i_1}, B_{i_2}, ... , B_{i_N} is an finite subcover of K. □

Proposition 2.52. *Let $a, b \in \mathbb{R}$ and suppose that $a < b$. Then the closed interval $[a, b]$ is compact.*

Proof. We will attempt a proof by contradiction. Assume that there is an open cover $\{B_i\}_{i \in I}$ of $[a, b]$ which doesn't have any finite subcover. Divide the interval $[a, b]$ in two halves and choose one of them, which doesn't have a finite subcover. Denote the chosen half interval by $[a_1, b_1]$. Again, split this interval in two halves, choose one of them which doesn't have a finite subcover and denote it by $[a_2, b_2]$. Continue this process indefinitely. In this way we construct a sequence of intervals

$$[a_1, b_1] \supset [a_2, b_2] \supset \cdots \supset [a_n, b_n] \supset \ldots ,$$

such that $b_n - a_n = \frac{b-a}{2^n}$ for all $n \in \mathbb{N}$ and none of them has a finite subcover.

The sequence $\{a_n\}$ is monotonically increasing and bounded above by b, therefore it is convergent. Similarly, the sequence $\{b_n\}$ is monotonically decreasing and bounded below by a, and hence it is convergent. The two sequences have the same limit, because $\{b_n - a_n\} \to 0$. Denote this common limit point by a_0.

The point $a_0 \in [a, b]$ must be covered by one of the sets B_i, $i \in I$. Let's choose, $a_0 \in B_{i_0}$. The set B_{i_0} is open, hence there exists $r > 0$ such that $(a_0 - r, a_0 + r) \subset B_{i_0}$. Choose $n_0 \in \mathbb{N}$ such that $\frac{b-a}{2^{n_o}} < r$. By the fact that $a_0 \in [a_{n_0}, b_{n_0}]$ we conclude that

$$[a_{n_0}, b_{n_0}] \subset (a_0 - r, a_0 + r) \subset B_{i_0} .$$

However, this shows that $[a_{n_0}, b_{n_0}]$ has a finite subcover, which contradicts the above constructions of these intervals.

$\square$

To generalize Proposition 2.52 from $\mathbb{R}$ to $\mathbb{R}^m$ we introduce the notation of closed rectangular box.

Definition 2.53. Consider the closed intervals $[a_1, b_1]$, ... , $[a_m, b_m]$. We define the closed rectangular box as the Cartesian product

$$[a_1, b_1] \times ... \times [a_m, b_m] = \{x \in \mathbb{R}^m \mid a_k \leq x_k \leq b_k, \ \forall \, 1 \leq k \leq m\}.$$

Similarly, the open rectangular box is defines as

$$(a_1, b_1) \times ... \times (a_m, b_m) = \{x \in \mathbb{R}^m \mid a_k < x_k < b_k, \ \forall \, 1 \leq k \leq m\}.$$

To measure the largest difference between the points of a set, we introduce the notion of the diameter.

Definition 2.54. Let X be a metric space and $A \subset X$. The diameter of A is defined by

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

Observe that the diameter of a ball is twice the radius and the diameter of the rectangular box from Definition 2.53 is given by

$$\operatorname{diam}([a_1, b_1] \times ... \times [a_m, b_m]) = \left(\sum_{k=1}^{m} (b_k - a_k)^2 \right)^{\frac{1}{2}} .$$

Any box can be fit within a closed ball with radius equal to half of the diameter of the box.

Proposition 2.55. *Every closed rectangular box in $\mathbb{R}^m$ is compact.*

Proof. The proof is very similar to the proof of Proposition 2.52. We will highlight the differences.

Consider a rectangular box $R = [a_1, b_1] \times ... \times [a_m, b_m]$. Assume that R has an open cover $\{B_i\}_{i \in I}$, which doesn't have any finite subcover. Divide the box R in a number of 2^m equal boxes, by cutting each interval $[a_k, b_k]$ in half. Choose one of them which doesn't have a finite subcover and name it R_1. Note that

$$R_1 = [a_{1_1}, b_{1_1}] \times ... \times [a_{m_1}, b_{m_1}] ,$$

where each interval is a half of the original. Also note that

$$\operatorname{diam}(R_1) = \frac{1}{2} \operatorname{diam}(R) .$$

Continue this process and at each step we find a rectangular box

$$R_n = [a_{1_n}, b_{1_n}] \times ... \times [a_{m_n}, b_{m_n}] ,$$

which doesn't have a finite subcover and

$$\operatorname{diam}(R_n) = \frac{1}{2^n} \operatorname{diam}(R) .$$

As in Proposition 2.52, for each coordinate we get a limit of the endpoints, so we obtain $a_o = (a_{1_o}, ..., a_{m_o})$.

The point $a_0 \in R$ must be covered by one of the sets B_i, $i \in I$. Let's choose, $a_0 \in B_{i_0}$. The set B_{i_0} is open, hence there exists $r > 0$ such that $B_r(a_0) \subset B_{i_0}$. Choose $n_0 \in \mathbb{N}$ such that $\frac{\operatorname{diam}(R)}{2^{n_o}} < r$. By the fact that $a_0 \in R_{n_0}$ we conclude that

$$R_{n_0} \subset B_r(a_0) \subset B_{i_0} .$$

However, this shows that R_{n_0} has a finite subcover, which contradicts the above constructions of these rectangular boxes.

□

Before we prove characterizations of compact sets, we need some additional results.

Theorem 2.56. *Let X be a metric space. If $K \subset X$ is compact, then K is bounded.*

Proof. Fix any $x \in K$. Consider the following open cover of K:

$$K \subset \bigcup_{n=1}^{\infty} B_n(x) ,$$

where $B_n(x)$ is an open ball of radius n. By the compactness of K, we can select a finite subcover. Let n_0 be the largest radius in the subcover. Then

$$K \subset B_{n_0}(x) ,$$

which shows that K is bounded.

□

Theorem 2.57. *Let X be a metric space. If $K \subset X$ is compact, then K is closed.*

Proof. We will prove that K^c is open. Let $x \in K^c$ be an arbitrary point. For each $y \in K$ consider $r_y = \frac{1}{2}d(x,y)$ and with it the open cover of K:

$$K \subset \bigcup_{y\in K} B_{r_y}(y)\,.$$

Using the compactness of K, we can select $y_1, ..., y_n \in K$ such that

$$K \subset \bigcup_{i=1}^{n} B_{r_{y_i}}(y_i)\,.$$

Let

$$r = \min\left\{r_{y_1}, ..., r_{y_n}\right\}\,.$$

Then $r > 0$ as the minimum of finitely many positive numbers and, moreover,

$$B_r(x) \bigcap \left(\bigcup_{i=1}^{n} B_{r_{y_i}}(y_i)\right) = \emptyset\,.$$

Hence, $K \cap B_r(x) = \emptyset$, so $B_r(x) \subset K^c$. As x can be chosen as any point of K^c, we conclude that K^c is open and therefore K is closed.

□

The previous two theorems imply the following corollary.

Corollary 2.58. *Let X be a metric space. If $K \subset X$ is compact, then K is closed and bounded.*

Note that the converse of the corollary is not always true: If we consider the discrete metric in $\mathbb{R}$, the interval $[0,1]$ is not compact. Indeed, with the discrete metric $B_{\frac{1}{2}}(x) = \{x\}$, so from the covering with open balls of radius $\frac{1}{2}$, we cannot extract a finite subcover.

Theorem 2.59. *Let X be a metric space. If $A \subset K \subset X$, A is closed and K is compact, then A is compact.*

Proof. Let $\{B_i\}_{i\in I}$ be an open cover of A. As A is closed, we know that A^c is open and we can extend the open cover of A to an open cover of K:

$$K \subset \left(\bigcup_{i\in I} B_i\right) \bigcup A^c\,.$$

The compactness of K implies that there are $i_1, ..., i_n \in I$ such that

$$K \subset \left(\bigcup_{k=1}^{n} B_{i_k}\right) \bigcup A^c\,.$$

Therefore,

$$A \subset \bigcup_{k=1}^{n} B_{i_k}\,,$$

which shows that we were able to select a finite subcover.

$\square$

Theorem 2.60. *Let X be a metric space and $\{K_i\}_{i\in I}$ be a collection of nonempty compact subsets of X, which satisfy the finite intersection property: any finitely many of these sets have nonempty intersection. Then,*

$$\bigcap_{i\in I} K_i \neq \emptyset\,.$$

Proof. Without loss of generality we can assume that X is compact, because otherwise we fix any $i_0 \in I$ and work with $\tilde{X} = X \cap K_{i_0} = K_{i_0}$ and $\tilde{K}_i = K_i \cap K_{i_0}$.

By contradiction, assume that

$$\bigcap_{i\in I} K_i = \emptyset\,.$$

Therefore, by the DeMorgan laws

$$X = \bigcup_{i\in I} (K_i)^c\,,$$

which is an open cover of X. Using the compactness of X, select a finite subcover:

$$X = \bigcup_{k=1}^{n} (K_{i_k})^c\,.$$

Again, by the DeMorgan laws we obtain

$$\bigcap_{k=1}^{n} K_{i_k} = \emptyset\,,$$

which contradicts the finite intersection property.

$\square$

Corollary 2.61. *Let X be a metric space and $\{K_n\}_{n\in\mathbb{N}}$ be a collection of nested nonempty compact subsets of X, which means that $K_{n+1} \subset K_n$ for every $n \in \mathbb{N}$. Then*

$$K = \bigcap_{n=1}^{\infty} K_n \neq \emptyset\,.$$

Moreover, if $\{\mathrm{diam}(K_n)\} \to 0$, then K contains exactly one element.

Proof. Every nested collection of nonempty sets satisfy the finite intersection property, hence the fact that $K \neq \emptyset$ follows from Theorem 2.60. In case of $\{\mathrm{diam}(K_n)\} \to 0$, assume that we have $x, y \in K$. Therefore, $x, y \in K_n$ for every $n \in \mathbb{N}$, so

$$d(x,y) \leq \mathrm{diam}(K_n), \ \forall \ n \in \mathbb{N}\,.$$

This leads to $d(x,y) = 0$ and ultimately to $x = y$.

□

The following theorem characterizes the compact sets in $\mathbb{R}^m$.

Theorem 2.62. *(The Heine-Borel theorem)*
Let $K \subset \mathbb{R}^m$ a nonempty set. The following statements are equivalent:

(a) *K is compact.*

(b) *K is closed and bounded.*

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

(d) *Every sequence from K has a subsequence convergent to a point in K.*

Proof. First we show that (a) $\Leftrightarrow$ (b). Indeed, if K is compact, then by Corollary 2.58 is closed and bounded. If K is closed and bounded, then K is a closed subset of a sufficiently large closed rectangular box, which is compact and therefore, by Theorem 2.59 we conclude that K is compact.

Next we will show that (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (b).

For (b) $\Rightarrow$ (c), as K is bounded, every infinite subset of it has a limit point. This is the Bolzano-Weierstrass theorem in $\mathbb{R}^m$. By the fact that K is closed, the limit point is in K.

For (c) $\Rightarrow$ (d), let $\{x_n\}$ be a sequence from K. If it has a constant subsequence, then this is the convergent subsequence we are looking for. If it doesn't have a constant subsequence, then the range S of $\{x_n\}$ has infinitely many elements, and by (c) it has a limit point in K. Lemma 2.35 gives us a sequence from S converging to this limit point. This sequence can be rearranged as a subsequence of $\{x_n\}$.

For (d) $\Rightarrow$ (b), assume first that K is not bounded. Then for all $n \in \mathbb{N}$ there exists $x_n \in K$ such that $||x_n|| > n$. The sequence $\{x_n\}$, which we just constructed, cannot have any bounded subsequence, hence cannot have any convergent subsequence, which contradicts (d). Assume now that K is not closed. Then, K has a limit point $x \notin K$. Therefore, we can construct a sequence $\{x_n\}$ from K, converging to x. However, every subsequence will converge to x, hence we cannot have a subsequence converging to a point in K and this contradicts (d).

□

Exercises

Exercise 2.30. Prove that any compact metric space is separable.

Exercise 2.31. Show that in $\mathbb{R}^m$, the collection of open balls with rational radius and centers at points with rational coordinates is countable.

Exercise 2.32. Show that every open set $A \subset \mathbb{R}^m$ is the union of a, at most countable, collection of open balls.

Exercise 2.33. Let $A \subset \mathbb{R}^m$. Prove that every open cover of A contains a, at most countable, subcover.

Exercise 2.34. Let $A \subset \mathbb{R}^m$ and assume that for every $x \in A$ there exists an open ball $B_r(x)$ such that $A \cap B_r(x)$ is at most countable. Show that A is at most countable.

Exercise 2.35. Let X be a metric space and $A \subset X$. Show that

$$\operatorname{diam}(A) = \operatorname{diam}(\overline{A}) .$$

2.5 Cauchy sequences and complete metric spaces

Definition 2.63. Let X be a metric space and $\{x_n\}$ be a sequence in X. We say that $\{x_n\}$ is a Cauchy sequence if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that for all $n, m > N$ we have $d(x_n, x_m) < \varepsilon$.

The following equivalent definition of a Cauchy sequence is also useful.

We say that $\{x_n\}$ is a Cauchy sequence if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that for all $n > N$ and for all $k \in \mathbb{N}$ we have $d(x_n, x_{n+k}) < \varepsilon$.

The notion of Cauchy sequences will be an important tool in studying convergent sequences, when we don't know the limit of the sequence.

Theorem 2.64. *Let X be a metric space and $\{x_n\}$ be a sequence in X. If $\{x_n\}$ is convergent, then $\{x_n\}$ is a Cauchy sequence.*

Proof. Assume that $\{x_n\} \to x$. Then, for every $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that, for all $n > N$ we have $d(x_n, x) < \frac{\varepsilon}{2}$. Hence, for all $n, m > N$ we have

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,,$$

which shows that $\{x_n\}$ is a Cauchy sequence.

□

Example 2.65. This example shows that the converse of the previous theorem is not always true. Consider $X = (0, 1]$ and $x_n = \frac{1}{n}$ for all $n \in \mathbb{N}$. Taking into consideration that for $n, k \in \mathbb{N}$ we have

$$d(x_n, x_{n+k}) = \left|\frac{1}{n} - \frac{1}{n+k}\right| = \frac{k}{(n+k)n} < \frac{1}{n}\,,$$

we can define $N(\varepsilon) = \frac{1}{\varepsilon}$, and this proves that $\{\frac{1}{n}\}$ is a Cauchy sequence. However, $\{\frac{1}{n}\}$ is not convergent in X, because its limit $0 \notin X$.

The proof of the next theorem is left as an exercise.

Theorem 2.66. *Let X be a metric space and $\{x_n\}$ be a sequence in X. If $\{x_n\}$ is a Cauchy sequence, then $\{x_n\}$ is a bounded sequence.*

Definition 2.67. A complete metric space is a metric space in which every Cauchy sequence is convergent.

Theorem 2.68. *Any compact metric space is a complete metric space.*

Proof. Let X be a compact metric space and let $\{x_n\}$ be a Cauchy sequence in X. For every $n \in \mathbb{N}$ define the set

$$A_n = \{x_n, x_{n+1}, ...\}.$$

The definition of a Cauchy sequence implies that $\{\text{diam}(A_n)\} \to 0$. Let $K_n = \overline{A_n}$. By the fact that a set and its closure have the same diameter, it follows that $\{\text{diam}(K_n)\} \to 0$. Moreover, the sets K_n are compact, as closed subsets of a compact metric space and $K_{n+1} \subset K_n$ for all $n \in \mathbb{N}$. Therefore, by Corollary 2.61, we obtain an $x \in X$ such that

$$\bigcap_{n=1}^{\infty} K_n = \{x\}\,.$$

The fact that $\{x_n\} \to x$ follows from the observation

$$d(x_n, x) \leq \text{diam}(K_n),\ \forall\ \ n \in \mathbb{N}\,.$$

□

Corollary 2.69. $\mathbb{R}^m$ *is a complete metric space.*

Proof. Let $\{x_n\}$ be a Cauchy sequence in $\mathbb{R}^m$. By Theorem 2.66 it follows that $\{x_n\}$ is bounded and hence it can be included in a closed rectangular box. However, any closed rectangular box is a compact set, hence $\{x_n\}$ is convergent by Theorem 2.68.

□

The proof of the next proposition is left as an exercise.

Proposition 2.70. *Let X be a complete metric space and $B_{r_n}[x_n]$ be a collection of nested closed balls. If $\{r_n\} \to 0$, then*

$$\bigcap_{n=1}^{\infty} B_{r_n}[x_n] \neq \emptyset\,.$$

Definition 2.71. Let X be a metric space and $A \subset X$. We say that A is nowhere dense in X if for any open ball $B_r(x)$, the intersection $A \cap B_r(x)$ is not dense in $B_r(x)$.

The proof of the next proposition is left as an exercise.

Proposition 2.72. *Let X be a metric space and $A \subset X$. Then A is nowhere dense in X if and only if $\overline{A}$ has empty interior.*

Theorem 2.73. *(Baire's theorem)*
A complete metric space cannot be represented as a union of countably many nowhere dense sets.

Proof. Let us suppose that

$$X = \bigcup_{n=1}^{\infty} A_n \, ,$$

where A_n is nowhere dense for all $n \in \mathbb{N}$. Consider any $x_0 \in X$. The set A_1 is nowhere dense, so there exists $B_{r_1}[x_1] \subset B_1(x_0)$ such that $r_1 < \frac{1}{2}$ and $A_1 \cap B_{r_1}[x_1] = \emptyset$.

Now we use the fact that A_2 is nowhere dense, and get $B_{r_2}[x_2] \subset B_{r_1}(x_1)$ such that $r_2 < \frac{1}{2^2}$ and $A_2 \cap B_{r_2}[x_2] = \emptyset$.

Continuing in this way, we get a nested sequence of closed balls, $\{B_{r_n}[x_n]\}$ such that $r_n \leq \frac{1}{2^n}$ and $A_n \cap B_{r_n}[x_n] = \emptyset$.

Proposition 2.70 implies that there exist $x \in X$ such that

$$x \in \bigcap_{n=1}^{\infty} B_{r_n}[x_n] \, .$$

By construction, $x \notin A_n$ for all $n \in \mathbb{N}$, hence

$$x \notin \bigcup_{n=1}^{\infty} A_n = X \, ,$$

which is a contradiction to $x \in X$.

In conclusion, the assumption $X = \bigcup_{n=1}^{\infty} A_n$ cannot be true.

□

Theorem 2.74. *Let X be a complete metric space which does not have isolated points. Then X is uncountable.*

Proof. Consider any $x \in X$. The proof relies on showing that the set $\{x\}$ is nowhere dense in X. For this consider any open ball $B_r(x)$. By the fact that x is not isolated, there exists $y \in B_r(x)$ such that $x \neq y$. Then, for some $0 < r' < \frac{1}{2}d(x, y)$, we have $B_{r'}(y) \subset B_r(x)$, but $\{x\} \cap B_{r'}(y) = \emptyset$. In this way we proved that $\{x\}$ is nowhere dense for any $x \in X$. Hence X cannot be written as the union of countably many sets made of one point each. □

Definition 2.75. Let X be a metric space and $P \subset X$. We say that P is a perfect set if it is closed and every point of P is a limit point of P.

Corollary 2.76. *Let X be a complete metric space and $\emptyset \neq P \subset X$ be a perfect set. Then P is uncountable.*

Proof. As a closed subset of a complete metric space, P is a complete metric space with the metric inherited from X. Each point of P is a limit point, so P does not have isolated points. The fact that P is uncountable follows now from Theorem 2.74. □

Exercises

Exercise 2.36. Let X be a complete metric space and $Y \subset X$ be a closed set. Show that, with the metric inherited from X, Y is a complete metric space.

Exercise 2.37. Prove Theorem 2.66.

Exercise 2.38. Let X be a metric space and $\{x_n\}$ be a sequence in X. Assume that there exists a sequence $\{r_n\}$ of real numbers converging to 0 such that

$$d(x_n, x_{n+k}) < r_n, \ \forall \ n, k \in \mathbb{N}\,.$$

Prove that $\{x_n\}$ is a Cauchy sequence.

Exercise 2.39. Prove Proposition 2.70.

Exercise 2.40. Let X be a metric space at $A \subset X$ be a set which is open and dense. Show that $B = X \setminus A$ is nowhere dense.

Exercise 2.41. Let X be a complete metric space and for every $n \in \mathbb{N}$ let $A_n \subset X$ be open and dense in X. Prove that

$$\bigcap_{n=1}^{\infty} A_n \neq \emptyset\,.$$

Exercise 2.42. Let X be a metric space and $\{x_n\}$ be a Cauchy sequence in X. Assume that there is a subsequence $\{x_{n_k}\}$ convergent to x. Prove that $\{x_n\} \to x$.

Exercise 2.43. Let X be a complete, separable metric space and $A \subset X$. A point $x \in X$ is called a condensation point of A, if every neighborhood of x contains uncountably many points of A. Let P be the set of condensation points of A and assume that $P \neq \emptyset$. Prove that P is a perfect set and $A \setminus P$ is at most countable.

Exercise 2.44. Prove that every countable, closed subset of $\mathbb{R}^m$ has isolated points.

Exercise 2.45. Is there a nonempty perfect set in $\mathbb{R}$ which contains no rational number?

Exercise 2.46. Let $0 < \varepsilon < 1$. Explain how can we cover $\mathbb{Q} \cap [0, 1]$ by a collection of disjoint intervals with total length less than ε.

Exercise 2.47. Let X be a complete metric space, $0 < \lambda < 1$ and $\{x_n\}$ be a sequence in X. Show that if

$$d(x_{n+2}, x_{n+1}) \leq \lambda d(x_{n+1}, x_n)\,,\ \forall\, n \in \mathbb{N}\,,$$

then $\{x_n\}$ is a convergent sequence.

Exercise 2.48. Let $a, b \in \mathbb{R}$, $a < b$. Define the sequence $\{x_n\}$ by

$$x_0 = a,\ x_1 = b,\ x_{n+1} = \frac{1}{2}\Big(x_n + x_{n-1}\Big),\ \forall\, n \in \mathbb{N}\,.$$

(i) Prove that $\{x_n\}$ is convergent.
(ii) Find the limit of $\{x_n\}$.

2.6 The Cantor set and a little measure theory on $\mathbb{R}$.

Let $C_0 = [0,1]$. We remove the open interval $(\frac{1}{3}, \frac{2}{3})$ from C_0 and get

$$C_1 = \left[0, \frac{1}{3}\right] \bigcup \left[\frac{2}{3}, \frac{3}{3}\right] .$$

Consider the following two functions:

$$\begin{aligned} f(x) &= \frac{x}{3} \\ g(x) &= \frac{x}{3} + \frac{2}{3} . \end{aligned}$$

Observe that

$$C_1 = f(C_0) \bigcup g(C_0) .$$

We continue the process by removing the middle one-third of each interval from C_1 and get

$$C_2 = \left[0, \frac{1}{9}\right] \bigcup \left[\frac{2}{9}, \frac{3}{9}\right] \bigcup \left[\frac{6}{9}, \frac{7}{9}\right] \bigcup \left[\frac{8}{9}, \frac{9}{9}\right] .$$

Observe again that

$$C_2 = f(C_1) \bigcup g(C_1) .$$

Continuing in this way we obtain a collection of sets $\{C_n\}_{n \in \mathbb{N}}$ with the following properties:

(1) Each C_n is compact as a union of a number of 2^n pairwise disjoint, closed and bounded intervals:

$$C_n = \bigcup_{k=1}^{2^n} C_{n_k} .$$

(2) $C_{n+1} = f(C_n) \bigcup g(C_n)$ and $C_{n+1} \subset C_n$ for all $n \in \mathbb{N}$.

(3) The length of each C_{n_k} is $\frac{1}{3^n}$.

(4) Each interval C_{n_k} can be identified by its left endpoint a_{n_k}, which is in the form

$$a_{n_k} = \sum_{i=1}^{n} \varepsilon_{i_k} \cdot 2 \cdot \frac{1}{3^i} ,$$

where ε_{i_k} equals 0 or 1. We can see that there are 2^n possibilities to choose the ε_{i_k} constants.

By Theorem 2.60, the intersection of these sets is nonempty, hence the following definition makes sense.

Definition 2.77. The Cantor set (more precisely, the middle third Cantor set) is defined as:

$$C = \bigcap_{n=1}^{\infty} C_n \,.$$

Theorem 2.78. *The Cantor set is a perfect set.*

Proof. First, the Cantor set is closed, as an intersection of closed sets. Let any $x \in C$. Then, $x \in C_n$ for all $n \in \mathbb{N}$ and hence $x \in C_{n_k}$ for a certain $1 \le k \le 2^n$. Denote one of the endpoints of the chosen C_{n_k} by x_n. In this way we can construct a sequence $\{x_n\}$ with the following properties valid for all $n \in \mathbb{N}$:

(a) $x_n \in C_n \subset C$.

(b) $|x - x_n| \le \frac{1}{3^n}$.

Therefore, $\{x_n\} \to x$ and hence x is a limit point of C. This proof is valid for any point of C, hence C is a perfect set.

□

Corollary 2.79. *The Cantor set is uncountable and nowhere dense.*

Proof. The proof of the fact that the Cantor set is uncountable is an immediate consequence of Theorem 2.78 and Corollary 2.76.

The fact that the Cantor set is nowhere dense follows from the fact that it doesn't contain any interval and hence has an empty interior.

□

The proofs of the following two propositions are left as exercises.

Proposition 2.80. *For the Cantor set the following relation holds:*

$$C = f(C) \bigcup g(C) \,.$$

Proposition 2.81. $x \in C$ *if and only if*

$$x = \sum_{n=1}^{\infty} \frac{x_n}{3^n} \,,$$

where each x_n *equals* 0 *or* 2.

Definition 2.82. We call the intervals and their, at most countable, disjoint unions, elementary sets. We define the measure of an elementary set in the following way:

$$\mu((a,b)) = \mu((a,b]) = \mu([a,b)) = \mu([a,b]) = b - a$$

$$\mu\left(\bigcup_n I_n\right) = \sum_n \mu(I_n)\,, \text{ if } I_n \text{ are intervals and } I_n \cap I_m = \emptyset\,, \text{ when } n \neq m.$$

Also, we define $\mu(\emptyset) = 0$.

Definition 2.83. Let $A \subset \mathbb{R}$. We define the outer measure of A by

$$\mu^*(A) = \inf \left\{ \sum_n \mu(I_n) \,, \;\; A \subset \bigcup_n I_n \;,\; I_n \text{ interval} \right\} .$$

The outer measure has the following properties:

- If E is an elementary set, then $\mu(E) = \mu^*(E)$.
- If $A = \bigcup_n A_n$, then
$$\mu^*(A) \leq \sum_n \mu^*(A_n) \,.$$
- If $A \subset \mathbb{R}$ and $x \in \mathbb{R}$, then $\mu^*(x + A) = \mu^*(A)$.
- If $A \subset \mathbb{R}$ and $c \in \mathbb{R}$, then $\mu^*(cA) = |c| \, \mu^*(A)$.
- If $A \subset B$, then $\mu^*(A) \leq \mu^*(B)$.

Definition 2.84. If $A \subset [a, b]$, we say that A is Lebesgue measurable if

$$\mu^*(A) + \mu^*([a, b] \setminus A) = b - a \,.$$

If $A \subset [a, b]$ is Lebesgue measurable, we define its Lebesgue measure as $\mu(A) = \mu^*(A)$.
If $A \subset \mathbb{R}$, we say that A is Lebesgue measurable if $A \bigcap [n, n+1]$ is Lebesgue measurable for all $n \in \mathbb{Z}$ and define

$$\mu(A) = \sum_{n \in \mathbb{Z}} \mu \left(A \bigcap [n, n+1] \right) \,.$$

Example 2.85. The following sets are Lebesgue measurable in $\mathbb{R}$:

- The open sets.
- The closed sets.
- The Borel sets, which are defined as the, at most countable, unions and intersections of open or closed sets.

Actually, all the sets we will be dealing with will be Lebesgue measurable. To have an idea of what a non-measurable set means, here is a short description of a non-measurable set:

In $[0, 1)$ define the following equivalence relation:

$$x \sim y \;\text{ if }\; x - y \in \mathbb{Q} \,.$$

For each $x \in [0, 1)$ we can define the equivalence class of x by

$$[x] = \{ y \in [0, 1) \mid \; x \sim y \} \,.$$

If $x_1 \sim x_2$, then $[x_1] = [x_2]$, otherwise $[x_1] \cap [x_2] = \emptyset$. Using the Axiom of Choice, select one element from each equivalence class and by this create a new set S. This set is non-measurable.

Definition 2.86. Let $A \subset \mathbb{R}$. We say that A is a set of measure zero if

$$\mu^*(A) = 0\,.$$

The following Lemma follows from the definition of the outer measure:

Lemma 2.87. *For $A \subset \mathbb{R}$ we have $\mu^*(A) = 0$ if and only if for all $\varepsilon > 0$ there exists a, at most countable, collection of intervals $\{I_n\}$ such that*

$$A \subset \bigcup_n I_n \ \ \textit{and} \ \ \sum_n \mu(I_n) < \varepsilon\,.$$

Proposition 2.88. *Any set of measure zero is Lebesgue measurable.*

Proof. Let $A \subset [a, b]$ with $\mu^*(A) = 0$. The properties of the outer measure imply that

$$b - a = \mu^*([a, b]) = \mu^*\left(A \bigcup([a, b] \setminus A)\right) \leq \mu^*(A) + \mu^*([a, b] \setminus A) = \mu^*([a, b] \setminus A)\,.$$

Hence, $\mu^*([a, b] \setminus A) \geq b - a$, but also $\mu^*([a, b] \setminus A) \leq \mu^*([a, b]) = b - a$, which shows that $\mu^*([a, b] \setminus A) = b - a$. Therefore, A is Lebesgue measurable.
If $A \subset \mathbb{R}$ and $\mu^*(A) = 0$, then $\mu^*(A \cap [n, n + 1]) = 0$, for all $n \in \mathbb{Z}$. Hence, $A \cap [n, n + 1]$ is Lebesgue measurable for all $n \in \mathbb{Z}$ and this implies that A is Lebesgue measurable.

□

Proposition 2.89. *The Cantor set is of measure zero, hence Lebesgue measurable.*

Proof. From the contruction of the Cantor set

$$C = \bigcap_{n=1}^{\infty} C_n\,,$$

where each C_n is the union of 2^n pairwise disjoint closed intervals of length $\frac{1}{3^n}$. Hence,

$$\mu^*(C) \leq \frac{2^n}{3^n},\ \forall\ \ n \in \mathbb{N},$$

which implies that $\mu^*(C) = 0$.

□

The proof of the following proposition is left as an exercise.

Proposition 2.90. *If $A \subset \mathbb{R}$ is finite or countable, then A is of measure zero.*

We finish this section with a definition.

Definition 2.91. We say that a property holds almost everywhere, if it is true outside of a set of measure zero.

Exercises

Exercise 2.49. Prove Proposition 2.81.

Exercise 2.50. Prove Proposition 2.90.

Exercise 2.51. Let A be the set of all $x \in [0,1]$ whose decimal expansion contains only the digits 3 and 7. Which of the following properties does A have: uncountable, dense in $[0,1]$, open, closed, compact, perfect, measurable?

Exercise 2.52. Repeat the construction of the Cantor set, but now dividing the $[0,1]$ interval in 5 equal parts and taking out the open interval from the middle. Compare the properties of this set with the properties of the Cantor set.

Exercise 2.53. (1) Show that if $A \subset [0,1]$ is Lebesgue measurable, than $[0,1] \backslash A$ is Lebesgue measurable.
(2) Show that $\nu^*([0,1] \cap \mathbb{Q}) = 0$ and that $\mathbb{Q}$ is a measurable set.
(3) Is $[0,1] \cap \mathbb{I} = [0,1] \setminus \mathbb{Q}$ measurable?

Exercise 2.54. Let $A \subset [a,b]$. Assume that there exists $0 \leq \rho < 1$ such that for any $(c,d) \subset [a,b]$ we have

$$\mu^*(A \cap (c,d)) \leq \rho(d-c)\,.$$

Show that A has measure zero.

Exercise 2.55. Show that a countable union of sets of measure zero is of measure zero.

Chapter 3

Continuous functions

3.1 Limits and continuity at a point

Definition 3.1. Let (X, d_X) and (Y, d_Y) be metric spaces, $D \subset X$, $f : D \to Y$ and $x_0 \in X$ be a limit point of D. We say that

$$\lim_{x \to x_0} f(x) = y_0 \,,$$

if for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, x_0) > 0$ such that $x \in D$ and $0 < d_X(x, x_0) < \delta$ imply that $d_Y(f(x), y_0) < \varepsilon$.

Example 3.2. Consider $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ and $x_0 = 2$. Let's prove, using the definition, that

$$\lim_{x \to 2} x^2 = 4 \,.$$

Consider any $\varepsilon > 0$. When we are searching for the $\delta > 0$, we can always restrict the search to $\delta \leq 1$. Let's estimate $|x^2 - 4|$ by imposing for now that $|x - 2| < 1$, which means $1 < x < 3$. Therefore,

$$|x^2 - 4| = |x + 2| \cdot |x - 2| < 5 \cdot |x - 2| \,.$$

We see that if we define

$$\delta = \delta(\varepsilon, 2) = \min\{1, \frac{\varepsilon}{5}\} \,,$$

and require that $|x - 2| < \delta$, then

$$|x^2 - 4| < 5 \cdot \frac{\varepsilon}{5} = \varepsilon \,.$$

In the next theorem we will use convergent sequences to characterize limits.

Theorem 3.3. *Let (X, d_X) and (Y, d_Y) be metric spaces, $D \subset X$, $f : D \to Y$ and $x_0 \in X$ be a limit point of D. Then,*

$$\lim_{x \to x_0} f(x) = y_0$$

if and only if

$$\lim_{n \to \infty} f(x_n) = y_0$$

for all sequences $\{x_n\}$ from D, such that $x_n \neq x_0$ for all $n \in \mathbb{N}$ and $\{x_n\} \to x_0$.

Proof. First suppose that

$$\lim_{x \to x_0} f(x) = y_0 \,.$$

By definition, for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, x_0) > 0$ such that for all $x \in D$ with $0 < d_X(x, x_0) < \delta$ it follows $d_Y(f(x), y_0) < \varepsilon$.

Now choose a sequence $\{x_n\}$ from D, such that $x_n \neq x_0$ for all $n \in \mathbb{N}$ and $\{x_n\} \to x_0$. By convergence, for the $\delta > 0$ chosen above, there exists $N = N(\delta) = N(\varepsilon) > 0$ such that for all $n > N$ we have $d_X(x_n, x_0) < \delta$. Therefore, if $n > N$, we can substitute x_n for x in the definition of the limit of the function at x_0. The conclusion is that for any $\varepsilon > 0$ we were able to find $N = N(\varepsilon) > 0$ such that if $n > N$, then $d_Y(f(x_n), y_0) < \varepsilon$. This shows that $\{f(x_n)\} \to y_0$.

To prove the reverse implication, let us assume, by contradiction, that

$$\lim_{x \to x_0} f(x) \neq y_0 \,.$$

The negation the definition of the limit implies that there exists $\varepsilon > 0$ such that for all $\delta > 0$ there exists $x_\delta \in D$ such that $0 < d_X(x_\delta, x_0) < \delta$, but $d_Y(f(x_\delta, y_0) \geq \varepsilon$. By choosing $\delta = \frac{1}{n}$ for all $n \in \mathbb{N}$, we construct a sequence $\{x_n\}$ in D such that

$$0 < d_X(x_n, x_0) < \frac{1}{n} \quad \text{and} \quad d_Y(f(x_n), y_0) \geq \varepsilon \,.$$

Therefore, $\{x_n\} \to x_0$, but $\{f(x_n\}\}$ doesn't converge to y_0, which is a contradiction with the assumption.

□

The uniqueness of limits of sequences implies the following corollary.

Corollary 3.4. *If the limit of a function at a point exists, it is a unique quantity.*

Definition 3.5. Let X be a metric space, $c \in \mathbb{R}$ and $f, g : X \to \mathbb{R}$. We can define the functions $f + g, cf, f \cdot g, \frac{f}{g} : X \to \mathbb{R}$ in the following way:

$$\begin{aligned}
(f+g)(x) &= f(x) + g(x) \\
(cf)(x) &= xf(x)
\end{aligned}$$

$$(f \cdot g)(x) = f(x) \cdot g(x)$$
$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}\,.$$

The properties of convergent sequences implies the following properties of limits of functions. The proof of the theorem is left as an exercise.

Theorem 3.6. *Let X be a metric space, $D \subset X$, x_0 a limit point of D, $c \in \mathbb{R}$ and $f, g : X \to \mathbb{R}$. If*

$$\lim_{x \to x_0} f(x) = y_0 \quad \textit{and} \quad \lim_{x \to x_0} g(x) = z_0\,,$$

then

(a) $\displaystyle\lim_{x \to x_0} (f+g)(x) = y_0 + z_0$

(b) $\displaystyle\lim_{x \to x_0} (cf)(x) = c\, y_0$

(c) $\displaystyle\lim_{x \to x_0} (f \cdot g)(x) = y_0 \cdot z_0$

(d) $\displaystyle\lim_{x \to x_0} \left(\frac{f}{g}\right)(x) = \frac{y_0}{z_0}$, *if $z_0 \neq 0$ and $g(x) \neq 0$ in a neighborhood of x_0.*

Definition 3.7. Let X and Y be metric spaces, $D \subset X$, $f : D \to Y$ and $x_0 \in D$. We say that f is continuous at x_0 if for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, x_0) > 0$ such that $x \in D$ and $d_X(x, x_0) < \delta$ imply that $d_Y(f(x), f(x_0)) < \varepsilon$.

This definition implies immediately the following proposition.

Proposition 3.8. *Let X and Y be metric spaces, $D \subset X$, $f : D \to Y$ and $x_0 \in D$.*

If x_0 is an isolated point of D, then f is continuous at x_0.

If x_0 is a limit point of D , then f is continuous at x_0 if and only if

$$\lim_{x \to x_0} f(x) = f(x_0)\,.$$

Remark 3.9. Note that, in order to talk about the continuity of the function f at x_0, it is necessary that x_0 be an element of the domain of f.

Consider $f(x) = \frac{\sin x}{x}$ and $x_0 = 0$. The domain of the function is $D = \mathbb{R} \setminus \{0\}$, so at this time we can talk about the limit of the function at 0, but not about the continuity. Observe that

$$\lim_{x \to 0} \frac{\sin x}{x} = 1\,.$$

To talk about the continuity of f at 0, we have to extend the domain of the function to 0:

$$f(x) = \begin{cases} \frac{\sin x}{x} & \text{if } x \neq 0 \\ \alpha & \text{if } x = 0\,. \end{cases}$$

The function f will be continuous at 0 if and only if $\alpha = 1$.

Definition 3.10. Let X and Y be metric spaces, $D \subset X$, $f : D \to Y$ and $x_0 \in D$. We say that f is discontinuous at x_0 if it is not continuous at x_0.

Note that, in order that f to be discontinuous at x_0, the point x_0 must be a limit point of the domain of f.

To characterize the points of discontinuity for functions of a real variable, we need the following sided limits.

Definition 3.11. Let $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ and $x_0 \in \mathbb{R}$ be a limit point of D. We say that the left-sided limit

$$\lim_{x \nearrow x_0} f(x) = y_0$$

exists, if for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, x_0) > 0$ such that $x \in D$ and $0 < x_0 - x < \delta$ imply that $|f(x) - y_0| < \varepsilon$.

We say that the right-sided limit

$$\lim_{x \searrow x_0} f(x) = y_0$$

exists, if for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, x_0) > 0$ such that $x \in D$ and $0 < x - x_0 < \delta$ imply that $|f(x) - y_0| < \varepsilon$.

The proof of the following theorem easily follows from the definition of the limit of a function at a point.

Theorem 3.12. *Let $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ and $x_0 \in D$ be a limit point of D. Then f is continuous at x_0 if and only if*

$$\lim_{x \nearrow x_0} f(x) = \lim_{x \searrow x_0} f(x) = f(x_0).$$

Definition 3.13. Let $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ and $x_0 \in D$ be a limit point of D.

(1) We say that f has a removable discontinuity at x_0 if

$$\lim_{x \nearrow x_0} f(x) = \lim_{x \searrow x_0} f(x) \neq f(x_0).$$

(2) We say that f has a jump discontinuity (or a discontinuity of first kind) at x_0, if the two sided limit exist, but
$$\lim_{x \nearrow x_0} f(x) \neq \lim_{x \searrow x_0} f(x).$$

(3) We say that f has a discontinuity of second kind, if it is not removable or jump discontinuity.

Example 3.14. Consider the function
$$f(x) = \begin{cases} x \sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 2 & \text{if } x = 0\,. \end{cases}$$
Observing that
$$\lim_{x \to 0} x \sin(\frac{1}{x}) = 0 \neq 2 = f(0)\,,$$
we conclude that f has a removable discontinuity at 0.

Example 3.15. Consider the function
$$f(x) = \begin{cases} x - 1 & \text{if } x < 3 \\ x^2 + 1 & \text{if } x \geq 3\,. \end{cases}$$
Observing that
$$\lim_{x \nearrow 3} f(x) = 2 \neq 10 = \lim_{x \searrow 3} f(x)\,,$$
we conclude that f has a jump discontinuity at 3.

Example 3.16. Consider the function
$$f(x) = \begin{cases} \sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0\,. \end{cases}$$
Observe that
$$\lim_{x \to 0} f(x) \ \text{ does not exist }.$$
Indeed, is we consider the sequence $x_n = \frac{1}{n\pi}$, $n \in \mathbb{N}$, then
$$\lim_{n \to \infty} f(x_n) = 0\,.$$
However, if we consider the sequence $y_n = \frac{2}{(4n+1)\pi}$, $n \in \mathbb{N}$, then
$$\lim_{n \to \infty} f(y_n) = 1\,.$$
Therefore, f has a discontinuity of second kind at 0.

Definition 3.17. Let $f : (a, b) \to \mathbb{R}$.
We say that f is monotonically increasing on (a, b) if $a < x_1 < x_2 < b$ implies $f(x_1) \le f(x_2)$.
We say that f is monotonically decreasing on (a, b) if $a < x_1 < x_2 < b$ implies $f(x_1) \ge f(x_2)$.

Theorem 3.18.

(i) *Let $f : (a, b) \to \mathbb{R}$ be monotonically increasing. Then for every $x_0 \in (a, b)$ the sided limits exist, and*

$$\sup_{x<x_0} f(x) = \lim_{x\nearrow x_0} f(x) \le f(x_0) \le \lim_{x\searrow x_0} f(x) = \inf_{x_0<x} f(x)\,.$$

(ii) *Let $f : (a, b) \to \mathbb{R}$ be monotonically decreasing. Then for every $x_0 \in (a, b)$ the sided limits exist, and*

$$\inf_{x<x_0} f(x) = \lim_{x\nearrow x_0} f(x) \ge f(x_0) \ge \lim_{x\searrow x_0} f(x) = \sup_{x_0<x} f(x)\,.$$

Proof. (i) Consider

$$A = \{f(x) \mid a < x < x_0\}.$$

The monotone increasing property of f implies that the set A is bounded from above by $f(x_0)$, so by the Dedekind completeness of $\mathbb{R}$, the $\sup A = \alpha$ exists. Therefore,

$$\sup_{x<x_0} f(x) = \alpha \le f(x_0)\,,$$

and for any $\varepsilon > 0$, there exists $a < x_\varepsilon < x_0$ such that

$$\alpha - \varepsilon < f(x_\varepsilon) \le \alpha\,.$$

Define

$$\delta = \delta(\varepsilon, x_0) = x_0 - x_\varepsilon\,.$$

Then, $x_\varepsilon = x_0 - \delta$ and, by the monotone increasing property of f, for all $x_\varepsilon < x < x_0$ we have

$$\alpha - \varepsilon < f(x_\varepsilon) \le f(x) \le \alpha < \alpha + \varepsilon\,,$$

which implies that

$$|f(x) - \alpha| < \varepsilon\,.$$

Hence,

$$\lim_{x\nearrow x_0} f(x) = \alpha = \sup_{x<x_0} f(x) \le f(x_0)\,.$$

The proof of

$$f(x_0) \le \lim_{x\searrow x_0} f(x) = \inf_{x_0<x} f(x)$$

is almost identical.

(ii) The proof is similar to the proof of (i).

□

The above theorem implies the following result about discontinuities of monotone functions.

Corollary 3.19. *Let $f : (a, b) \to \mathbb{R}$ be a monotone increasing or decreasing function. Then f can have only jump discontinuities.*

Theorem 3.20. *Let $f : (a, b) \to \mathbb{R}$ be a monotone increasing or decreasing function. Then f can have at most countably many discontinuity points.*

Proof. Let us assume that f has at least two discontinuity points. Let us pick any two of them, $a < x_1 < x_2 < b$. Then, by Theorem 3.18 we obtain

$$\lim_{x \nearrow x_1} f(x) < \lim_{x \searrow x_1} f(x) \leq \lim_{x \nearrow x_2} f(x) < \lim_{x \searrow x_2} f(x)\,.$$

By the density of rational numbers, we can choose $r_1, r_2 \in \mathbb{Q}$ such that

$$\lim_{x \nearrow x_1} f(x) < r_1 < \lim_{x \searrow x_1} f(x) \leq \lim_{x \nearrow x_2} f(x) < r_2 < \lim_{x \searrow x_2} f(x)\,.$$

Hence we can establish a one-to-one correspondence between the discontinuity points and a subset of $\mathbb{Q}$, which implies that we can have at most countably many discontinuities.

$\square$

For limits involving infinities, we have the following definitions.

Definition 3.21. Let $f : \mathbb{R} \to \mathbb{R}$.

(1) We say that
$$\lim_{x \to +\infty} f(x) = \alpha$$
if for all $\varepsilon > 0$ there exists $M = M(\varepsilon) > 0$ such that
$$|f(x) - \alpha| < \varepsilon\,,\ \forall\ x > M\,.$$

(2) We say that
$$\lim_{x \to -\infty} f(x) = \alpha$$
if for all $\varepsilon > 0$ there exists $M = M(\varepsilon) > 0$ such that
$$|f(x) - \alpha| < \varepsilon\,,\ \forall\ x < -M\,.$$

(3) If $x_0 \in \mathbb{R}$, we say that
$$\lim_{x \to x_0} f(x) = +\infty$$
if for all $M > 0$ there exists $\delta = \delta(M) > 0$ such that
$$f(x) > M\,,\ \forall\ 0 < |x - x_0| < \delta\,.$$

(4) We say that
$$\lim_{x \to x_0} f(x) = -\infty$$
if for all $M > 0$ there exists $\delta = \delta(M) > 0$ such that
$$f(x) < -M \,, \ \forall \ \ 0 < |x - x_0| < \delta \,.$$

(5) We say that
$$\lim_{x \to +\infty} f(x) = +\infty$$
if for all $N > 0$ there exists $M = M(N) > 0$ such that
$$f(x) > N \,, \ \forall \ \ x > M \,.$$

(6) We say that
$$\lim_{x \to -\infty} f(x) = +\infty$$
if for all $N > 0$ there exists $M = M(N) > 0$ such that
$$f(x) > N \,, \ \forall \ \ x < -M \,.$$

(7) We say that
$$\lim_{x \to +\infty} f(x) = -\infty$$
if for all $N > 0$ there exists $M = M(N) > 0$ such that
$$f(x) < -N \,, \ \forall \ \ x > M \,.$$

(8) We say that
$$\lim_{x \to -\infty} f(x) = -\infty$$
if for all $N > 0$ there exists $M = M(N) > 0$ such that
$$f(x) < -N \,, \ \forall \ \ x < -M \,.$$

Exercises

Exercise 3.1. Let $f : \mathbb{R} \to \mathbb{R}$ defined by
$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin Q \,. \end{cases}$$
Show that f has discontinuities of second kind at every point.

Exercise 3.2. Let $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin Q\,. \end{cases}$$

Show that f is continuous at $x = 0$ and has discontinuities of second kind at every $x \neq 0$.

Exercise 3.3. Let $f : [0, 1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{1}{n} & \text{if } x = \frac{m}{n} \quad \text{in lowest terms} \\ 1 & \text{if } x = 0 \\ 0 & \text{if } x \notin Q\,. \end{cases}$$

Show that f is continuous at irrational points and has removable discontinuity at every $x \in \mathbb{Q} \cap [0, 1]$.

Exercise 3.4. Let X be a metric space, $D \subset X$, $x_0 \in X$ be a limit point of D and $f, g, h : D \to \mathbb{R}$. Assume that there exists $r > 0$ such that

$$f(x) \le g(x) \le h(x),\ \forall\, x \in B_r(x_0) \cap D\,,\ x \neq x_0\,.$$

Prove that if

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x) = \alpha\,,$$

then

$$\lim_{x\to x_0} g(x) = \alpha\,.$$

Exercise 3.5. Prove, without the L'Hopital's rule, that

$$\lim_{x\to 0} \frac{\sin x}{x} = 1\,.$$

Exercise 3.6. Prove that

$$\lim_{x\to 0} x \sin\left(\frac{1}{x}\right) = 0\,.$$

Exercise 3.7. Let $f : \mathbb{R} \to \mathbb{R}$ and $x_0 \in \mathbb{R}$. Show that if f is continuous at x_0, then $|f|$ is continuous at x_0.

Exercise 3.8. Let $f, g : \mathbb{R} \to \mathbb{R}$ and $x_0 \in \mathbb{R}$. Show that if f and g are continuous at x_0, then $\max\{f, g\}$ and $\min\{f, g\}$ are continuous at x_0.

Hint:

$$\max\{f(x), g(x)\} = \frac{f(x) + g(x) + |f(x) - g(x)|}{2},$$
$$\min\{f(x), g(x)\} = \frac{f(x) + g(x) - |f(x) - g(x)|}{2}.$$

Exercise 3.9. Let X be a metric space, $f : X \to \mathbb{R}$ and $x_0 \in X$. Show that if f is continuous at x_0 and $f(x_0) > 0$ then there exists $r > 0$ such that

$$f(x) > \frac{f(x_0)}{2}, \quad \forall\, x \in B_r(x_0).$$

Exercise 3.10. Let $f : [0, +\infty) \to \mathbb{R}$ be a monotonically increasing function. Prove that either $\lim_{x \to +\infty} f(x)$ exists, or $\lim_{x \to +\infty} f(x) = +\infty$.

3.2 Continuous functions on a set

In this section we consider the domain of the function be the entire metric space X. When working with metric spaces, we can always restrict the metric space to a subset, if necessary.

Definition 3.22. Let X, Y be metric spaces, $D \subset X$ and $f : X \to Y$. We say that f is continuous on D, if f is continuous at every $x \in D$.

Theorem 3.23. *Let X, Y be metric spaces and $f : X \to Y$ be a function. Then f is continuous on X if and only if for every open set $V \subset Y$, the inverse image $f^{-1}(V)$ is open in X.*

Proof. First, suppose that f is continuous on X and let $V \subset Y$ be an open set.
If $f^{-1}(V) = \emptyset$, then it is open. Otherwise, consider any $x \in f^{-1}(V)$. Then $y = f(x) \in V$ and by the openess of V, there exists $\varepsilon > 0$ such that $B_\varepsilon(y) \subset V$. The continuity if f at x implies that there exists $\delta = \delta(\varepsilon, x) > 0$ such that $f(B_\delta(x) \subset B_\varepsilon(y) \subset V$. Hence, $B_\delta(x) \subset f^{-1}(V)$. In conclusion, we proved that any point of $f^{-1}(V)$ is an interior point and therefore $f^{-1}(V)$ is open.

Second, suppose that for every open set $V \subset Y$, the inverse image $f^{-1}(V)$ is open. Consider any $x \in X$ and $\varepsilon > 0$. Let $V = B_\varepsilon(f(x))$ which is an open ball in Y. Then $f^{-1}(B_\varepsilon(f(x))$ is open. However,

$$x \in f^{-1}(B_\varepsilon(f(x)),$$

so there exists $\delta > 0$ such that

$$B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x)) .$$

This show that for all $x' \in B_\delta(x)$ we have $f(x') \in B_\varepsilon(f(x))$, which means that f is continuous at x. But, $x \in X$ has been chosen as an arbitrary point, so f is continuous at every point of X.

□

The proof of the following corollary depends in the relation

$$X \setminus f^{-1}(W) = f^{-1}(Y \setminus W) .$$

Corollary 3.24. *Let X, Y be metric spaces and $f : X \to Y$ be a function. Then f is continuous on X if and only if for every closed set $W \subset Y$, the inverse image $f^{-1}(W)$ is closed in X.*

Remark 3.25. Note that the open and closed property of the inverse image is relative to the domain of f, which we assume to be the metric space X. If we consider $f(x) = \sqrt{x}$, then we have to set the metric spaces $X = [0, +\infty)$ and $Y = \mathbb{R}$. The inverse image of $V = (-5, 3)$ is $f^{-1}(V) = [0, 9)$, which is open in X.

Theorem 3.26. *Let X and Y be metric spaces and $f : X \to Y$ be a continuous function. If X is compact, then $f(X)$ is compact.*

Proof. Consider an open cover of $f(X)$:

$$f(X) \subset \bigcup_{i\in I} V_i \,.$$

Therefore,

$$X \subset f^{-1}(f(X)) \subset f^{-1}\Big(\bigcup_{i\in I} V_i\Big) = \bigcup_{i\in I} f^{-1}(V_i) \,.$$

Using the continuity of f, we can observe that $\{f^{-1}(V_i)\}$ forms an open cover of X. The compactness of X allows us to select an finite subcover:

$$X \subset \bigcup_{k=1}^{n} f^{-1}\Big(V_{i_k}\Big) = f^{-1}\Big(\bigcup_{k=1}^{n} V_{i_k}\Big) \,.$$

Then,

$$f(X) \subset f\left(f^{-1}\Big(\bigcup_{k=1}^{n} V_{i_k}\Big)\right) \subset \bigcup_{k=1}^{n} V_{i_k} \,,$$

which provides the finite subcover of $f(X)$. Hence, $f(X)$ is compact. $\square$

Next, we show two very important consequences of Theorem 3.26.

Theorem 3.27. *Let X be a compact metric space and $f : X \to \mathbb{R}$ be a continuous function on X. Then there exist $x_*, x^* \in X$ such that*

$$f(x_*) = \inf_{x\in X} f(x) \quad \textit{and} \quad f(x^*) = \sup_{x\in X} f(x) \,.$$

Note: When the infimum and supremum are attained, we call them minimum and maximum and use the notations:

$$f(x_*) = \min_{x\in X} f(x) \,, \ \text{ and } \ f(x^*) = \max_{x\in X} f(x) \,.$$

Proof. The set $f(X)$ is compact in $\mathbb{R}$, therefore is bounded and closed. The boundedness implies that the infimum and supremum exist and are finite. By the properties of infimum, for every $\varepsilon = \frac{1}{n}$, $n \in \mathbb{N}$ there exists $x_n \in X$ such that

$$\inf_{x\in X} f(x) \le f(x_n) < \inf_{x\in X} f(x) + \frac{1}{n} \,.$$

The sequence $\{x_n\}$ is from the compact space X, therefore we can select a convergent subsequence and its limit:

$$\{x_{n_k}\} \to x_* \in X \,.$$

The continuity of the function f implies that

$$\{f(x_{n_k})\} \to f(x_*)\,,$$

and therefore, from the inequality above, we find

$$\inf_{x\in X} f(x) = f(x_*)\,.$$

The proof for the supremum is similar, and it is left as an exercise.

$\square$

Theorem 3.28. *Let X be a compact metric space and Y be a metric space. If $f : X \to Y$ is one-to-one, onto and continuous, then $f^{-1} : Y \to X$ is continuous.*

Proof. Since f is one-to-one and onto, we have $(f^{-1})^{-1} = f$. Therefore, we have to show that for every open set $U \subset X$, it follows that $f(U)$ is open.

Let $U \subset X$ be an open set. Then U^c is closed and also compact as a closed subset of a compact metric space. Therefore, by Theorem 3.26, we know that $f(U^c)$ is compact in Y, hence closed. Again, since f is one-to-one and onto, we obtain that $f(U^c) = f(U)^c$, and by this, $f(U)$ is open.

$\square$

Example 3.29. In general the compactness is necessary. Consider $X = [0, 2\pi)$ and $Y = S^1$, the unit circle in $\mathbb{R}^2$. The function $f : X \to Y$, $f(t) = (\cos t, \sin t)$ is one-to-one, onto and continuous. However, its inverse function is not continuous at the point $(1, 0)$. Indeed,

$$f^{-1}\Big(\cos\big(\frac{1}{n}\big), \sin\big(\frac{1}{n}\big)\Big) = \frac{1}{n} \to 0\,,$$

and

$$f^{-1}\Big(\cos\big(2\pi - \frac{1}{n}\big), \sin\big(2\pi - \frac{1}{n}\big)\Big) = 2\pi - \frac{1}{n} \to 2\pi \notin X\,.$$

See also Exercise 3.11.

Definition 3.30. Let X be a metric space and $A, B \subset X$.
We say that A and B are separated if $\overline{A} \cap B = \emptyset$ and $A \cap \overline{B} = \emptyset$.
We say that $C \subset X$ is a connected set, if we cannot write C as the union of two non-empty separated sets.

In the next theorem we use the expression "interval" for any of the following: $[a, b]$, $[a, b)$, $(a, b]$, (a, b), $(-\infty, a)$, $(-\infty, a]$, $(b, +\infty)$, $[b, +\infty)$.

Proposition 3.31. *The only connected sets in $\mathbb{R}$ are the singletons and the intervals.*

Proof. The proof follows from the following result: If $C \subset \mathbb{R}$ is connected, $x, y \in C$ and $x < z < y$, then $z \in C$. To prove this result, assume that $z \notin C$ and this leads to a contradiction with C being connected:

$$C = A \bigcup B\,, \quad A = C \bigcap (-\infty, z)\,, \quad B = C \bigcap (z, +\infty)\,.$$

We leave the remaining details of the proof to the reader.

□

Theorem 3.32. *Let X, Y be metric spaces and $f : X \to Y$ be a continuous mapping. If $C \subset X$ is a connected set, then $f(C) \subset Y$ is connected, too.*

Proof. Consider a connected $C \subset X$ and assume that $f(C) \subset Y$ is not connected. Then there exist non-empty separated sets $A, B \subset Y$ such that $f(C) = A \bigcup B$. Let

$$U = C \bigcap f^{-1}(A)\,, \quad V = C \bigcap f^{-1}(B)\,.$$

Then $U \neq \emptyset$, $V \neq \emptyset$ and $C = U \bigcup V$.
Since $A \subset \overline{A}$, we have $f^{-1}(A) \subset f^{-1}(\overline{A})$ and hence $U \subset f^{-1}(\overline{A})$. The continuity of f implies that $f^{-1}(\overline{A})$ is closed and therefore $\overline{U} \subset f^{-1}(\overline{A})$. It follows now that $f(\overline{U}) \subset \overline{A}$.
However, $f(V) = B$ and $\overline{A} \bigcap B = \emptyset$, which imply that $f(\overline{U}) \bigcap f(V) = \emptyset$. Observe that

$$f(\overline{U} \bigcap V) \subset f(\overline{U}) \bigcap f(V) = \emptyset\,.$$

We can conclude now that $\overline{U} \bigcap V = \emptyset$.
Similarly, we can prove that $U \bigcap \overline{V} = \emptyset$, which implies that U and V are separated, and this contradicts the connectedness of C.

□

We can prove now the Intermediate Value Theorem for continuous functions.

Theorem 3.33. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. If $f(a) < c < f(b)$, then there exists $x \in (a, b)$ such that $f(x) = c$.*

Proof. By Theorem 3.26 and Theorem 3.32, $f([a, b])$ is a closed, bounded interval. Therefore, if $f(a) < c < f(b)$, then $c \in f([a, b])$, so there exists $x \in [a, b]$ such that $f(x) = c$. Note that x cannot be a or b, so $x \in (a, b)$.

□

Exercises

Exercise 3.11. If $X \subset \mathbb{R}$, can we remove the compactness from the assumptions of Theorem 3.28?

Exercise 3.12. Let X, Y be metric spaces and $f : X \to Y$ be a continuous function. Prove that for all $A \subset X$ we have

$$f(A) \subset f(\overline{A}) \subset \overline{f(A)}\,.$$

Exercise 3.13. Let X, Y be metric spaces and $f : X \to Y$ be a continuous and onto function. Prove that if $A \subset X$ is dense in X, then $f(A)$ is dense in Y.

Exercise 3.14. Let X be a metric space and $f : X \to \mathbb{R}$ be a continuous function. Prove that the following set is closed: $Z(f) = \{x \in X : \;\; f(x) = 0\}\,.$

Exercise 3.15. Let $f : [0, 1] \to [0, 1]$ be a continuous function. Show that there exists $x \in [0, 1]$ such that $f(x) = x$.

Exercise 3.16. Let $f : [a, b] \to \mathbb{R}$ be a continuous functions such that its range is at most countable. Show that f is a constant function.

Exercise 3.17. Let $f : [0, 2] \to \mathbb{R}$ be a continuous functions. Prove that if $f(0) = f(2)$, then there exists $x \in [0, 1]$ such that $f(x) = f(x + 1)$.

Exercise 3.18. Let $f : [a, b] \to \mathbb{R}$ be a function. We say that f has the Intermediate Value Property, if $a \le x_1 < x_2 \le b$ and $f(x_1) < y < f(x_2)$ implies that there exists $x_1 < x < x_2$ such that $f(x) = y$.

(a) Give an example of a function which has the Intermediate Value Property, but which is not continuous.

(b) Show that if a function has the Intermediate Value Property, then it can have only discontinuities of second kind.

Exercise 3.19. Let $f : [a, b] \to \mathbb{R}$ be monotonically increasing. Prove that if f satisfies the Intermediate Value Property, then f is continuous.

Exercise 3.20. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Define $g : [a, b] \to \mathbb{R}$ by

$$g(x) = \begin{cases} f(a) & \text{if } \; x = a \\ \max_{a \le t \le x} f(t) & \text{if } \; a < x \le b\,. \end{cases}$$

Prove that g is monotonically increasing and continuous on $[a, b]$.

Exercise 3.21. Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that if $A \subset \mathbb{R}$ is open, then $f(A)$ is open (functions with this property are called open functions). Prove that if f is a continuous and open function, then f is monotonically increasing or decreasing on $\mathbb{R}$.

Exercise 3.22. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function such that $f(x) > 0$ for all $x \in \mathbb{R}$ and $\lim_{x\to\pm\infty} f(x) = 0$. Prove that there exists $x_0 \in \mathbb{R}$ such that $f(x_0) \geq f(x)$ for all $x \in \mathbb{R}$.

Exercise 3.23. Let $f : \mathbb{R} \to \mathbb{R}$ be a function which satisfies the Intermediate Value Property. Assume that for each $r \in \mathbb{Q}$, the set $f^{-1}(\{r\})$ is closed. Show that f is continuous on $\mathbb{R}$.

Exercise 3.24. Prove that any connected metric space with at least two points is uncountable.

3.3 Uniform, absolute and Lipschitz continuity

Definition 3.34. Let X, Y be metric spaces and $f : X \to Y$.

(i) We say that f is uniformly continuous on X if for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that $d(f(x_1), f(x_2)) < \varepsilon$ for all $x_1, x_2 \in X$ with $d(x_1, x_2) < \delta$.

(ii) We say that f is Lipschitz continuous on X if there exists $L \geq 0$ such that

$$d(f(x_1), f(x_2)) \leq L\, d(x_1, x_2)\,, \quad \forall\ x_1, x_2 \in X\,.$$

The number L is called the Lipschitz constant of f.

(iii) If f is Lipschitz continuous and $0 \leq L < 1$, then we say that f is a contraction.

Theorem 3.35. *Let X, Y be metric spaces and $f : X \to Y$. Regarding the continuity of f on X, the following implications hold:*

$$\textit{Lipschitz continuous} \implies \textit{uniformly continuous} \implies \textit{continuous}\,.$$

Proof. If f is Lipschitz continuous, then for any $\varepsilon > 0$ we can choose $\delta = \delta(\varepsilon) = \frac{\varepsilon}{L+1}$, which shows that f is uniformly continuous on X.

If f is uniformly continuous, then for any $x_0 \in X$ and $\varepsilon > 0$ we can choose $\delta(\varepsilon, x_0) = \delta(\varepsilon)$, which shows that f is continuous at x_0.

□

Example 3.36. There are functions which are uniformly continuous, but not Lipschitz continuous. Consider $f : [0, +\infty) \to \mathbb{R}$, $f(x) = \sqrt{x}$.

First, let's prove that f is uniformly continuous. We will show that for $\varepsilon > 0$ and $x_0 \geq 0$ we can choose the $\delta(\varepsilon)$ independently of x_0. Let's fix an arbitrary $\varepsilon > 0$.
If $x_0 = 0$, then we can choose $\delta = \delta(\varepsilon) = \varepsilon^2$, because if $0 \leq x < \varepsilon^2$, then $\sqrt{x} < \varepsilon$.
If $0 < x_0 < \varepsilon^2$ and $0 \leq x < \varepsilon^2$,

$$|\sqrt{x} - \sqrt{x_0}| < \max\{\sqrt{x}, \sqrt{x_0}\}\,,$$

which shows that we can choose again $\delta = \delta(\varepsilon) = \varepsilon^2$. If $0 < x_0 < \varepsilon^2$ and $\varepsilon^2 \leq x$ or $\varepsilon^2 \leq x_0$, then

$$|\sqrt{x} - \sqrt{x_0}| = \frac{|x - x_0|}{\sqrt{x} + \sqrt{x_0}} \leq \frac{1}{\varepsilon}|x - x_0|\,,$$

which shows that we can choose again $\delta = \delta(\varepsilon) = \varepsilon^2$. In conclusion, to cover all cases, we can choose $\delta = \delta(\varepsilon) = \varepsilon^2$.

Second, let's assume that f is Lipschitz continuous and for any $n \in \mathbb{N}$ consider $x_{1_n} = \frac{1}{n^2}$ and $x_{2_n} = \frac{1}{(n+1)^2}$. Then the Lipschitz condition implies that

$$\frac{1}{n} - \frac{1}{n+1} \leq L\left(\frac{1}{n^2} - \frac{1}{(n+1)^2}\right)\,,$$

which leads to

$$\frac{n^2+n}{2n+1} \leq L\,,\ \forall\ \ n \in \mathbb{N}\,.$$

This gives a contradiction, since the supremum of the left side is $+\infty$.

Example 3.37. There are functions which are continuous, but not uniformly continuous. Consider $f : [0, +\infty) \to \mathbb{R}$, $f(x) = x^2$. For $n \in \mathbb{N}$ consider $x_{0_n} = n$. Let $\varepsilon > 0$ and assume that we were able to choose a $\delta = \delta(\varepsilon) > 0$. Then, if $|x - n| < \delta$, then the inequality

$$|x^2 - n^2| = |x + n| \cdot |x - n| \leq (2n + \delta)\delta < \varepsilon$$

leads to

$$\delta^2 + 2n\delta - \varepsilon^2 < 0\,,$$

which gives

$$0 < \delta < \sqrt{n^2 + \varepsilon^2} - n = \frac{\varepsilon^2}{\sqrt{n^2 + \varepsilon^2} + n}\,.$$

As the infimum of the right hand side is 0, we conclude that no choice of $\delta(\varepsilon) > 0$, independent of n, is possible. Hence, the function is not uniformly continuous.

Theorem 3.38. *Let X and Y be metric spaces and $f : X \to Y$ be uniformly continuous on X. Then, if $\{x_n\}$ is a Cauchy sequence in X, then $\{f(x_n\}$ is a Cauchy sequence in Y.*

Proof. Let $\varepsilon > 0$. By the uniform continuity of f, there exists $\delta = \delta(\varepsilon) > 0$ such that $d(f(x), f(y) < \varepsilon$ for all $x, y \in X$ with $d(x, y) < \delta$.

By the definition of a Cauchy sequence, for the $\delta = \delta(\varepsilon) > 0$ chosen above, there exists $N = N(\delta) = N(\varepsilon) > 0$ such that $d(x_n, x_m) < \delta$ for all $n, m > N$.

In conclusion, for the $\varepsilon > 0$ we found $N = N(\varepsilon) > 0$ such that $d(f(x_n), f(x_m)) < \varepsilon$ for all $n, m > N$, and this shows that $\{f(x_n\}$ is a Cauchy sequence.

□

Theorem 3.39. *Let $f : (a, b) \to \mathbb{R}$ be a uniformly continuous function on (a, b). Then*

$$\lim_{x \searrow a} f(x) \quad \textit{and} \quad \lim_{x \nearrow b} f(x)$$

exist, and we can extend f to a continuous function $\hat{f} : [a, b] \to \mathbb{R}$.

Proof. Consider a sequence $\{x_n\}$ in (a, b) such that $\{x_n\} \to a$. While this sequence is convergent in $\mathbb{R}$, it is only a Cauchy sequence in (a, b), which is the domain of f. By Theorem 3.38, $\{f(x_n)\}$ is a Cauchy sequence in $\mathbb{R}$. Therefore, $\{f(x_n)\} \to \alpha \in \mathbb{R}$.

We have to show that for any other sequence $\{x'_n\}$ in (a, b) such that $\{x'_n\} \to a$ we still get $\{f(x'_n)\} \to \alpha$.
By the uniform continuity of f, for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that

$$d(f(x), f(y) < \frac{\varepsilon}{2} \text{ for all } x, y \in X \text{ with } d(x, y) < \delta\,.$$

From the fact that both sequences, $\{x_n\}$ and $\{x'_n\}$, converge to a, we get that for the $\delta > 0$, chosen above, there exists $N_1 = N_1(\delta) = N_1(\varepsilon) > 0$ such that for all $n > N_1$ we have

$$|x_n - a| < \frac{\delta}{2} \text{ and } |x'_n - a| < \frac{\delta}{2}.$$

Thereofore,

$$|x_n - x'_n| \leq |x_n - a| + |x'_n - a| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

By the fact that $\{f(x_n)\} \to \alpha$, there exists $N_2 = N_2(\varepsilon)$ such that for all $n > N_2$ we have

$$|f(x_n) - \alpha| < \frac{\varepsilon}{2}.$$

Let $N = \max\{N_1, N_2\}$. Therefore, for all $n > N$ we have

$$|f(x'_n) - \alpha| \leq |f(x'_n) - f(x_n)| + |f(x_n) - \alpha| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and this shows that $\{f(x'_n)\} \to \alpha$. In conclusion, we can define

$$\lim_{x \searrow a} f(x) = \alpha.$$

In a similar way, we can prove that the left sided limit at b exists:

$$\lim_{x \nearrow b} f(x) = \beta.$$

The continuous extension of f can be defined as

$$\hat{f}(x) = \begin{cases} f(x) & \text{if } x \in (a, b) \\ \alpha & \text{if } x = a \\ \beta & \text{if } x = b. \end{cases}$$

$\square$

Example 3.40. The function $f : (0,1) \to \mathbb{R}$, $f(x) = \frac{1}{x}$ doesn't have a continuous extension to $[0,1]$. What goes wrong if you try adapting the proof of the above theorem to this function?

Theorem 3.41. *Let X, Y be metric spaces and $f : X \to Y$. If X is compact and f is continuous on X, then f is uniformly continuous on X.*

Proof. Let $\varepsilon > 0$ be arbitrary, but fixed. For any $x \in X$ we can choose $\delta(\varepsilon, x) > 0$ such that if $d(z, x) < \delta(\varepsilon, x)$, then $d(f(z), f(x)) < \frac{\varepsilon}{2}$. Consider the following open cover of X:

$$X \subset \bigcup_{x \in X} B_{\frac{1}{2}\delta(\varepsilon,x)}(x).$$

The compactness of X implies that we can select a finite subcover

$$X \subset \bigcup_{k=1}^{n} B_{\frac{1}{2}\delta(\varepsilon,x_k)}(x_k)\,.$$

Let

$$\delta(\varepsilon) = \min\{\frac{1}{2}\delta(\varepsilon,x_k) \mid 1 \le k \le n\}\,.$$

To check the uniform continuity of f let $x, z \in X$ such that $d(x,z) < \delta(\varepsilon)$. From the finite subcover, we select a point x_k such that $x \in B_{\frac{1}{2}\delta(\varepsilon,x_k)}(x_k)$. Then, $d(x,x_k) < \frac{1}{2}\delta(\varepsilon,x_k)$, but also

$$d(z,x_k) \le d(z,x) + d(x,x_k) < \delta(\varepsilon) + \frac{1}{2}\delta(\varepsilon,x_k) \le \delta(\varepsilon,x_k)\,.$$

Therefore,

$$d(f(x),f(z)) \le d(f(x),f(x_k)) + d(f(x_k),z) \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,,$$

and by this we proved that f is uniformly continuous.

□

For functions $f : [a,b] \to \mathbb{R}$ we introduce another notion of continuity.

Definition 3.42. A function $f : [a,b] \to \mathbb{R}$ is called absolutely continuous on $[a,b]$, if for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for any collection of pairwise disjoint intervals $\{(c_k,d_k),\ 1 \le k \le n\}$ from $[a,b]$, with

$$\sum_{k=1}^{n}(d_k - c_k) < \delta,$$

it follows

$$\sum_{k=1}^{n}|f(d_k) - f(c_k)| < \varepsilon\,.$$

Note that, as the number of intervals is independent of ε, in Definition 3.42, we can have countably many intervals.

The proof of the following theorem is left as an exercise.

Theorem 3.43. *Let $f : [a,b] \to \mathbb{R}$. Regarding the continuity of f on $[a,b]$, the following implications hold:*

$$\textit{Lipschitz continuous} \implies \textit{absolutely continuous} \implies \textit{uniformly continuous}\,.$$

Example 3.44. The function from Example 3.36 is absolutely continuous, but not Lipschitz continuous. We will prove the absolute continuity part in the chapter for integration.

Example 3.45. The Cantor function.
This is an example of a uniformly continuous function, which is not absolutely continuous.

We will use the Cantor set, as defined in Section 2.6. We remind that it can be defined as

$$C = \left\{ x \in [0,1] \,\middle|\, x = \sum_{n=1}^{\infty} \frac{2a_n}{3^n}\,, \quad \text{where} \quad a_n = 0 \text{ or } 1 \right\}.$$

We define the Cantor function $f : [0,1] \to [0,1]$ as

$$f(x) = \begin{cases} \sum_{n=1}^{\infty} \frac{a_n}{2^n} & \text{if } \; x = \sum_{n=1}^{\infty} \frac{2a_n}{3^n} \in C \\ \\ \sup\{f(c) \mid c \leq x, c \in C\} & \text{if } \; x \notin C\,. \end{cases}$$

Notice that

$$f\left(\frac{1}{3^n}\right) = f\left(\frac{2}{3^{n+1}} + \frac{2}{3^{n+2}} + \dots\right) = \frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \dots = \frac{1}{2^n}$$

and

$$f\left(\frac{2}{3^n}\right) = \frac{1}{2^n}\,.$$

This relation can be generalized to the endpoints of the deleted intervals during the construction of the Cantor set, and shows that the Cantor function is constant over these deleted intervals.

Observing that

$$\sum_{n=1}^{\infty} \frac{2a_n}{3^n} \leq \sum_{n=1}^{\infty} \frac{2b_n}{3^n} \Longleftrightarrow \sum_{n=1}^{\infty} \frac{a_n}{2^n} \leq \sum_{n=1}^{\infty} \frac{b_n}{2^n},$$

leads to the fact that the Cantor function is monotonically increasing on the Cantor set. This implies that the Cantor function is monotonically increasing on $[0,1]$.

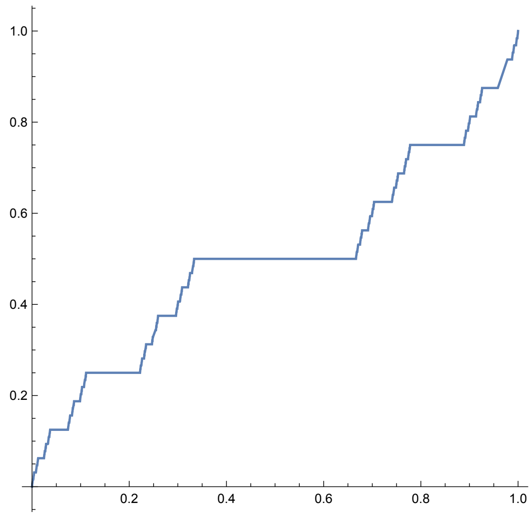
Moreover, the points in the Cantor set satisfy the following relation

$$\left| \sum_{n=1}^{\infty} \frac{2a_n}{3^n} - \sum_{n=1}^{\infty} \frac{2b_n}{3^n} \right| < \frac{1}{3^m} \Longrightarrow a_n = b_n, \forall\, 1 \leq n \leq m\,,$$

which implies that

$$\left| f\left(\sum_{n=1}^{\infty} \frac{2a_n}{3^n}\right) - f\left(\sum_{n=1}^{\infty} \frac{2b_n}{3^n}\right) \right| < \frac{1}{2^m}.$$

Hence, the Cantor function is continuous on the Cantor set and, by the fact that the Cantor set is a perfect set, we can conclude that the extension of the Cantor function to $[0,1]$ is continuous on $[0,1]$.



Let us list the properties of the Cantor function in the following proposition.

Proposition 3.46. *The Cantor function has the following properties:*

(a) *It is monotonically increasing and uniformly continuous on* $[0, 1]$.

(b) *It is not absolutely continuous on* $[0, 1]$.

(c) *It is constant over each interval excluded from the Cantor set.*

(d) *Maps a set of measure* 0 *into a set of measure* 1.

Proof. (a) In Example 3.45 we highlighted the facts leading to the monotone increasing and continuity properties. Moreover, as a function, which is continuous on a compact set, the Cantor function is uniformly continuous.

(b) The set C_n from the construction of the Cantor set contains 2^n intervals, each of length $\frac{1}{3^n}$. Let us denote

$$C_n = [c_1, d_1] \bigcup [c_2, d_2] \bigcup \cdots \bigcup [c_{2^n}, d_{2^n}] .$$

Then,

$$\sum_{k=1}^{2^n} (d_k - c_k) = \frac{2^n}{3^n} .$$

We leave the reader to verify that

$$f(d_k) - f(c_k) = \frac{1}{2^n} , \ \forall \ 1 \le k \le 2^n.$$

Therefore,

$$\sum_{k=1}^{2^n} (f(d_k) - f(c_k)) = \frac{2^n}{2^n} = 1 ,$$

which shows that we can choose a collection of intervals with arbitrary small total length, for which the Cantor function gives a fixed total variation of 1. This shows that the Cantor function is not absolutely continuous.

(c) It follows from the discussion in Example 3.45.

(d) It follows from the observation that $f(C) = [0, 1]$.

$\square$

Theorem 3.47. *Let $f : [a, b] \to \mathbb{R}$ be an absolutely continuous function and $A \subset [a, b]$ with $\mu(A) = 0$. Then, $\mu(f(A)) = 0$.*

Proof. Let $\varepsilon > 0$. By the absolute continuity of f, there exists $\delta = \delta(\varepsilon) > 0$ such that for any collection, finite or countable, of pairwise disjoint intervals $\{(c_k, d_k)\}$ from $[a, b]$, with

$$\sum_k (d_k - c_k) < \delta,$$

it follows

$$\sum_k |f(d_k) - f(c_k)| < \varepsilon\,.$$

By the fact that A has measure zero, consider an open cover of A by intervals with total length less than δ. For this cover, combine the overlapping intervals, so we have an open cover with pairwise disjoint intervals. Let us denote this covering of A again by $\{(c_k, d_k)\}$. Therefore,

$$f(A) \subset \bigcup_k f\Big((c_k, d_k)\Big)\,.$$

Note that for each $k \in \mathbb{N}$, $f\Big((c_k, d_k)\Big)$ is an interval (not necessarily open). The function f is continuous on $[c_k, d_k]$, so it attains its maximum and minimum. If f is not a constant function on $[c_k, d_k]$, lets's denote by $c_k \le x_k < y_k \le d_k$ two points where it attains its minimum and maximum. If f is a constant on $[c_k, d_k]$, then use $x_k = c_k$ and $y_k = d_k$. In this way we obtained a collection of pairwise, disjoint intervals $\{x_k, y_k\}$ such that

$$\sum_k (y_k - x_k) \le \sum_k (d_k - x_k) < \delta\,,$$

and

$$A \subset \bigcup_k [f(x_k), f(y_k)]\ ,\ \text{or}\ \ [f(y_k), f(x_k)]\ .$$

By the absolute continuity of f we have

$$\sum_k |f(y_k) - f(x_k)| < \varepsilon\,,$$

which means that we managed to cover A by intervals with total lenght less then ε.

In conclusion, we obtained that $\mu(f(A)) = 0$.

$\square$

Exercises

Exercise 3.25. Let X, Y and Z be metric spaces, $f : X \to Y$, $g : Y \to Z$ and $h = g \circ f$. Assume that Y is compact and g is one-to-one, onto and continuous. Prove that:

(a) If h is continuous, then f is continuous.

(b) If h is uniformly continuous, then f is uniformly continuous.

Exercise 3.26. Let $A \subset \mathbb{R}$ be a bounded set and $f : A \to \mathbb{R}$ be a uniformly continuous functions. Show that $f(A)$ is bounded.

Exercise 3.27. Let $f : [0, +\infty) \to \mathbb{R}$ be a continuous function. Prove that if $\lim_{x\to+\infty} = L \in \mathbb{R}$, then f is uniformly continuous on $[0, +\infty)$.

Exercise 3.28. Let $f : (a, b) \to \mathbb{R}$ be a continuous function. Show that f is uniformly continuous if and only if f can be extended to a continuous function on $[a, b]$.

Exercise 3.29. Which of the following function are uniformly continuous?

$$\begin{aligned}
&(a)\ \ f : (0,1) \to \mathbb{R},\ \ f(x) = \frac{x^2+1}{x}\\
&(b)\ \ f : (1,2) \to \mathbb{R},\ \ f(x) = \frac{x^2+1}{x}\\
&(c)\ \ f : (0,1) \to \mathbb{R},\ \ f(x) = \sin\frac{1}{x}\\
&(d)\ \ f : (0,1) \to \mathbb{R},\ \ f(x) = x\sin\frac{1}{x}\\
&(e)\ \ f : [0,+\infty) \to \mathbb{R},\ \ f(x) = \sqrt{x}
\end{aligned}$$

Exercise 3.30. Let X be a complete metric space and $f : X \to X$ be a contraction. Prove that there exists a unique $x_0 \in X$ such that $f(x_0) = x_0$.

Chapter 4

Differentiable functions

4.1 Definitions and properties

Definition 4.1. Let $f : [a, b] \to \mathbb{R}$ and $x \in [a, b]$.
If $x \in (a, b)$ and $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ exists, then we say that f is differentiable at x and use the notation

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} .$$

If $x = a$ and $\lim_{h\searrow 0} \frac{f(a+h)-f(a)}{h}$ exists, then we say that f is differentiable at a and use the notation

$$f'(a) = \lim_{h\searrow 0} \frac{f(a+h) - f(a)}{h} .$$

If $x = b$ and $\lim_{t\nearrow 0} \frac{f(b+h)-f(b)}{h}$ exists, then we say that f is differentiable at b and use the notation

$$f'(b) = \lim_{h\nearrow 0} \frac{f(b+h) - f(b)}{h} .$$

We call $f'(x)$ the derivative of f at x.
We say that f is differentiable on $[a, b]$ if it is differentiable at any $x \in [a, b]$.

Except some special cases, we will not distinguish between x being in the interior of the interval or being one of the endpoints. In case of the endpoints, we will automatically consider the one-sided limits.

Lemma 4.2. *Let $f : [a, b] \to \mathbb{R}$ and $x \in [a, b]$. If f is differentiable at x, then there exists a function $\phi(h)$ defined on a small neighborhood of 0 such that*

$$f(x + h) - f(x) = \big(f'(x) + \phi(h)\big)h ,$$

and

$$\lim_{h\to 0} \phi(h) = 0.$$

Proof. If we define

$$\phi(h) = \frac{f(x+h) - f(x)}{h} - f'(x) \ ,$$

then the conclusion of the lemma follows from the definition.

□

Theorem 4.3. *Let $f : [a,b] \to \mathbb{R}$ and $x \in [a,b]$. If f is differentiable at x, then f is continuous at x.*

Proof. By Lemma 4.2 we get that

$$\lim_{h\to 0} f(x+h) - f(x) = 0\,.$$

Therefore,

$$\lim_{h\to 0} f(x+h) = f(x)\,,$$

and this implies the continuity of f at x. □

Theorem 4.4. *Let $f, g : [a,b] \to \mathbb{R}$, $x \in [a,b]$ and $c \in \mathbb{R}$. If f and g are differentiable at x, then $f+g$, cf, fg and $\frac{f}{g}$ (if $g(x) \neq 0$) are differentiable at x and*

(1) $(f+g)'(x) = f'(x) + g'(x)\,.$

(2) $(cf)'(x) = cf'(x)\,.$

(3) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)\,.$

(4) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}\,, \quad$ *if* $g(x) \neq 0.$

Proof. (1) The proof immediately follows from

$$\lim_{h\to 0} \frac{(f+g)(x+h) - (f+g)(x)}{h} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} + \lim_{h\to 0} \frac{g(x+h) - g(x)}{h}\,.$$

(2) Again, as in case of (1), the proof is evidently follows from

$$\lim_{h\to 0} \frac{(cf)(x+h) - (cf)(x)}{h} = c \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}\,.$$

(3) By Lemma 4.2, we use the functions ϕ and ψ are associated to f and g, and get

$$\begin{aligned}
&\frac{f(x+h)g(x+h)-f(x)g(x)}{h}\\
=&\frac{\Big(f(x)+\big(f'(x)+\phi(h)\big)h\Big)\Big(g(x)+\big(g'(x)+\psi(h)\big)h\Big)-f(x)g(x)}{h}\\
=&\big(f'(x)+\phi(h)\big)g(x)+f(x)\big(g'(x)+\psi(h)\big)\,.
\end{aligned}$$

Letting $h \to 0$, the formula (3) follows.

(4) Similarly to (3), we write

$$\begin{aligned}
&\frac{\frac{f(x+h)}{g(x+h)}-\frac{f(x)}{g(x)}}{h}=\frac{f(x+h)g(x)-f(x)g(x+h)}{hg(x)g(x+h)}\\
=&\frac{\Big(f(x)+\big(f'(x)+\phi(h)\big)h\Big)g(x)-f(x)\Big(g(x)+\big(g'(x)+\psi(h)\big)h\Big)}{hg(x)g(x+h)}\\
=&\frac{\big(f'(x)+\phi(h)\big)g(x)-f(x)\big(g'(x)+\psi(h)\big)}{g(x)g(x+h)}\,.
\end{aligned}$$

Now, formula (4) follows from letting $h \to 0$, using that

$$\lim_{h\to 0}\phi(h)=\lim_{h\to 0}\psi(h)=0\,,$$

and also the fact that differentiability implies continuity, i.e.

$$\lim_{h\to 0} g(x+h)=g(x)\,.$$

□

Theorem 4.5. *(The Chain Rule) Let $f:[a,b]\to[c,d]$ and $g:[c,d]\to\mathbb{R}$. If f is differentiable at $x\in[a,b]$ and g is differentiable at $f(x)$, then $g\circ f$ is differentiable at x and*

$$(g\circ f)'(x)=g'(f(x))\cdot f'(x)\,.$$

Proof. Let us use the notation $y=f(x)$. By Lemma 4.2 we have

$$f(x+h)=f(x)+\big(f'(x)+\phi(h)\big)h$$

and

$$g(y+t) = g(y) + \big(g'(y) + \psi(t)\big)t\,.$$

Therefore,

$$\begin{aligned} g \circ f(x+h) &= g\Big(f(x) + \big(f'(x) + \phi(h)\big)h\Big) \\ &= g(f(x)) + \left(g'(f(x) + \psi\Big(\big(f'(x)+\phi(h)\big)h\Big)\right)\Big(f'(x) + \phi(h)\Big)h\,. \end{aligned}$$

Hence,

$$\begin{aligned} &\frac{g \circ f(x+h) - g(f(x))}{h} \\ =&\left(g'(f(x)) + \psi\Big(\big(f'(x)+\phi(h)\big)h\Big)\right)\Big(f'(x) + \phi(h)\Big)\,. \end{aligned}$$

Letting $h \to 0$, and using the properties of ϕ and ψ, we obtain the chain rule for differentiation.

□

Example 4.6. Let

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

If $x \neq 0$, then f is differentiable at x as a composition of two differentiable functions, and

$$f'(x) = \cos\frac{1}{x} \cdot \left(-\frac{1}{x^2}\right)\,.$$

We proved in Example 3.16 that $\lim_{x\to 0} f(x)$ does not exist, hence f is not continuous at $x = 0$ and therefore cannot be differentiable at $x = 0$.

Example 4.7. Let

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

If $x \neq 0$, then f is differentiable at x as a product and composition of differentiable functions and

$$f'(x) = -\frac{1}{x}\cos\frac{1}{x} + \sin\frac{1}{x}\,.$$

From $\lim_{x\to 0} f(x) = 0$ it follows that f is continuous at $x = 0$.
The limit, as $x \to 0$, of the difference quotient

$$\frac{f(x) - f(0)}{x} = \frac{x\sin\frac{1}{x} - 0}{x} = \sin\frac{1}{x}$$

does not exists, which shows that f is not differentiable at 0.

Example 4.8. Let

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0\,. \end{cases}$$

If $x \neq 0$, then f is differentiable at x as a product and composition of differentiable functions and

$$f'(x) = -\cos \frac{1}{x} + 2x \sin \frac{1}{x}\,.$$

From $\lim_{x\to 0} f(x) = 0$ it follows that f is continuous at $x = 0$.
Calculating the limit, as $x \to 0$, of the difference quotient

$$\lim_{x\to 0} \frac{f(x) - f(0)}{x} = \lim_{x\to 0} \frac{x^2 \sin \frac{1}{x} - 0}{x} = \lim_{x\to 0} x \sin \frac{1}{x} = 0$$

implies that

$$f'(x) = \begin{cases} -\cos \frac{1}{x} + 2x \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0\,. \end{cases}$$

Note that f is differentiable on $\mathbb{R}$, but the derivative is not continuous at 0. Moreover, the discontinuity of f' at 0 is of second kind.

Theorem 4.9. *Let $f : [a, b] \to \mathbb{R}$ and $x_0 \in (a, b)$. If f has a local maximum (or minimum) at x_0 and f is differentiable at x_0, then $f'(x_0) = 0$.*

Proof. Suppose that f has local minimum at x_0. Then, there exists $\delta > 0$ such that

$$(x_0 - \delta, x_0 + \delta) \subset [a, b]$$

and

$$f(x) \geq f(x_0)\,,\ \forall\, x \in (x_0 - \delta, x_0 + \delta)\,.$$

Therefore,

$$\lim_{x \nearrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

and

$$\lim_{x \searrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \geq 0\,.$$

The differentiability of f at x_0 implies that the two sided limits have to be equal to $f'(x_0)$ and therefore $f'(x_0) = 0$.

□

Remark 4.10. Note that, if $x_0 = a$ or $x_0 = b$, the Theorem 4.9 does not hold. Look at $f : [0, 1] \to \mathbb{R}$, $f(x) = 2x$. It has a maximum at $x_0 = 1$, but $f'(1) = 2 \neq 0$. Also, see Exercise 4.1.

Theorem 4.11. *Let $f : [a, b] \to \mathbb{R}$ be differentiable on $[a, b]$. Then f' has the Intermediate Value Property.*

Proof. Let $a \leq x_1 < x_2 \leq y$ and $f'(x_1) < c < f'(x_2)$. Consider $g(x) = f(x) - cx$, which is differentiable on $[a, b]$. Then, $g'(x_1) = f'(x_1) - c < 0$ and therefore there exists $\delta_1 > 0$ such that

$$\frac{g(x) - g(x_1)}{x - x_1} < 0\,,\ \forall\, x \in (x_1, x_1 + \delta_1)\,.$$

Hence, there exists $z_1 \in (x_1, x_2)$ such that $g(x_1) > g(z_1)$.
Similarly, there exists $z_2 \in (x_1, x_2)$ such that $g(z_2) < g(x_2)$. The function g is continuous on $[x_1, x_2]$, so it will attains its minimum, which by the above facts cannot be at x_1 or x_2. Hence, there exists $x_3 \in (x_1, x_2)$ such that $g'(x_3) = 0$. However, this implies that $f'(x_3) = c$. $\square$

Corollary 4.12. *If $f : [a, b] \to \mathbb{R}$ is differentiable on $[a, b]$, then f' can have only discontinuities of second kind.*

Exercises

Exercise 4.1. Let $f : [a, b] \to \mathbb{R}$. Prove that, if f has a local maximum (or minimum) at $x_0 = a$ or $x_0 = b$ and f is differentiable at x_0, then

$$f'(x_0)(x - x_0) \leq 0 \quad (\text{ or } f'(x_0)(x - x_0) \geq 0\,)$$

for all $x \in [a, b]$. (These are called variational inequalities).

Exercise 4.2. Let $f : \mathbb{R} \to \mathbb{R}$ and $x_0 \in \mathbb{R}$. Show that if f is continuous at x_0 and $|f|$ is differentiable at x_0, then f is differentiable at x_0.

Exercise 4.3. Suppose that f is differentiable on (a, b) and f' is monotone increasing on (a, b). Prove that f' is continuous on (a, b).

Exercise 4.4. Let $p \in \mathbb{Z}$ and

$$f(x) = \begin{cases} x^p \sin\frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0\,. \end{cases}$$

(a) For what values of p is f continuous on $\mathbb{R}$?

(b) For what values of p is f differentiable on $\mathbb{R}$?

(c) For what values of p is f' continuous on $\mathbb{R}$?

(d) For what values of p is f differentiable twice on $\mathbb{R}$?

4.2 Mean Value Theorems

Theorem 4.13. *(Rolle's Theorem)*
Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists $x_0 \in (a, b)$ such that $f'(x_0) = 0$.

Proof. If f is constant on $[a, b]$, then $f'(x) = 0$ for all $x \in (a, b)$.
If f is not constant on $[a, b]$, then it attains its maximum and minimum of $[a, b]$. From $f(a) = f(b)$ it follows that one of them must occur inside (a, b). Therefore, there exists a local maximum or minimum point $x_0 \in (a, b)$, and therefore $f'(x_0) = 0$.

□

Theorem 4.14. *(Lagrange's Theorem)*
Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $x_0 \in (a, b)$ such that

$$f(b) - f(a) = f'(x_0)(b - a)\,.$$

Proof. Consider

$$h(x) = x\big(f(b) - f(a)\big) - f(x)\big(b - a\big)\,.$$

Then h is continuous on $[a, b]$, differentiable on (a, b) and simple calculations show that $h(a) = h(b)$. Therefore, by Theorem 4.13, there exists $x_0 \in (a, b)$ such that $h'(x_0) = 0$. Noting that

$$h'(x) = \big(f(b) - f(a)\big) - f'(x)\big(b - a\big)\,,$$

it follows that

$$f(b) - f(a) - f'(x_0)(b - a) = 0\,.$$

□

Theorem 4.15. *(Cauchy's Theorem)*
Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $x_0 \in (a, b)$ such that

$$g'(x_0)\Big(f(b) - f(a)\Big) = f'(x_0)\Big(g(b) - g(a)\Big)\,.$$

Proof. Consider

$$h(x) = g(x)\Big(f(b) - f(a)\Big) - f(x)\Big(g(b) - g(a)\Big)\,.$$

The rest of the proof is similar to the proof of Theorem 4.14.

□

We will list some important consequences of the mean value theorems.

Theorem 4.16. *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then,*

(1) *If $f'(x) = 0$ for all $x \in (a, b)$, then $f =$ constant on $[a, b]$.*

(2) *If* $f'(x) \geq 0$ *for all* $x \in (a,b)$*, then* f *is monotonically increasing on* $[a,b]$*.*

(3) *If* $f'(x) \leq 0$ *for all* $x \in (a,b)$*, then* f *is monotonically decreasing on* $[a,b]$*.*

Proof. (1) Consider any $a < x_1 \leq b$. The function f is continuous on $[a, x_1]$ and is differentiable on (a, x_1), so we can apply Theorem 4.14 to obtain $x_0 \in (a, x_1)$ such that

$$f(x_1) - f(a) = f'(x_0)(x_1 - a) = 0\,.$$

In conclusion, $f(x_1) = f(a)$ for all $a < x_1 \leq b$, which shows that f is a constant function on $[a,b]$.

(2) Consider any $a \leq x_1 < x_2 \leq b$. Apply Theorem 4.14 for f on the interval $[x_1, x_2]$ and find $x_1 < x_0 < x_2$ such that

$$f(x_2) - f(x_1) = f'(x_0)(x_2 - x_1) \geq 0\,.$$

Therefore, we proved that for all $a \leq x_1 < x_2 \leq b$ we have $f(x_1) \leq f(x_2)$, which shows that f is monotonically increasing on $[a,b]$.

(3) The proof is similar to (2).

□

Although, we state and prove the following theorems for an interval on the right side of $x = a$, similar assumptions can be used for the left side or for a symmetric interval around it.

Theorem 4.17. *(Taylor's Theorem)*
Let $f : [a,b] \to \mathbb{R}$ *be a function and* $n \in \mathbb{N}$*. Assume that* $f, f', \ldots, f^{(n)}$ *are continuous on* $[a,b]$ *and* $f^{(n+1)}$ *exists on* (a,b)*. Then for any* $a < x \leq b$ *there exists* $a < c_x < x$ *such that*

$$f(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(c_x)}{(n+1)!}(x-a)^{n+1}\,.$$

Proof. Let us use the notation

$$T(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n\,,$$

which is called the Taylor polynomial of degree n.

Fix any $a < x \leq b$ and let $M \in \mathbb{R}$ be such that

$$\frac{f(x) - T(x)}{(x-a)^{n+1}} = M\,.$$

Define the function $g : [a,b] \to \mathbb{R}$ as,

$$g(t) = f(t) - T(t) - M(t-a)^{n+1}\,.$$

By the choice of M we have

$$g(a) = g(x) = 0\,,$$

so by Theorem 4.13, we obtain $a < x_1 < x$ such that

$$g'(x_1) = 0\,.$$

Now we have that $g'(a) = g'(x_1) = 0$, so we apply again Theorem 4.13 for g', to get $a < x_2 < x_1$ such that

$$g''(x_2) = 0\,.$$

Continuing in this way, after $n+1$ steps we get $a < x_{n+1} < x_n$ such that

$$g^{(n+1)}(x_{n+1}) = 0\,.$$

Observe that

$$g^{(n+1)}(t) = f^{(n+1)}(t) - (n+1)!\,M\,,$$

which shows that

$$M = \frac{f^{(n+1)}(x_{n+1})}{(n+1)!}\,.$$

Therefore, we can choose $c_x = x_{n+1}$ and this finishes the proof.

$\square$

Corollary 4.18. *Under the assumptions of Theorem 4.17, for all $x \in [a,b]$ we have*

$$|f(x) - T(x)| \leq \frac{(x-a)^{n+1}}{(n+1)!} \sup\left\{|f^{(n+1)}(t)| \,:\, t \in (a,b)\right\}\,.$$

Theorem 4.19. *(L'Hôpital's Theorem) Let $-\infty \leq a < b \leq +\infty$ and $f, g : (a,b) \to \mathbb{R}$. Assume that:*

(1) *f and g are differentiable on (a,b).*

(2) $\displaystyle\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = \alpha \in \overline{\mathbb{R}}$.

(3) $\displaystyle\lim_{x \searrow a} f(x) = 0$ *and* $\displaystyle\lim_{x \searrow a} g(x) = 0$
or $\displaystyle\lim_{x \searrow a} g(x) = +\infty$.

Then,

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \alpha\,.$$

Proof. We will show the proof only in the case, when $a \in \mathbb{R}$, $\alpha \in \mathbb{R}$, $\lim_{x \searrow a} f(x) = 0$ and $\lim_{x \searrow a} g(x) = 0$. We leave the proof of the other cases to the reader.

Let $\varepsilon > 0$. Using (2), we find $\delta = \delta(\varepsilon) > 0$ such that

$$\alpha - \varepsilon < \frac{f'(x)}{g'(x)} < \alpha + \varepsilon\,,\ \forall\, a < x < a + \delta\,.$$

Let $a < x_1 < x < a + \delta$ be arbitrary. By Theorem 4.15, there exists $x_1 < t < x$ such that

$$\frac{f(x) - f(x_1)}{g(x) - g(x_1)} = \frac{f'(t)}{g'(t)}\,.$$

Therefore,

$$\alpha - \varepsilon < \frac{f(x) - f(x_1)}{g(x) - g(x_1)} < \alpha + \varepsilon\,,\ \forall\, a < x_1 < x < a + \delta\,.$$

Let $x_1 \to a$ and get

$$\alpha - \varepsilon < \frac{f(x)}{g(x)} < \alpha + \varepsilon\,,\ \forall\, a < x < a + \delta\,.$$

Hence,

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \alpha\,.$$

□

Remark 4.20. We have to be careful when we apply L'Hôpital's Theorem. Let's have a look at two examples. First, consider

$$\lim_{x \to 0} \frac{\sin(x) - x}{x^3}\,.$$

It is very common to see the following solution:
We realize that it is a $\frac{0}{0}$ case, so we apply L'Hôpital's Theorem and get

$$\lim_{x \to 0} \frac{\sin(x) - x}{x^3} = \lim_{x \to 0} \frac{\cos x - 1}{3x^2}\,,$$

which is still in the form of $\frac{0}{0}$, so we want to apply L'Hôpital's Theorem again. But, let's stop for a moment. Look at the equality above. Did L'Hôpital's Theorem say that the equality holds? No, because the assumption (2) is not checked, yet.

How can we solve this problem correctly? Start with

$$\lim_{x \to 0} \frac{\sin(x) - x}{x^3}\,,$$

and realize that it is in a $\frac{0}{0}$ indeterminate form. Look at the limit of $\frac{f'}{g'}$:

$$\lim_{x \to 0} \frac{\cos x - 1}{3x^2}\,,$$

and realize that it is still in a $\frac{0}{0}$ indeterminate form. Take another derivative and evaluate

$$\lim_{x\to 0} \frac{\sin x}{6x},$$

which is equal to $\frac{1}{6}$. Now we can say that based on L'Hôpital's Theorem we have

$$\frac{1}{6} = \lim_{x\to 0} \frac{\sin x}{6x} = \lim_{x\to 0} \frac{\cos x - 1}{3x^2},$$

and therefore

$$\frac{1}{6} = \lim_{x\to 0} \frac{\cos x - 1}{3x^2} = \lim_{x\to 0} \frac{\sin(x) - x}{x^3}.$$

Second, consider the limit

$$\lim_{x\to 0} \frac{x^2 \sin \frac{1}{x}}{\sin x}.$$

Indeed, it is an an indeterminate form $\frac{0}{0}$, so let's try applying L'Hôpital's Theorem. The limit of $\frac{f'}{g'}$ looks like

$$\lim_{x\to 0} \frac{2x \sin \frac{1}{x} - \cos \frac{1}{x}}{\cos x},$$

which does not exist. Can we say that the original limit does not exist? No. Actually,

$$\lim_{x\to 0} \frac{x^2 \sin \frac{1}{x}}{\sin x} = \lim_{x\to 0} \frac{x}{\sin x} \cdot x \sin \frac{1}{x} = 1 \cdot 0 = 0.$$

Exercises

Exercise 4.5. Let $f : (a, b) \to \mathbb{R}$ be differentiable on (a, b) and assume that there exists $0 \leq M < +\infty$ such that $|f'(x)| \leq M$ for all $x \in (a, b)$.

(1) Show that f is uniformly continuous on (a, b).

(2) Give an example of a function which has unbounded derivative, but still is uniformly continuous.

Exercise 4.6. Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b).

Show that if $f'(x) \neq 0$ for all $x \in (a, b)$, then f is strictly monotone increasing or decreasing on $[a, b]$.

Is it true that if f strictly monotone increasing then $f'(x) \neq 0$ for all $x \in (a, b)$?

(Hint: Strictly increasing means that $x_1 < x_2$ implies $f(x_1) < f(x_2)$.)

Exercise 4.7. Let $f : (0, +\infty) \to \mathbb{R}$ be differentiable on $(0, +\infty)$. Prove that, if $\lim_{x\to+\infty} f(x) = M \in \mathbb{R}$, then for all $\varepsilon > 0$ there exists $x(\varepsilon) > 0$ such that $|f'(x(\varepsilon))| < \varepsilon$.

Exercise 4.8. Let $f : [0, 1] \to [0, 1]$ be continuous on $[0, 1]$ and differentiable on $(0, 1)$. Show that if $f'(x) \neq 1$ for all $x \in (0, 1)$, then there exists a unique $x_0 \in [0, 1]$ such that $f(x_0) = x_0$.

Exercise 4.9. Let $f : (0, 1) \to \mathbb{R}$ be differentiable on $(0, 1)$. Assume that $|f'(x)| < \pi$ for all $x \in (0, 1)$. Show that $\lim_{n\to\infty} f(\frac{1}{n})$ exists.

Exercise 4.10. Let $f : (0, +\infty) \to \mathbb{R}$ be differentiable on $(0, +\infty)$.
Assume that $f'(x) \geq c > 0$ for all $x > 0$. Prove that $\lim_{x\to+\infty} f(x) = +\infty$.

Exercise 4.11. Suppose that f is differentiable on $[a, b]$ and f' is continuous on $[a, b]$. Prove that f is absolutely continuous on $[a, b]$.

Exercise 4.12. Let $f : [a, b] \to \mathbb{R}$ be differentiable on $[a, b]$. Assume that $f(a) = 0$ and that, for some $M > 0$, we have $|f'(x)| \leq M|f(x)|$ for all $x \in [a, b]$. Prove that $f \equiv 0$ on $[a, b]$.

Exercise 4.13. Let $f : [a, b] \to [c, d]$ be one-to-one and onto. Also, assume that f is continuous on $[a, b]$ and differentiable on (a, b).

(1) Show that f is strictly monotone.

(2) Show that f^{-1} is continuous on $[c, d]$.

(3) Show that f^{-1} is differentiable at every $x \in (a, b)$ where $f'(x) \neq 0$ and

$$\left(f^{-1}\right)' (f(x)) = \frac{1}{f'(x)} .$$

(4) At how many points can f^{-1} fail to be differentiable?

Exercise 4.14. Let $f, g : [a, +\infty) \to \mathbb{R}$ be continuous on $[a, +\infty)$ and differentiable on $(a, +\infty)$. Assume that $f(a) = g(a)$ and $f'(x) \leq g'(x)$ for all $a < x$.
Prove that:

(1) $f(x) \leq g(x)$ for all $a \leq x$.

(2) $1 + \ln x \leq x$ for all $1 \leq x$.

(3) $0 \leq \sin x \leq x$ for all $0 \leq x \leq 1$.

(4) $1 - \frac{x^2}{2} \leq \cos x \leq 1$ for all $0 \leq x \leq 1$.

Exercise 4.15. Let $f : [a, b] \to \mathbb{R}$ and $a < c < b$. Assume that f is continuous on $[a, b]$ and differentiable on $(a, b) \setminus \{c\}$. Show that if $\lim_{x\to c} f'(x) = \gamma \in \mathbb{R}$, then f is differentiable at c and $f'(c) = \gamma$.

Exercise 4.16. Let

$$x_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln n\,,\ \forall\ n \in \mathbb{N}\,.$$

Show that $\{x_n\}$ is convergent. What can you say about its limit?
Hint: Use Lagrange's theorem.

4.3 Differentiability of monotone functions

Definition 4.21. Let $f : [a, b] \to \mathbb{R}$ and $x_0 \in [a, b]$. We define:

$$\lambda_L(f, x_0) = \liminf_{x \nearrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \searrow 0} \inf \left\{ \frac{f(x) - f(x_0)}{x - x_0} \mid x \in (x_0 - h, x_0) \cap [a, b] \right\}.$$

$$\Lambda_L(f, x_0) = \limsup_{x \nearrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \searrow 0} \sup \left\{ \frac{f(x) - f(x_0)}{x - x_0} \mid x \in (x_0 - h, x_0) \cap [a, b] \right\}.$$

$$\lambda_R(f, x_0) = \liminf_{x \searrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \searrow 0} \inf \left\{ \frac{f(x) - f(x_0)}{x - x_0} \mid x \in (x_0, x_0 + h) \cap [a, b] \right\}.$$

$$\Lambda_R(f, x_0) = \limsup_{x \searrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \searrow 0} \sup \left\{ \frac{f(x) - f(x_0)}{x - x_0} \mid x \in (x_0, x_0 + h) \cap [a, b] \right\}.$$

Note that, the infimum increases as h decreases to 0, so $\liminf$ means the supremum of infimum. Similarly, $\limsup$ means the infimum of supremum.

We always have $\lambda_L(f, x) \le \Lambda_L(f, x)$ and $\lambda_R(f, x) \le \Lambda_R(f, x)$.

Moreover, f is differentiable at x_0 if $\lambda_L(f, x_0) = \Lambda_L(f, x_0) = \lambda_R(f, x_0) = \Lambda_R(f, x_0)$.

Definition 4.22. Let $f : [a, b] \to \mathbb{R}$ be a continuous function and $a \le x_0 \le b$. We say that x_0 is dominated from right by f, if there exists $x_0 < x_1 \le b$ such that $f(x_0) < f(x_1)$. Similarly, we say that x_0 is dominated from left by f if there exists $a \le x_1 < x_0$ such that $f(x_1) > f(x_0)$.

Lemma 4.23. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then, the set of all points dominated from right by f is the union of a pairwise disjoint, at most countable, collection of open intervals (relative to $[a, b]$), (a_i, b_i) or $[a, b_i)$, such that $f(a_i) \le f(b_i)$.*

Similarly, the set of all points dominated from left by f is the union of a pairwise disjoint, at most countable, collection of open intervals (relative to $[a, b]$), (a_i, b_i) or $(a, b_i]$, such that $f(a_i) \ge f(b_i)$.

Proof. We will consider the case of points dominated from right by f. The continuity of f implies that if x_0 is dominated from right by f, then there exists an open interval around x_0 with points dominated from right by f. Hence, this set is open relative to $[a, b]$ and therefore is the union of a collection of pairwise disjoint, at most countable collection of open intervals.

Let us assume the for one of these intervals we have $f(a_i) > f(b_i)$. Then, by the Intermediate Value Theorem, there exists $a_i < x_i < b_i$ such that

$$f(x_i) = \frac{1}{2}(f(a_i) + f(b_i)) > f(b_i).$$

The set

$$K = \{x \in [a_i, b_i] \mid f(x) = \frac{1}{2}(f(a_i) + f(b_i))\}$$

is compact, so there exists $x_0 = \max K$. Note that, $b_i \notin K$, so $a_i < x_i \le x_0 < b_i$. Moreover, x_0 is dominated from right by f and therefore there exists $x > x_0$ such that $f(x_0) < f(x)$. If $x < b_i$, then the Intermediate Value Theorem gives a point in K larger than x_0, which is a contradiction. Also, $x \neq b_i$, because $f(x) > f(x_0) > f(b_i)$. Hence, $b_i < x$, but this shows that b_i is dominated from right, which is a contradiction, because b_i is not part of set of points dominated from right by f.

□

Lemma 4.24. *Let $f : [a, b] \to \mathbb{R}$ be a monotone increasing function, $0 < c < C < \infty$ and*

$$A(c, C) = \left\{ x \in [a, b] \,\middle|\, \lambda_L(f, x) < c < C < \Lambda_R(f, x) \right\} .$$

Then, for all $(\alpha, \beta) \subset [a, b]$ we have

$$\mu^*\Big(A(c, C) \cap (\alpha, \beta)\Big) \le \frac{c}{C}(\beta - \alpha) . \tag{4.1}$$

Proof. Let us fix an interval $(\alpha, \beta) \subset [a, b]$. If $A(c, C) \cap (\alpha, \beta) = \emptyset$, then (4.1) is trivial.

If $A(c, C) \cap (\alpha, \beta) \neq \emptyset$, then let $x_0 \in A(c, C) \cap (\alpha, \beta)$ be an arbitrary point. From $\lambda_L < c$ it follows that there exists $x < x_0$ such that

$$\frac{f(x) - f(x_0)}{x - x_0} < c ,$$

which implies that

$$f(x) - cx > f(x_0) - cx_0 .$$

This means that x_0 is dominated from left by $f(x) - cx$ and therefore

$$A(c, C) \cap (\alpha, \beta) \subset \bigcup_i (\alpha_i, \beta_i) ,$$

such that

$$f(\beta_i) - f(\alpha_i) \le c(\beta_i - \alpha_i) ,$$

as it is described by Lemma 4.23.

If for some index i, there is point $x_i \in A(c, C) \cap (\alpha_i, \beta_i)$, then, by similar arguments as above, x_i is dominated from right by $f(x) - Cx$ and therefore

$$A(c, C) \cap (\alpha_i, \beta_i) \subset \bigcup_n (\alpha_{i,n}, \beta_{i,n}) .$$

Hence,

$$A(c, C) \cap (\alpha, \beta) \subset \bigcup_{i,n} (\alpha_{i,n}, \beta_{i,n}) ,$$

and

$$f(\beta_{i,n}) - f(\alpha_{i,n}) \ge C(\beta_{i,n} - \alpha_{i,n}) , \; \forall\, i, n .$$

Therefore,

$$\sum_{i,n}(\beta_{i,n}-\alpha_{i,n}) \le \frac{1}{C}\sum_{i,n}\left(f(\beta_{i,n})-f(\alpha_{i,n})\right) \le \frac{1}{C}\sum_{i}\left(f(\beta_i)-f(\alpha_i)\right)$$
$$\le \frac{c}{C}\sum_{i}(\beta_i-\alpha_i) \le \frac{c}{C}(\beta-\alpha)\,.$$

This last estimate clearly implies (4.1).

□

Lemma 4.25. *Let* $f:[a,b]\to\mathbb{R}$ *be a monotone increasing function,* $0<c<C<\infty$ *and*

$$B(c,C)=\left\{x\in[a,b]\;\middle|\;\lambda_R(f,x)<c<C<\Lambda_L(f,x)\right\}\,.$$

Then, for all $(\alpha,\beta)\subset[a,b]$ *we have*

$$\mu^*\Big(B(c,C)\cap(\alpha,\beta)\Big)\le\frac{c}{C}(\beta-\alpha)\,. \tag{4.2}$$

Proof. Let $f^*:[-b,-a]\to\mathbb{R}$ defined by $f^*(x)=-f(-x)$. Observe that for any $x\in[a,b]$ we have

$$\lambda_R(f,x)=\lambda_L(f^*,-x)\ \text{ and }\ \Lambda_L(f,x)=\Lambda_R(f^*,-x)\,.$$

We apply now Lemma 4.24 to f^*, from which (4.2) follows.

□

Theorem 4.26. *Let* $f:[a,b]\to\mathbb{R}$ *be a monotone increasing (or decreasing) function. Then* f *is differentiable almost everywhere.*

Proof. Let us assume that f is monotone increasing. Also, we can assume that f is continuous, because it can have at most countable many discontinuities, which can be removed by subtracting Heaviside type jump functions (see the section about Riemann-Stieltjes integrals).

Let us define the following sets:

$$Z_1=\left\{x\in[a,b]\;\middle|\;\Lambda_R(f,x)=+\infty\right\},$$
$$Z_2=\left\{x\in[a,b]\;\middle|\;\lambda_L(f,x)<\Lambda_R(f,x)\right\},$$
$$Z_3=\left\{x\in[a,b]\;\middle|\;\lambda_R(f,x)<\Lambda_L(f,x)\right\}.$$

If we show that all these sets have measure zero, then

$$Z=Z_1\cup Z_2\cup Z_3$$

has also measure zero. If $x\in[a,b]\setminus Z$, then

$$\Lambda_R(f,x)\le\lambda_L(f,x)\le\Lambda_L(f,x)\le\lambda_R(f,x)\le\Lambda_R(f,x)\ne+\infty\,,$$

which means that f is differentiable at x.

Let $c > 0$ be an arbitrary number. If $x_0 \in Z_1$, then there exists $x > x_0$ such that

$$\frac{f(x) - f(x_0)}{x - x_0} > c\,.$$

Hence,

$$f(x) - cx > f(x_0) - cx_0\,,$$

which means that x_0 is dominated from right by $f(x) - cx$. Therefore, by Lemma 4.23 Z_1 is the union of a pairwise disjoint, at most countable, collection of open intervals (relative to $[a, b]$), (a_i, b_i) or $[a, b_i)$, such that

$$f(b_i) - f(a_i) \geq C(b_i - a_i)\,.$$

Then,

$$\sum_i (b_i - a_i) \leq \sum_i \frac{f(b_i) - f(a_i)}{c} \leq \frac{f(b) - f(a)}{c}\,.$$

But, $c > 0$ can be chosen arbitrarily large, so $\mu(Z_1) = 0$.

Using the sets $A(c, C)$ from Lemma 4.24 we can write Z_2 as a countable union

$$Z_2 = \bigcup \left\{ A(c, C) \;\middle|\; c < C\,;\; c\,, C \in \mathbb{Q} \right\}.$$

Therefore, we have to show that $\mu(A(c, C)) = 0$. Let us use the notation

$$\mu^*(A(c, C)) = \nu\,.$$

Then, for any $\varepsilon > 0$ there is a, at most countable, collection of open intervals (a_i, b_i) such that

$$\sum_i (b_i - a_i) < \nu + \varepsilon\,.$$

Then by Lemma 4.24 we have

$$\nu \leq \frac{c}{C} \sum (b_i - a_i) < \frac{c}{C}(\nu + \varepsilon)\,.$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$\nu \leq \frac{c}{C}\nu\,,$$

which implies that $\nu = 0$.

Therefore, $\nu(Z_2) = 0$ and similarly $\nu(Z_3) = 0$.

$\square$

Exercises

Exercise 4.17. Let $\{r_n\}$ is an enumeration of the rational numbers in $[0,1]$. Define the function $f : [0,1] \to \mathbb{R}$ by

$$f(x) = \sum_{r_n < x} \frac{1}{2^n}, \ \forall\, x \in [0,1].$$

(1) Prove that f is monotone increasing.

(2) Prove that f is continuous at every irrational number and discontinuous at every rational number.

(3) What can we say about the differentiability of f?

Exercise 4.18. Consider the Cantor function defined in Example 3.45.

(1) Describe the differentiability property of the Cantor function.

(2) If a continuous function has derivative equal to zero almost everywhere, can we conclude that the function is constant?

Chapter 5

Integration

In this chapter we will talk about several notions of integrals.

5.1 The Riemann integral

The Riemann integral is probably the simplest and most commonly used. This integral is part of the prerequisite knowledge, so we will skip the proofs of the elementary properties.

Definition 5.1. Consider a closed and bounded interval $[a,b]$. A partition P of $[a,b]$ is a set of points $P = \{x_0, x_1, \dots, x_n\}$ such that $a = x_0 < x_1 < \cdots < x_n = b$.
If we have two partitions P and Q of the same interval $[a,b]$, we say that Q is a refinement of P, if $P \subset Q$.

Definition 5.2. Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $P = \{x_0, x_1, \dots, x_n\}$ be a partition of $[a,b]$. We define the following quantities:

$$m_i(f) = \inf\{f(x) \mid x_{i-1} \le x \le x_i\}\,,$$

$$M_i(f) = \sup\{f(x) \mid x_{i-1} \le x \le x_i\}\,,$$

$$\Delta x_i = x_i - x_{i-1}\,.$$

The norm of the partition P is defined as

$$||P|| = \max\{\Delta x_i \mid 1 \le i \le n\}\,.$$

The lower Riemann sum of f associated to the partition P is

$$\underline{S}(f,P) = \sum_{i=1}^{n} m_i(f)\Delta x_i\,.$$

The upper Riemann sum of f associated to the partition P is

$$\overline{S}(f,P) = \sum_{i=1}^{n} M_i(f)\Delta x_i\,.$$

The lower Riemman sum of f is

$$\underline{S}(f) = \sup\{\underline{S}(f,P) \mid P \text{ is any partition of } [a,b]\}.$$

The upper Riemman sum of f is

$$\overline{S}(f) = \inf\{\overline{S}(f,P) \mid P \text{ is any partition of } [a,b]\}.$$

We say that f is Riemann integrable on $[a,b]$ if

$$\underline{S}(f) = \overline{S}(f)\,.$$

If f is Riemann integrable, then the common value of the upper and lower Riemann sum is denoted by

$$\int_a^b f(x)dx\,,$$

and it is called the Riemann integral of f.
Also, we define

$$\int_a^a f(x)\,dx = 0\,.$$

and

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx\,.$$

Example 5.3. Let $f : [a,b] \to \mathbb{R}$, $f(x) = c$ for all $x \in [a,b]$. For any partition P of $[a,b]$ we have $m_i(f) = M_i(f) = c$ and

$$\underline{S}(f,P) = \overline{S}(f,P) = c(b-a).$$

Therefore,

$$\underline{S}(f) = \overline{S}(f) = c(b-a)\,,$$

so the constant function is Riemann integrable and hence

$$\int_a^b c\,dx = c(b-a)\,.$$

Example 5.4. Let $f : [a, b] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in [a, b] \cap \mathbb{Q} \\ 0 & \text{if } x \in [a, b] \cap \mathbb{I} \,. \end{cases}$$

For any partition P of $[a, b]$, the density of rationals and irrationals shows that $m_i(f) = 0$ and $M_i(f) = 1$, so

$$\underline{S}(f, P) = 0\,, \quad \text{and} \quad \overline{S}(f, P) = b - a\,.$$

Therefore,

$$0 = \underline{S}(f) < \overline{S}(f) = 1\,,$$

and hence the function f is not Riemann integrable.

To prove the properties of Riemann integrals, we need the following lemma. We leave its proof as an exercise.

Lemma 5.5. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then the following statements hold:*

(1) *If P_1, P_2 are partitions of $[a, b]$, $P_1 \subset P_2$, then*

$$\underline{S}(f, P_1) \le \underline{S}(f, P_2) \ \ \textit{and} \ \ \overline{S}(f, P_1) \ge \overline{S}(f, P_2)\,.$$

(2) *If P and Q are any two partitions of $[a, b]$, then*

$$\underline{S}(f, P) \le \overline{S}(f, Q)\,.$$

(3)

$$\underline{S}(f) \le \overline{S}(f)\,.$$

(4) *$\underline{S}(f) = \overline{S}(f)$ if and only if for all $\varepsilon > 0$ there exists a partition P_ε such that*

$$\overline{S}(f, P_\varepsilon) - \underline{S}(f, P_\varepsilon) < \varepsilon\,.$$

Theorem 5.6. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f is Riemann integrable on $[a, b]$.*

Proof. The function f is a continuous function defined on a compact interval, so f is bounded and uniformly continuous.

Let $\varepsilon > 0$. By the uniform continuity of f, there exists $\delta = \delta(\varepsilon) > 0$ such that for all $x, y \in [a, b]$ with $|x - y| < \delta$ we have $|f(x) - f(y)| < \frac{\varepsilon}{2(b-a)}$.

Let $P_\varepsilon = \{x_0, x_1, \dots, x_n\}$ be a partition of $[a,b]$ such that $||P_\varepsilon|| < \delta(\varepsilon)$. Then for all $1 \le i \le n$ and $x_{i-1} \le x < y \le x_i$ we have

$$|f(x) - f(y)| < \frac{\varepsilon}{2(b-a)}\,.$$

Hence,

$$M_i(f) - m_i(f) \le \frac{\varepsilon}{2(b-a)}\,,\ \forall\, 1 \le i \le n\,.$$

Therefore,

$$\begin{aligned}\overline{S}(f,P_\varepsilon) - \underline{S}(f,P_\varepsilon) &= \sum_{i=1}^{n}(M_i(f) - m_i(f))\Delta x_i \le \frac{\varepsilon}{2(b-a)}\sum_{i=1}^{n}\Delta x_i \\ &= \frac{\varepsilon}{2(b-a)}(b-a) = \frac{\varepsilon}{2} < \varepsilon\,.\end{aligned}$$

Hence, by Lemma 5.5 (4), it follows that f is Riemann integrable.

□

For the main theorem regarding Riemann integrals, we need the following definition and lemma.

Definition 5.7. Let $f : [a,b] \to \mathbb{R}$ and $x \in [a,b]$. For any $h > 0$ we define the oscillation of f on the interval $(x-h, x+h)$ as

$$\operatorname{osc}(f)(x-h,x+h) = \sup\Big\{|f(x_1) - f(x_2)|\ \Big|\ x_1, x_2 \in (x-h,x+h) \cap [a,b]\Big\}\,.$$

We define the oscillation of f at x as

$$\operatorname{osc}(f)(x) = \lim_{h \searrow 0}\ \operatorname{osc}(f)(x-h,x+h)\,.$$

Notice that, if $0 < h_1 < h_2$, then

$$\operatorname{osc}(f)(x-h_1,x+h_1) \le \operatorname{osc}(f)(x-h_2,x+h_2)\,,$$

so we can define $\operatorname{osc}(f)(x)$ also as

$$\operatorname{osc}(f)(x) = \inf_{h>0}\ \operatorname{osc}(f)(x-h,x+h)\,.$$

Example 5.8. Let $f : [-1,1] \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} 1 & \text{if } -1 \le x < 0 \\ 3 & \text{if } 0 \le x \le 1\,. \end{cases}$$

In this case, $\operatorname{osc}(f)(0) = 2$ and $\operatorname{osc}(f)(x) = 0$ if $x \neq 0$.

Lemma 5.9. *Let $f : [a,b] \to \mathbb{R}$ and $x \in [a,b]$.*
Then f is continuous at x if and only if $\mathrm{osc}(f)(x) = 0$.

Proof. First, suppose that f is continuous at x. Then, for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that $|f(x) - f(y)| < \frac{\varepsilon}{2}$ for all $y \in (x-\delta, x+\delta) \cap [a,b]$.
Therefore, for all $x_1, x_2 \in (x-\delta, x+\delta) \cap [a,b]$ we have

$$|f(x_1) - f(x_2)| \le |f(x_1) - f(x)| + |f(x) - f(x_2)| < \varepsilon\,.$$

Hence,

$$\mathrm{osc}(f)(x-\delta, x+\delta) \le \varepsilon\,.$$

Note that, as we mentioned eralier, if $0 < h < \delta$, then

$$\mathrm{osc}(f)(x-h, x+h) \le \mathrm{osc}(f)(x-\delta, x+\delta)\,.$$

In conclusion, we showed that for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all $0 < h < \delta$ we have

$$\mathrm{osc}(f)(x-h, x+h) \le \varepsilon\,,$$

which implies that $\mathrm{osc}(f)(x) = 0$.

Suppose now that $\mathrm{osc}(f)(x) = 0$. Then, for any $\varepsilon > 0$ there exists $H = H(\varepsilon) > 0$ such that for all $0 < h < H$ we have

$$\mathrm{osc}(f)(x-h, x+h) < \varepsilon\,.$$

By considering $x_2 = x$ in Definition 5.7 and fixing $h > 0$ with $0 < h < H$, we get

$$|f(x_1) - f(x)| < \varepsilon,\,,\ \forall\, x_1 \in (x-h, x+h) \cap [a,b]\,,$$

which shows that f is continuous at x.

$\square$

Lemma 5.10. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $\gamma > 0$. Then the set*

$$D_\gamma = \left\{ x \in [a,b] \,\middle|\, \mathrm{osc}(f)(x) \ge \gamma \right\}$$

is compact.

Proof. As D_γ is a subset of the compact set $[a,b]$, it is enough to show that it is closed, which is equivalent to saying that its complement is open in $[a,b]$. So, consider

$$D_\gamma^c = \left\{ x \in [a,b] \,\middle|\, \mathrm{osc}(f)(x) < \gamma \right\}\,.$$

Let $x \in D_\gamma^c$. Then, there exists $\varepsilon > 0$ such that

$$\mathrm{osc}(f)(x) < \gamma - \varepsilon\,,$$

and by the properties of infimums, there exists $h_\varepsilon > 0$ such that

$$\operatorname{osc}(f)(x - h_\varepsilon, x + h_\varepsilon) < \gamma\,.$$

Therefore, for any $z \in (x - \frac{h_\varepsilon}{2}, x + \frac{h_\varepsilon}{2}) \cap [a, b]$ we have

$$(z - \frac{h_\varepsilon}{2}, z + \frac{h_\varepsilon}{2}) \subset (x - h_\varepsilon, x + h_\varepsilon)\,,$$

and hence

$$\operatorname{osc}(f)(z - \frac{h_\varepsilon}{2}, z + \frac{h_\varepsilon}{2}) < \gamma\,.$$

It follows that $\operatorname{osc}(f)(z) < \gamma$ and $z \in D_\gamma^c$, which implies that D_γ^c is open in $[a, b]$.

□

Theorem 5.11. *(The Riemann-Lebesgue Theorem)*
Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then the following two statements are equivalent:

(1) *f is Riemann integrable on $[a, b]$.*

(2) *The set of discontinuities of f has measure zero.*

Proof. First, suppose that f is Riemann integrable on $[a, b]$. By Lemma 5.9 we can write the set of discontinuity points of f as:

$$D = \left\{ x \in [a, b] \,\middle|\, \operatorname{osc}(f)(x) > 0 \right\}\,.$$

We continue the proof by considering

$$D = \bigcup_{k=1}^{\infty} D_k\,,$$

where

$$D_k = \left\{ x \in [a, b] \,\middle|\, \operatorname{osc}(f)(x) \geq \frac{1}{k} \right\}\,.$$

If we show that the measure of each D_k is zero, then the measure of D is also zero, as a countable union of sets of measure zero.

Fix $k \in \mathbb{N}$ and let $\varepsilon > 0$. Then, by the Riemann integrability of f, there exists a partition $P = P_\varepsilon = \{x_0, x_1, \ldots, x_n\}$ such that

$$\overline{S}(f, P) - \underline{S}(f, P) < \frac{\varepsilon}{2k}\,.$$

If $x \in D_k \setminus P$, then $x \in (x_{i-1}, x_i)$ for some $1 \leq i \leq n$. Therefore, there exists $h > 0$ such that $(x - h, x + h) \subset (x_{i-1}, x_i)$ and hence

$$M_i(f) - m_i(f) \geq \frac{1}{k}\,.$$

We obtained that

$$\frac{\varepsilon}{2k} > \overline{S}(f,P) - \underline{S}(f,P) \geq \sum_{(x_{i-1},x_i)\cap D_k \neq \emptyset} (M_i(f) - m_i(f))\Delta x_i$$
$$\geq \frac{1}{k} \sum_{(x_{i-1},x_i)\cap D_k \neq \emptyset} \Delta x_i \,.$$

Therefore,

$$\sum_{(x_{i-1},x_i)\cap D_k \neq \emptyset} \Delta x_i < \frac{\varepsilon}{2}\,,$$

and by this we covered $D_k \setminus P$ by intervals of total length less then $\frac{\varepsilon}{2}$.

The partition P has finitely many points, so we can cover $D_k \cap P$ by intervals of total length also less then $\frac{\varepsilon}{2}$ and by this we can cover the entire D_k with intervals of total length less then ε. In conclusion, D_k has measure zero and this implies that $\mu(D) = 0$.

Suppose now that f is continuous on $[a,b]$, except a set of measure zero. Let

$$M = \sup\left\{ f(x) \,\middle|\, x \in [a,b] \right\} \text{ and } m = \inf\left\{ f(x) \,\middle|\, x \in [a,b] \right\}\,.$$

We can assume that $M - m > 0$, because otherwise f is a constant function and it is Riemann integrable.

Let us use the same notation D for the set of discontinuities of f. As in the first part of this proof, let

$$D = \left\{ x \in [a,b] \,\middle|\, \operatorname{osc}(f)(x) > 0 \right\}\,,$$

and for any $\gamma > 0$, let

$$D_\gamma = \left\{ x \in [a,b] \,\middle|\, \operatorname{osc}(f)(x) \geq \gamma \right\}\,.$$

From $D_\gamma \subset D$ it follows that $\mu(D_\gamma) = 0$. Moreover, by Lemma 5.10, D_γ is compact.

Let $\varepsilon > 0$. We want to find a partition P of $[a,b]$ such that

$$\overline{S}(f,P) - \underline{S}(f,P) < \varepsilon\,.$$

Using the properties of sets of measure zero and the compactness of the sets D_γ, we can find a finite collection of open intervals $\{I_1, \ldots, I_p\}$ such that for $\gamma_\varepsilon = \frac{\varepsilon}{2(b-a)}$ we have

$$D_{\gamma_\varepsilon} \subset \bigcup_{k=1}^{p} I_k$$

and

$$\sum_{k=1}^{p} \mu(I_k) < \frac{\varepsilon}{2(M-m)}\,.$$

If $z \in [a,b] \setminus \bigcup_{k=1}^{p} I_k$, then $\operatorname{osc}(f)(z) < \gamma_\varepsilon$, and hence there exists $h_z > 0$ such that $\operatorname{osc}(f)(z-h_z, z+h_z) < \gamma_\varepsilon$. The set $[a,b] \setminus \bigcup_{k=1}^{p} I_k$ is compact, so from its cover by these intervals $(z-h_z, z+h_z)$ we can select a finite subcover:

$$[a,b] \setminus \bigcup_{k=1}^{p} I_k \;\subset\; \bigcup_{j=1}^{m} (z_j - h_j, z_j + h_j)\,.$$

In this way we obtained an open cover of the interval $[a,b]$:

$$[a,b] \subset \bigcup_{k=1}^{p} I_k \;\bigcup\; \bigcup_{j=1}^{m} (z_j - h_j, z_j + h_j)\,.$$

We select the partition $P = \{x_0, x_1, \ldots, x_n\}$ in such a way that each interval $[x_{i-1}, x_i]$ is contained in one of the above intervals covering $[a,b]$. Then,

$$\overline{S}(f,P) - \underline{S}(f,P) = \sum_k (M_k - m_k)\Delta x_k + \sum_j (M_j - m_j)\Delta x_j\,,$$

where $\sum_k$ is the sum over those intervals which are contained in $\bigcup_k I_k$ and the $\sum_j$ contains the remaining ones. Therefore,

$$\sum_k (M_k - m_k)\Delta x_k \le (M-m)\sum_k \Delta x_k < (M-m)\sum_{k=1}^{p} \mu(I_k) < \frac{\varepsilon}{2}$$

and

$$\sum_j (M_j - m_j)\Delta x_j \le \gamma_\varepsilon \sum_j \Delta x_j \le \frac{\varepsilon}{2(b-a)}(b-a) = \frac{\varepsilon}{2}\,.$$

In conclusion,

$$\overline{S}(f,P) - \underline{S}(f,P) < \varepsilon\,,$$

and this proves that f is Riemann integrable.

□

Corollary 5.12. *Let $f : [a,b] \to \mathbb{R}$ be a montone increasing or decreasing function. Then, f is Riemann integrable.*

Proof. A monotone function defined on a closed interval is bounded and, moreover, can have at most countably many discontinuity points, which form a set of measure zero. Hence, f is Riemann integrable.

□

We extend now the Riemann integrability to unbounded functions.

Definition 5.13. Let $f : [a, b) \to \mathbb{R}$ be a function such that f is Riemann integrable on every $[a, t]$ interval, when $a < t < b$ and

$$\lim_{t \nearrow b} f(t) = +\infty \quad \text{or} \quad \lim_{t \nearrow b} f(t) = -\infty \, .$$

We say that f is improper Riemann integrable on $[a, b]$ if

$$\lim_{t \nearrow b} \int_a^t f(x)\, dx$$

exists and is finite. In this case, we use the notation

$$\lim_{t \nearrow b} \int_a^t f(x)\, dx = \int_a^b f(x)\, dx \, .$$

Similarly, if $f : (a, b] \to \mathbb{R}$ and f is Riemann integrable on every interval $[t, b]$, when $a < t < b$ and

$$\lim_{t \searrow a} f(t) = +\infty \quad \text{or} \quad \lim_{t \searrow a} f(t) = -\infty \, ,$$

we say that f is improper Riemann integrable if

$$\lim_{t \searrow a} \int_t^b f(x)\, dx$$

exists and is finite. Again, we use the notation

$$\lim_{t \searrow a} \int_t^a f(x)\, dx = \int_a^b f(x)\, dx \, .$$

If f is improper Riemann integrable on $[a, b]$, we also use the expression that the improper integral is convergent.

Moreover, if there are at most finitely many points $x_1, ..., x_n \in [a, b]$ at which f has unbounded one-sided limits, we say that f is improper Riemann integrable on $[a, b]$ if:

(1) f is bounded in the complement of any open cover of $x_1, ..., x_n$.

(2) We can write $\int_a^b f(x)\, dx$ as a sum of finitely many convergent improper Riemann integrals (as defined above).

Remark 5.14. Note that, when we say that f is Riemann integrable on $[a, b]$, we always assume that f is bounded on $[a, b]$.

Similarly, when we say that f is improper Riemann integrable, we always assume that there are at most finitely many points $x_1, ..., x_n \in [a, b]$ at which f has unbounded limits, and f is bounded in the complement of any open cover of these points.

Corollary 5.15. *If f is improper Riemann integrable on $[a, b]$, then the set of discontinuities of f is of measure zero.*

Proof. Without loss of generality we can assume that the only point where f is unbounded is $x = a$. Then f is Riemann integrable on every interval $[a+\frac{1}{n}, b]$, where $n > \frac{1}{b-a}$. By Theorem 5.11, the set of discontinuity points of f on $[a + \frac{1}{n}, b]$ is of measure zero, and therefore, as a union of countable many sets of measure zero, the set of discontinuities on $[a, b]$ is also a set of measure zero.

□

Definition 5.16. Let $f : [a, +\infty) \to \mathbb{R}$ be a function such that f is Riemann integrable on every $[a, t]$ interval, when $a < t < +\infty$. We say that f is improper Riemann integrable on $[a, +\infty)$ if

$$\lim_{t\to+\infty} \int_a^t f(x)\,dx$$

exists and is finite. In this case, we use the notation

$$\lim_{t\nearrow b} \int_a^t f(x)\,dx = \int_a^{+\infty} f(x)\,dx\,.$$

Similarly, if $f : (-\infty, b] \to \mathbb{R}$ and f is Riemann integrable on every interval $[t, b]$, for $-\infty < t < b$, we say that f is improper Riemann integrable if

$$\lim_{t\to-\infty} \int_t^b f(x)\,dx$$

exists and is finite. Again, we use the notation

$$\lim_{t\searrow a} \int_t^a f(x)\,dx = \int_{-\infty}^b f(x)\,dx\,.$$

The next theorem lists the most elementary properties of the Riemann integral. These properties are usually proven first for Riemann integrable functions and then extended, if possible, to the improper Riemann integrable functions. The details of the proof are left as an exercise.

Theorem 5.17. *Let f and g be Riemann integrable functions on $[a, b]$, $a < c < b$ and $\lambda \in \mathbb{R}$. Then the following properties hold:*

(1) *$f + g$ is Riemann integrable and*

$$\int_a^b f(x) + g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx\,.$$

(2)
$$\int_a^b \lambda f(x)\,dx = \lambda \int_a^b f(x)\,dx\,.$$

(3)
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\,.$$

(4) *If $f(x) \geq 0$ for all $x \in [a,b]$, except maybe finitely many points, then*
$$\int_a^b f(x)\,dx \geq 0\,.$$

(5) *If f is Riemann integrable on $[a,b]$, then $|f|$ is also Riemann integrable on $[a,b]$ and*
$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx\,.$$

Remark 5.18. When trying to extend property (5) to improper Riemann integrals $\int_a^\infty$ or $\int_{-\infty}^b$, we have to allow the right hand side to be $+\infty$, if needed. As an example, look at
$$f(x) = \frac{\sin(x)}{x}\,,\ x \geq 1\,.$$

Theorem 5.19. *(The Mean Value Theorem) Let $f : [a,b] \to \mathbb{R}$ be a Riemann integrable function and assume that $m \leq f(x) \leq M$ for all $x \in [a,b]$. Then:*

(1)
$$m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a)\,.$$

(2) *If f is continuous on $[a,b]$, then there exists $c \in [a,b]$ such that*
$$\int_a^b f(x)\,dx = f(c)(b-a)\,.$$

Proof. (1) Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a,b]$. Then the properties of lower and upper Riemann sums imply that
$$m(b-a) \leq \underline{S}(f,P) \leq \overline{S}(f,P) \leq M(b-a)\,,$$

from which (1) follows.

(2) Let f be continuous on $[a, b]$ and consider

$$m = \min\left\{f(x) \,\middle|\, x \in [a, b]\right\} \quad \text{and} \quad M = \max\left\{f(x) \,\middle|\, x \in [a, b]\right\} .$$

From (1) it follows

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

Any continuous function has the Intermediate Value Property, hence there exists $c \in [a, b]$ such that

$$\frac{1}{b-a} \int_a^b f(x) = f(c) ,$$

from which the Mean Value Formula follows.

$\square$

Theorem 5.20. *Let f be a Riemann or improper Riemann integrable function on $[a, b]$ and define $F : [a, b] \to \mathbb{R}$ as*

$$F(x) = \int_a^x f(t)\, dt .$$

Then:

(1) *F is absolutely continuous on $[a, b]$.*

(2) *If f is continuous at $x \in [a, b]$, then F is differentiable at x and $F'(x) = f(x)$.*

(3) *F is differentiable and $F'(x) = f(x)$ almost everywhere on $[a, b]$.*

Proof. (1) First, let us consider that f is Riemann integrable. Then, f is bounded, and let $M = \sup\left\{\left|f(x)\right| \,\middle|\, x \in [a, b]\right\}$. Then, for every $a \leq x_0 < x \leq b$ we have

$$\left|F(x) - F(x_0)\right| = \left|\int_a^x f(t)\, dt - \int_a^{x_0} f(t)\, dt\right| = \left|\int_{x_0}^x f(t)\, dt\right| \leq M|x - x_0| , \tag{5.1}$$

which shows that f is Lipschitz continuous and therefore, absolutely continuous, too.

Second, consider that f is improper Riemann integrable and x_0 is a point at which

$$\lim_{x \to x_0} f(x) = +\infty .$$

For simplicity, we can consider that x_0 is the only such point in $[a, b]$, otherwise we split $[a, b]$ in a finite number of intervals in which we have only one such point. Also, we can assume that $x_0 = b$.

Let $\varepsilon > 0$. Using the convergence of the improper integral $\int_a^b f(t)\,dt$ we can find $0 < \delta_0 < b-a$ such that for all $x \in [b-\delta_0, b]$ we have $f(x) > 0$ and

$$\left| \int_{b-\delta_0}^{b} f(x)\,dx \right| < \frac{\varepsilon}{2}\,.$$

Let

$$M_0 = \sup\left\{ \left|f(x)\right| \,\middle|\, x \in [a, b-\delta_0] \right\}\,.$$

If $[x, y] \subset [a, b-\delta_0]$ then, similarly to (5.1),

$$|F(x) - F(y)| \le M_0 |x-y|\,.$$

We define

$$\delta = \delta(\varepsilon) = \min\left\{ \delta_0, \frac{\varepsilon}{2(M_0+1)} \right\}\,.$$

Let now $\{(c_1, d_1), \ldots, (c_n, d_n)\}$ be a collection of pairwise disjoint sub-intervals of $[a, b]$, such that

$$\sum_{k=1}^{n} (d_k - c_k) < \delta\,.$$

We can assume that for each interval we have $[c_k, d_k] \subset [b-\delta, b]$ or $[c_k, d_k] \cap (b-\delta, b) = \emptyset$, because otherwise we split them in smaller intervals.

For the intervals with $[c_k, d_k] \subset [b-\delta, b]$ we have

$$\sum_k |F(d_k) - F(c_k)| = \sum_k (F(d_k) - F(c_k)) \le \int_{b-\delta}^{b} f(x)\,dx < \frac{\varepsilon}{2}\,.$$

For the intervals with $[c_j, d_j] \cap (b-\delta, b) = \emptyset$ we have

$$\sum_j |F(d_j) - F(c_j)| \le M_0 \sum_j (d_j - c_j) < M_0 \frac{\varepsilon}{2(M_0+1)} < \frac{\varepsilon}{2}\,.$$

Hence,

$$\sum_{k=1}^{n} |F(d_n) - F(c_n)| < \varepsilon\,,$$

and this proves that F is absolutely continuous.

(2) Consider now that f is continuous at $x_1 \in [a, b]$ and let $\varepsilon > 0$. Then there exists $\delta = \delta(\varepsilon, x_1) > 0$ such that for all $x \in [x_1 - \delta, x_1 + \delta] \cap [a, b]$ we have $|f(x) - f(x_1)| < \varepsilon$.

If $x \in [x_1 - \delta, x_1 + \delta] \cap [a, b]$, then

$$\left| \frac{F(x) - F(x_1)}{x - x_1} - f(x_1) \right| = \left| \frac{1}{x - x_1} \int_{x_1}^{x} f(t)\,dt - \frac{1}{x - x_1} \int_{x_1}^{x} f(x_1)\,dt \right|$$

$$
\begin{aligned}
&\leq \frac{1}{|x-x_1|}\int_{x_1}^{x} |f(t)-f(x_1)|\,dt \\
&\leq \frac{1}{|x-x_1|}\varepsilon|x-x_1| = \varepsilon\,.
\end{aligned}
$$

Therefore,

$$\lim_{x\to x_1}\frac{F(x)-F(x_1)}{x-x_1} = f(x_1)\,,$$

which shows that $F'(x_1) = f(x_1)$.

(3) The function f is Riemann or improper Riemann integrable on $[a,b]$, so the set of discontinuities is of measure zero. Hence, in the complement of this set of measure zero, the result from (2) applies.

$\square$

Example 5.21. Consider $f:[-1,1]\setminus\{0\}\to\mathbb{R}$, given by the expression

$$f(x) = \frac{2}{3\sqrt[3]{x}}\,.$$

Notice that f is improper Riemann integrable on $[-1,1]$ because the improper integrals

$$\int_{-1}^{0}\frac{2}{3\sqrt[3]{x}}\,dx \quad\text{and}\quad \int_{0}^{1}\frac{2}{3\sqrt[3]{x}}\,dx$$

are convergent. Indeed,

$$\int_{-1}^{0} f(t)\,dt = \lim_{x\nearrow 0}\int_{-1}^{x}\frac{2}{3\sqrt[3]{t}}\,dt = \lim_{x\nearrow 0}\sqrt[3]{t^2}\Big|_{-1}^{x} = \lim_{x\nearrow 0}\sqrt[3]{x^2}-1 = -1\,.$$

Similarly,

$$\int_{0}^{1} f(t)\,dt = \lim_{x\searrow 0}\int_{x}^{1}\frac{2}{3\sqrt[3]{t}}\,dt = \lim_{x\searrow 0}\sqrt[3]{t^2}\Big|_{x}^{1} = \lim_{x\searrow 0}1-\sqrt[3]{x^2} = 1\,.$$

Define $F:[-1,1]\to\mathbb{R}$ as

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

If $x=-1$, the $F(-1)=0$.
If $-1<x<0$, then

$$F(x) = \int_{-1}^{x}\frac{2}{3\sqrt[3]{t}}\,dt = \sqrt[3]{x^2}-1\,.$$

If $x=0$, then

$$F(0) = \int_{-1}^{0} f(t)\,dt = -1\,.$$

If $0 < x \leq 1$, then

$$F(x) = \int_{-1}^{0} f(t)\,dt + \int_{0}^{x} f(t)\,dt = F(0) + \int_{0}^{x} \frac{2}{3\sqrt[3]{t}}\,dt = \sqrt[3]{x^2} - 1\,.$$

In conclusion, $F(x) = \sqrt[3]{x^2} - 1$ for all $x \in [-1, 1]$. Observe that, F is absolutely continuous on $[-1, 1]$ and differentiable everywhere, except $x = 0$.

Example 5.22. Consider $F(x) = \sqrt{x}$. We want to prove that F is absolutely continuous on $[0, 1]$. One way to do this is, to prove that F can be defined as the antiderivative of a Riemann or improper Riemann integrable function on $[0, 1]$. Indeed, let $f(x) = \frac{1}{2\sqrt{x}}$, which is improper Riemann integrable on $[0, 1]$ and

$$\sqrt{x} = \int_{0}^{x} \frac{1}{2\sqrt{t}}\,dt\,.$$

Theorem 5.23. *(The Fundamental Theorem of Calculus)*
Let $f, F : [a, b] \to \mathbb{R}$ such that:

(1) *f is Riemann integrable on $[a, b]$.*

(2) *F is continuous on $[a, b]$ and differentiable on (a, b).*

(3) *$F'(x) = f(x)$ for all $x \in (a, b)$.*

Then,

$$\int_{a}^{b} f(x)\,dx = F(b) - F(a)\,.$$

Proof. Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. Then, for each subinterval $[x_{i-1}, x_i]$ we apply Theorem 4.14 to get a point $t_i \in (x_{i-1}, x_i)$ such that

$$F(x_i) - F(x_{i-1}) = f(t_i)(x_i - x_{i-1})\,.$$

In this way we get

$$\begin{aligned} F(b) - F(a) &= (F(x_1) - F(x_0)) + (F(x_2) - F(x_1)) + \cdots + (F(x_n) - F(x_{n-1})) \\ &= \sum_{i=1}^{n} f(t_i)\Delta x_i\,. \end{aligned}$$

Hence, for all partitions P we have

$$\sum_{i=1}^{n} m_i(f)\Delta x_i \leq F(b) - F(a) \leq \sum_{i=1}^{n} M_i(f)\Delta x_i\,,$$

which leads to

$$\underline{S}(f,P) \le F(b) - F(a) \le \overline{S}(f,P)\,.$$

Taking the supremum of the lower sums and infimum of the upper sums, we get

$$\underline{S}(f) \le F(b) - F(a) \le \overline{S}(f)\,.$$

The Riemann integrability of f implies that all three quantities are equal, so

$$\int_a^b f(x)\,dx = F(b) - F(a)\,.$$

□

Definition 5.24. The function F, given in Theorem 5.23, it is called an antiderivative of f on the interval $[a,b]$.

Theorem 5.25. *(Integration by Parts)*
Let $f,g : [a,b] \to \mathbb{R}$ be differentiable functions on $[a,b]$. If $f' \cdot g$ and $f \cdot g'$ are Riemann integrable on $[a,b]$, then

$$\int_a^b f(x)g'(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)\,dx\,.$$

Proof. The product rule for derivatives implies that

$$f(x)g'(x) + f'(x)g(x) = \Big(f(x)g(x)\Big)'\,.$$

Each term is Riemann integrable, so by Theorem 5.23 we get

$$\int_a^b f(x)g'(x)\,dx + \int_a^b f(x)g'(x)\,dx = \int_a^b \Big(f(x)g(x)\Big)'\,dx = f(b)g(b) - f(a)g(a)\,.$$

The Integration by Parts formula now follows after moving the second integral from the left to the right side.

□

Theorem 5.26. *(Change of Variables)*
Let $f : [c,d] \to \mathbb{R}$ and $g : [a,b] \to [c,d]$. Assume that:

(1) *f is continuous on $[c,d]$.*

(2) *g is differentiable on $[a,b]$ and g' is continous on $[a,b]$.*

Then,

$$\int_a^b f\big(g(x)\big)g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)du\,.$$

Proof. Define $F : [c, d] \to \mathbb{R}$ as

$$F(x) = \int_{g(a)}^{x} f(u)\, du\,.$$

The continuity of f on $[c, d]$ implies that F is differentiable on $[c, d]$ and $F'(x) = f(x)$ for all $x \in [c, d]$. Then, the chain rule for derivatives implies that

$$\big(F(g(x))\big)' = F'\big(g(x)\big)\, g'(x) = f\big(g(x)\big)\, g'(x)\,,$$

and hence

$$\int_a^b f\big(g(x)\big) g'(x)\, dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(u) du\,.$$

□

Exercises

Exercise 5.1. Let $f, g : [a, b] \to \mathbb{R}$ be bounded functions.

(1) Show that if f is Riemann integrable and $f(x) = g(x)$ for every $x \in [a, b]$ except finitely many points, then g is Riemann integrable and $\int_a^b f(x)\, dx = \int_a^b g(x)\, dx$.

(2) Will the result from (1) remain true, if $f(x) = g(x)$ for every $x \in [a, b]$ except countably many points?

Exercise 5.2. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Assume that $\int_a^b f(x)\, dx = 0$.
(1) Show that there exists $c \in (a, b)$ such that $f(c) = 0$.
(2) Show that if $f(x) \geq 0$ for all $x \in [a, b]$, then $f(x) = 0$ for all $x \in [a, b]$.

Exercise 5.3. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Assume that $\int_a^b f(x)g(x)\, dx = 0$ for all Riemann integrable functions $g : [a, b] \to \mathbb{R}$. Show that $f(x) = 0$ for all $x \in [a, b]$.

Exercise 5.4. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Assume that $\int_a^x f(t)\, dt = 0$ for all $x \in [a, b]$. Show that $f(x) = 0$ for all $x \in [a, b]$.

Exercise 5.5. Let $f, g : [a, b] \to \mathbb{R}$ be continuous functions and assume that $g(x) \geq 0$ for all $x \in [a, b]$.
(1) Show that there exists $c \in [a, b]$ such that

$$\int_a^b f(x)g(x)\, dx = f(c) \int_a^b g(x)\, dx\,.$$

(2) Give an example showing that the assumption $g(x) \geq 0$ for all $x \in [a, b]$ cannot be dropped.

Exercise 5.6. Let $f, g : [a, b] \to \mathbb{R}$ be Riemann integrable functions and $p, q \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Show that:
(1) Hölder's inequality holds:

$$\left| \int_a^b f(x)g(x)\, dx \right| \leq \left(\int_a^b |f(x)|^p\, dx \right)^{\frac{1}{p}} \cdot \left(\int_a^b |g(x)|^q\, dx \right)^{\frac{1}{q}} .$$

(2) Minkowski's inequality holds:

$$\left(\int_a^b |f(x) + g(x)|^p\, dx \right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p\, dx \right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p\, dx \right)^{\frac{1}{p}} .$$

Exercise 5.7. Let $f : [1, +\infty) \to \mathbb{R}$ be a monotone decreasing function such that $f(x) > 0$ for all $x \geq 1$. For each $n \in \mathbb{N}$ define

$$x_n = \sum_{i=1}^{n} f(i) - \int_1^n f(t)\, dt .$$

Show that the sequence $\{x_n\}$ is convergent.

Exercise 5.8. Let $f : [0, 1] \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} n & \text{if } x = \frac{1}{n},\ \forall\, n \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

Is f Riemann or improper Riemann integrable?

Exercise 5.9. Let $f : [0, 1] \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} \frac{1}{n} & \text{if } x = \frac{m}{n} \in \mathbb{Q} \text{ in lowest terms} \\ 0 & \text{otherwise} \end{cases}$$

Is f Riemann or improper Riemann integrable?

Exercise 5.10. Give an example of a Riemann integrable function $f : [a, b] \to \mathbb{R}$, which does not have an antiderivative, as in Theorem 5.23.

Exercise 5.11. Let $f : [a, b] \to \mathbb{R}$ be a continuous function and suppose that

$$\int_a^x f(t)\, dt = \int_x^b f(t)\, dt\, ,\ \forall\, x \in [a, b] .$$

Show that $f(x) = 0$ for all $x \in [a, b]$.

Exercise 5.12. Let $f : [0,1] \to \mathbb{R}$ be continuous on $[0,1]$ and differentiable on $(0,1)$. Assume that $|f'(x)| \le M$ for all $x \in (0,1)$. Prove that for all $n \in \mathbb{N}$ holds

$$\left| \int_0^1 f(x)\,dx - \frac{1}{n} \sum_{k=1}^{n} f\left(\frac{k}{n}\right) \right| \le \frac{M}{n}\,.$$

Exercise 5.13. Let $f : (a,b] \to \mathbb{R}$ be a bounded function, which is Riemann integrable on $[c,b]$ for all $a < c < b$. Let $\{a_n\}$ be a sequence in (a,b) such that $\{a_n\} \to a$ and denote

$$y_n = \int_{a_n}^{b} f(x)\,dx\,,\ \forall\, n \in \mathbb{N}\,.$$

(1) Show that $\{y_n\}$ is a Cauchy sequence.
(2) Show that $\{y_n\}$ is convergent.
(3) Show that the sequence

$$\left\{ \int_{\frac{1}{n}}^{1} \sin\frac{1}{x}\,dx \right\}_{n\in\mathbb{N}}$$

is convergent.

Exercise 5.14. Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable on $[a,b]$. Prove that for any $\varepsilon > 0$ there exists a continuous function $g_\varepsilon : [a,b] \to \mathbb{R}$ such that

$$\int_a^b |f(x) - g_\varepsilon(x)|\,dx < \varepsilon\,.$$

Exercise 5.15. Let $f : [0,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2 \sin\frac{1}{x^2} & \text{if } 0 < x \le 1 \\ 0 & \text{if } x = 0\,. \end{cases}$$

Show that f is differentiable on $[0,1]$, but f' is not Riemann integrable.

Exercise 5.16. Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable function on $[a,b]$ and let

$$F(x) = \int_a^x f(t)\,dt\,,\ \forall\, x \in [a,b]\,.$$

Show that there exist two monotone increasing functions $\alpha, \beta : [a,b] \to \mathbb{R}$ such that

$$F(x) = \alpha(x) - \beta(x)\,,\ \forall\, x \in [a,b]\,.$$

Exercise 5.17. Show that $f(x) = \sqrt{x}$ is absolutely continuous on $[0, +\infty)$.

Exercise 5.18. Suppose that f is a continuous, nonnegative function on the interval $[0, 1]$. Let

$$M = \sup \left\{ f(x) \;\middle|\; x \in [0, 1] \right\}.$$

Prove that

$$\lim_{n \to \infty} \left(\int_0^1 \big(f(t)\big)^n \, dt \right)^{\frac{1}{n}} = M .$$

5.2 The Riemann-Stieltjes integral

5.2.1 Definitions and properties

The Riemann-Stieltjes integral is a generalization if the Riemann integral and has important applications in probability and statistics and in functional analysis, where it is used to describe the linear functionals on spaces of continuous functions.

Definition 5.27. Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $\alpha : [a,b] \to \mathbb{R}$ be a monotonically increasing function.

Consider a partition $P = \{x_0, x_1, ...x_n\}$ of the interval $[a,b]$. Similarly to the Riemann integral, we define the following quantities:

$$\begin{aligned}
&m_i(f) = \inf\{f(x) \mid x_{i-1} \le x \le x_i\} \\
&M_i(f) = \sup\{f(x) \mid x_{i-1} \le x \le x_i\} \\
&\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})\,.
\end{aligned}$$

The lower Riemann-Stieltjes sum of f associated to the partition P is

$$\underline{S}(f, P, \alpha) = \sum_{i=1}^{n} m_i(f)\Delta\alpha_i\,.$$

The upper Riemann-Stieltjes sum of f associated to the partition P is

$$\overline{S}(f, P, \alpha) = \sum_{i=1}^{n} M_i(f)\Delta\alpha_i\,.$$

The lower Riemman-Stieltjes sum of f is

$$\underline{S}(f, \alpha) = \sup\{\underline{S}(f, P, \alpha) \mid P \text{ is any partition of } [a,b]\}.$$

The upper Riemman-Stietjes sum of f is

$$\overline{S}(f, \alpha) = \inf\{\overline{S}(f, P, \alpha) \mid P \text{ is any partition of } [a,b]\}.$$

We say that f is Riemann-Stieltjes integrable with respect to α on $[a,b]$ if

$$\underline{S}(f, \alpha) = \overline{S}(f, \alpha)\,.$$

If f is Riemann-Stieltjes integrable, then the common value of the upper and lower Riemann-Stieltjes sums is denoted by

$$\int_a^b f(x)d\alpha(x)\,,$$

and it is called the Riemann-Stieltjes integral of f.
Also, we define

$$\int_a^a f(x)\,d\alpha(x) = 0\,.$$

and

$$\int_b^a f(x)\,d\alpha(x) = -\int_a^b f(x)\,d\alpha(x)\,.$$

Similarly to Lemma 5.5, we present some basic properties of the Riemann-Stieltjes sums.

Lemma 5.28. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $\alpha : [a,b] \to \mathbb{R}$ be a monotonically increasing function. Then, the following statements hold:*

(1) *If P_1, P_2 are partitions of $[a,b]$, $P_1 \subset P_2$, then*

$$\underline{S}(f, P_1, \alpha) \le \underline{S}(f, P_2, \alpha) \ \ \text{and} \ \ \overline{S}(f, P_1, \alpha) \ge \overline{S}(f, P_2, \alpha)\,.$$

(2) *If P and Q are any two partitions of $[a,b]$, then*

$$\underline{S}(f, P, \alpha) \le \overline{S}(f, Q, \alpha)\,.$$

(3)

$$\underline{S}(f, \alpha) \le \overline{S}(f, \alpha)\,.$$

(4) $\underline{S}(f,\alpha) = \overline{S}(f,\alpha)$ *if and only if for all $\varepsilon > 0$ there exists a partition P_ε such that*

$$\overline{S}(f, P_\varepsilon, \alpha) - \underline{S}(f, P_\varepsilon, \alpha) < \varepsilon\,.$$

Example 5.29. The Heaviside function, or unit step function, is a very important example for the increasing function α. It is defined as

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } 0 \le x \end{cases}\,.$$

Let us show that $f(x) = x + 2$ is Riemann-Stieltjes integrable with respect to H on the interval $[-1,1]$.

Let $\varepsilon > 0$. Consider a partition $P_\varepsilon = \{x_0 = -1, x_1 = -\frac{\varepsilon}{2}, x_2 = 0, x_3 = 1\}$. Only $\Delta\alpha_2 = 1$ is non-zero, so

$$\overline{S}(f, P_\varepsilon, \alpha) - \underline{S}(f, P_\varepsilon, \alpha) = \Big(2 - (-\frac{\varepsilon}{2} + 2)\Big) \cdot 1 = \frac{\varepsilon}{2} < \varepsilon\,.$$

This shows that f is Riemann-Stieltjes integrable with respect to H.

Observe that, $\overline{S}(f, P_\varepsilon, \alpha) = 2$ and no matter how we choose another partition we cannot get a lower value than this. Therefore,

$$\int_{-1}^1 (x+2)\,dH(x) = 2\,.$$

Example 5.30. Consider the function $f : [-1, 1] \to \mathbb{R}$,

$$f(x) = \begin{cases} x & \text{if } x \in [-1,1] \cap \mathbb{Q} \\ 0 & \text{if } x \in [-1,1] \cap \mathbb{I} \end{cases} .$$

This function f is not Riemann integrable on $[-1, 1]$, because it is only continuous at 0.

Let $\varepsilon > 0$. Consider the same partition as before,

$$P_\varepsilon = \{x_0 = -1, x_1 = -\frac{\varepsilon}{2}, x_2 = 0, x_3 = 1\} .$$

Then,

$$\overline{S}(f, P_\varepsilon, \alpha) - \underline{S}(f, P_\varepsilon, \alpha) = \Big(0 - (-\frac{\varepsilon}{2})\Big) \cdot 1 = \frac{\varepsilon}{2} < \varepsilon .$$

This shows that f is Riemann-Stieltjes integrable with respect to H.

Observe that, $\underline{S}(f, P, \alpha) = 0$ no matter how we choose a partition. Therefore,

$$\int_{-1}^{1} f(x)\, dH(x) = 0 .$$

Theorem 5.31. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function and $\alpha : [a, b] \to \mathbb{R}$ be a monotonically increasing function. Then f is Riemann-Stieltjes integrable with respect to α on $[a, b]$.*

Proof. We can assume that $\alpha(b) - \alpha(a) > 0$, because otherwise α is a constant function and the integrability result is trivial.

Let $\varepsilon > 0$. By the uniform continuity of f, there exists $\delta = \delta(\varepsilon) > 0$ such that for all $x, y \in [a, b]$ with $|x - y| < \delta$ we have $|f(x) - f(y)| < \frac{\varepsilon}{2(\alpha(b)-\alpha(a))}$.

Let $P_\varepsilon = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$ such that $||P_\varepsilon|| < \delta(\varepsilon)$. Then for all $1 \le i \le n$ and $x_{i-1} \le x < y \le x_i$ we have

$$|f(x) - f(y)| < \frac{\varepsilon}{2(\alpha(b) - \alpha(a))} .$$

Hence,

$$M_i(f) - m_i(f) \le \frac{\varepsilon}{2(\alpha(b) - \alpha(a))} , \ \forall\, 1 \le i \le n .$$

Therefore,

$$\begin{aligned} \overline{S}(f, P_\varepsilon, \alpha) - \underline{S}(f, P_\varepsilon, \alpha) &= \sum_{i=1}^{n} (M_i(f) - m_i(f)) \Delta\alpha_i \le \frac{\varepsilon}{2(\alpha(b) - \alpha(a))} \sum_{i=1}^{n} \Delta\alpha_i \\ &= \frac{\varepsilon}{2(\alpha(b) - \alpha(a))} (\alpha(b) - \alpha(a)) = \frac{\varepsilon}{2} < \varepsilon . \end{aligned}$$

Hence, by Lemma 5.28 (4), it follows that f is Riemann-Stietjes integrable with respect to α on $[a, b]$.

$\square$

The next theorem holds for both monotone increasing or decreasing functions f.

Theorem 5.32. *Let $f : [a,b] \to \mathbb{R}$ be a monotone increasing function and $\alpha : [a,b] \to \mathbb{R}$ be a monotonically increasing and continuous function. Then f is Riemann-Stieltjes integrable with respect to α on $[a,b]$.*

Proof. Let $\varepsilon > 0$. We can suppose that $\alpha(b) - \alpha(a) > 0$ and let $n \in \mathbb{N}$ such that

$$n > \frac{\big(\alpha(b) - \alpha(a)\big)\big(f(b) - f(a)\big)}{\varepsilon}.$$

By the continuity of α, we can choose a partition $P = \{x_0, x_1, \ldots, x_n\}$ such that $\Delta\alpha_i = \frac{1}{n}\big(\alpha(b) - \alpha(a)\big)$. Then,

$$\begin{aligned}
\overline{S}(f,P,\alpha) - \underline{S}(f,P,\alpha) &= \sum_{i=1}^{n}\big(f(x_i) - f(x_{i-1})\big)\Delta\alpha_i \\
&= \frac{\alpha(b) - \alpha(a)}{n}\sum_{i=1}^{n}\big(f(x_i) - f(x_{i-1})\big) \\
&= \frac{\alpha(b) - \alpha(a)}{n}\big(f(b) - f(a)\big) < \varepsilon\,.
\end{aligned}$$

Therefore, f is Riemann-Stieltjes integrable with respect to α on $[a,b]$.

□

Theorem 5.33. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $\alpha : [a,b] \to \mathbb{R}$ be a monotonically increasing function. Assume that f has finitely many discontinuity points in $[a,b]$ and α is continuous at each of these points. Then f is Riemann-Stieltjes integrable with respect to α on $[a,b]$.*

Proof. Let $\varepsilon > 0$ and $M = \sup\{|f(x)| \mid x \in [a,b]\}$. Let us cover the finite set of discontinuity points by pairwise disjoint interval $[u_j, v_j]$, $1 \leq j \leq m$, such that each discontinuity point lies in the interior relative to $[a,b]$ and

$$\sum_{j=1}^{m}\alpha(v_j) - \alpha(u_j) < \frac{\varepsilon}{2(M+1)}\,.$$

Let K be the subset of $[a,b]$ obtained by removing the interiors of these intervals, relative to $[a,b]$. Then K is compact and f is uniformly continuous on K. Therefore, there exists $\delta = \delta(\varepsilon) > 0$ such that $|f(x) - f(y)| < \frac{\varepsilon}{2(\alpha(b)-\alpha(a)+1)}$ if $x, y \in K$ and $|x - y| < \delta$.

Now choose a partition $P = \{x_0, x_1, \ldots, x_n\}$ in the following way: Each u_j, v_j occurs in P as consecutive points; Cover K by intervals $[x_k, y_k]$, each with length less then δ. Note that some x_k and y_k coincide with some u_j and v_j. Then

$$\overline{S}(f,P,\alpha) - \underline{S}(f,P,\alpha) = \sum_{j}(M_j - m_j)\Delta\alpha_j + \sum_{k}(M_k - m_k)\Delta\alpha_k$$

$$
\begin{aligned}
&\leq \sum_j 2M\Delta\alpha_j + \sum_k \frac{\varepsilon}{2(\alpha(b)-\alpha(a)+1)}\Delta\alpha_k \\
&\leq 2M\frac{\varepsilon}{2(M+1)} + \frac{\varepsilon}{2(\alpha(b)-\alpha(a)+1)}(\alpha(b)-\alpha(a)) < \varepsilon\,.
\end{aligned}
$$

Therefore, f is Riemann-Stieltjes integrable with respect to α on $[a,b]$.

□

For the next theorem we use the translated version of the Heaviside function:

$$
H(x-c) = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } c \leq x \end{cases}\,.
$$

Theorem 5.34. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $a < c \leq b$. If f is continuous from left at c, then f is Riemann-Stieltjes integrable with respect to $H(x-c)$ on $[a,b]$ and*

$$
\int_a^b f(x)\,dH(x-c) = f(c)\,.
$$

Proof. Let $\varepsilon > 0$. From $\lim_{x\nearrow c} f(x) = f(c)$ it follows that we can choose $\delta = \delta(\varepsilon) > 0$ such that $a < c-\delta$ and $|f(x)-f(c)| < \frac{\varepsilon}{4}$ if $c-\delta \leq x \leq c$. Choose the partition

$$
P_\varepsilon = \{a = x_0 < x_1 = c-\delta < x_2 = c < x_3 = b\}\,,\ \text{if } a < c < b\,,
$$

or

$$
P_\varepsilon = \{a = x_0 < x_1 = c-\delta < x_2 = b\}\,,\ \text{if } c = b\,.
$$

Then,

$$
\overline{S}(f,P_\varepsilon,\alpha) - \underline{S}(f,P_\varepsilon,\alpha) = (M_2 - m_2)\Delta\alpha_2 \leq \frac{\varepsilon}{2}\cdot 1 < \varepsilon\,.
$$

Therefore, f is Riemann-Stieltjes integrable with respect to $H(x-c)$ on $[a,b]$. To calculate the integral, note that for any refinement P of P_ε

$$
\inf\{f(x)\,\big|\, c-\delta \leq x \leq c\} \leq \underline{S}(f,P,\alpha) \leq \overline{S}(f,P,\alpha) \leq \sup\{f(x)\,\big|\, c-\delta \leq x \leq c\}\,,
$$

However, as δ decreases to zero, the continuity of f from left at c gives

$$
\int_a^b f(x)\,dH(x-c) = f(c)\,.
$$

□

Theorem 5.35. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $\alpha : [a,b] \to \mathbb{R}$ be a monotonically increasing function which is differentiable on $[a,b]$. If f and α' are Riemann integrable on $[a,b]$, then f is Riemann-Stieltjes integrable with respect to α on $[a,b]$ and*

$$
\int_a^b f(x)\,d\alpha(x) = \int_a^b f(x)\alpha'(x)\,dx\,.
$$

Proof. Let $\varepsilon > 0$. The Riemann integrability of α' implies that there exists a partition $P_\varepsilon = \{x_0, x_1, \dots, x_n\}$ such that

$$\overline{S}(\alpha', P_\varepsilon) - \underline{S}(\alpha', P_\varepsilon) < \varepsilon\,. \tag{5.2}$$

By the Largrange's Theorem 4.14, for each $1 \le i \le n$ there exists $t_i \in [x_{i-1}, x_i]$ such that $\Delta\alpha_i = \alpha'(t_i)\Delta x_i$. By inequality (5.2), for any $s_i \in [x_{i-1}, x_i]$ we have

$$\sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i < \varepsilon\,. \tag{5.3}$$

Since,

$$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i = \sum_{i=1}^{n} f(s_i)\alpha'(t_i)\Delta x_i\,,$$

inequality (5.3) implies that for any $s_i \in [x_{i-1}, x_i]$ we have

$$\left|\sum_{i=1}^{n} f(s_i)\Delta\alpha_i - \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i\right| < M\varepsilon\,, \tag{5.4}$$

where $M = \sup\big\{|f(x)| \,\big|\, x \in [a,b]\big\}$. Inequality (5.4) implies

$$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i < M\varepsilon + \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i\,,$$

and hence

$$\overline{S}(f, P_\varepsilon, \alpha) \le M\varepsilon + \overline{S}(f\alpha', P_\varepsilon)\,.$$

In a similar way, starting again from (5.4), we get

$$\overline{S}(f\alpha', P_\varepsilon) \le M\varepsilon + \overline{S}(f, P_\varepsilon, \alpha)\,.$$

In conclusion, we obtained that

$$\big|\overline{S}(f, P_\varepsilon, \alpha) - \overline{S}(f\alpha', P_\varepsilon)\big| \le M\varepsilon\,.$$

This last inequality holds also for any refinement of P_ε, so

$$\big|\overline{S}(f, \alpha) - \overline{S}(f\alpha')\big| \le M\varepsilon\,.$$

As $\varepsilon > 0$ is arbitrary, we obtain

$$\overline{S}(f, \alpha) = \overline{S}(f\alpha')\,.$$

Going back to (5.4) and restarting the estimates with

$$-M\varepsilon + \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i \le \sum_{i=1}^{n} f(s_i)\Delta\alpha_i\,,$$

will lead to

$$\underline{S}(f, \alpha) = \underline{S}(f\alpha') .$$

We know that $f\alpha'$ is Riemann integrable as a product of two Riemann integrable functions, so we finally get that f is Riemann-Stieltjes integrable with respect to α and

$$\int_a^b f(x)\, d\alpha(x) = \int_a^b f(x)\alpha'(x)\, dx .$$

$\square$

In the next theorem we list some fundamental properties of the Riemann-Stieltjes integral. The proofs are based on the properties of lower and upper Riemann-Stieltjes sums and we leave them as exercises.

Theorem 5.36. *Let $f, f_1, f_2 : [a, b] \to \mathbb{R}$ be bounded, $\alpha, \alpha_1, \alpha_2 : [a, b] \to \mathbb{R}$ be monotone increasing functions, $a < c < b$, and $\lambda \in \mathbb{R}$. Then, the following statements hold:*

(1) $\displaystyle\int_a^b f_1(x) + f_2(x)\, d\alpha(x) = \int_a^b f_1(x)\, d\alpha(x) + \int_a^b f_2(x)\, d\alpha(x) .$

(2) $\displaystyle\int_a^b \lambda f(x)\, d\alpha(x) = \lambda \int_a^b f(x)\, d\alpha(x) .$

(3) $\displaystyle\int_a^b f(x)\, d(\alpha_1(x) + \alpha_2(x)) = \int_a^b f(x)\, d\alpha_1(x) + \int_a^b f(x)\, d\alpha_2(x) .$

(4) $\displaystyle\int_a^b f(x)\, d(\lambda\alpha(x)) = \lambda \int_a^b f(x)\, d\alpha(x) .$

(5) $\displaystyle\int_a^b f(x)\, d\alpha(x) = \int_a^c f(x)\, d\alpha(x) + \int_c^b f(x)\, d\alpha(x) .$

(6) $\displaystyle\int_a^b f_1(x)\, d\alpha(x) \le \int_a^b f_2(x)\, d\alpha(x) \,,\ \textit{if}\ \ f_1(x) \le f_2(x) \,,\ \forall\, x \in [a, b] .$

Example 5.37. Let $f, \alpha : [0, 2] \to \mathbb{R}$ be defined as:

$$f(x) = x + 1$$

and

$$\alpha(x) = \begin{cases} x^2 & \text{if } 0 \le x < 1 \\ 2 & \text{if } x = 1 \\ 4\sqrt{x} & \text{if } 1 < x \le 2 \end{cases} .$$

Let's calculate

$$\int_0^2 f(x)\, d\alpha(x) .$$

By Theorem 5.36, we have

$$\int_0^2 f(x)\,d\alpha(x) = \int_0^1 f(x)\,d\alpha(x) + \int_1^2 f(x)\,d\alpha(x)\,.$$

On $[0,1]$ we can write $\alpha(x) = x^2 + H(x-1)$, so

$$\begin{aligned}\int_0^1 f(x)\,d\alpha(x) &= \int_0^1 (x+1)d(x^2) + \int_0^1 (x+1)dH(x-1)\\ &= \int_0^1 (x+1)2x\,dx + 2 = \frac{11}{3}\,.\end{aligned}$$

On $[1,2]$ we can write $\alpha(x) = 2\tilde{H}(x-1) + 4\sqrt{x}$, where

$$\tilde{H}(x-1) = \begin{cases} 0 & \text{if } 0 \le x \le 1 \\ 1 & \text{if } 1 < x \le 2 \end{cases}\,.$$

Therefore,

$$\begin{aligned}\int_1^2 f(x)\,d\alpha(x) &= \int_1^2 (x+1)d(4\sqrt{x}) + 2\int_1^2 (x+1)d\tilde{H}(x-1)\\ &= \int_1^2 (x+1)\frac{2}{\sqrt{x}}\,dx + 4 = \frac{28}{3}\,.\end{aligned}$$

Adding the two partial answers gives

$$\int_0^2 f(x)\,d\alpha(x) = 13\,.$$

5.2.2 Functions of bounded variation

Definition 5.38. A function $f : [a,b] \to \mathbb{R}$ is said to be of bounded variation, if there exists $M > 0$ such that for any partition $P = \{x_0, x_1, ..., x_n\}$ of $[a,b]$ we have

$$\sum_{i=1}^n \big|f(x_i) - f(x_{i-1})\big| \le M\,. \tag{5.5}$$

We define the total variation of f on $[a,b]$, and denote it by $V_a^b(f)$, the infimum of the constants M from (5.5). Equivalently,

$$V_a^b(f) = \sup\Big\{\sum_{i=1}^n \big|f(x_i) - f(x_{i-1})\big| \;\Big|\; \text{for any partition } P = \{x_0, x_1, ..., x_n\}\Big\}\,. \tag{5.6}$$

The next theorem contains some of the most basic properties of functions of bounded variation.

Theorem 5.39. *Let $f, g : [a,b] \to \mathbb{R}$ be functions of bounded variation, $a < c < b$ and $\lambda \in \mathbb{R}$. Then:*

(1) *$f + g$ is of bounded variation and*

$$V_a^b(f+g) \le V_a^b(f) + V_a^b(g)\,.$$

(2) *λf is of bounded variation and*

$$V_a^b(\lambda f) = |\lambda| V_a^b(f)\,.$$

(3) *f is a bounded variation of on $[a,c]$ and $[c,b]$ and*

$$V_a^b(f) = V_a^c(f) + V_c^b(f)\,.$$

Proof. (1) The proof evidently follows from the inequality

$$\big|(f+g)(x_i) - (f+g)(x_{i-1})\big| \le \big|f(x_i) - f(x_{i-1})\big| + \big|g(x_i) - g(x_{i-1})\big|\,.$$

(2) Again, the proof follows from as easy formula

$$\big|\lambda f(x_i) - \lambda f(x_{i-1})\big| = |\lambda|\,\big|f(x_i) - f(x_{i-1})\big|\,.$$

(3) Let P be any partition of $[a,b]$. If c is not a point in the partition, we can add it to the partition, because the sum in (5.6) increases with refined partitions. Now as c is a point in P (say $c = x_k$ for some $1 < k < n$), then the partition can be split into the union of a partition of $[a,c]$ and a partition of $[c,b]$ and we obtain:

$$\sum_{i=1}^{n}\big|f(x_i) - f(x_{i-1})\big| = \sum_{i=1}^{k}\big|f(x_i) - f(x_{i-1})\big| + \sum_{i=k}^{n}\big|f(x_i) - f(x_{i-1})\big|\,.$$

Therefore,

$$\sum_{i=1}^{n}\big|f(x_i) - f(x_{i-1})\big| \le V_a^c(f) + V_c^b(f)\,,$$

and by considering the supremum on the left-side we get

$$V_a^b(f) \le V_a^c(f) + V_c^b(f)\,.$$

To get the opposite inequality, consider and arbitrary $\varepsilon > 0$ and two partitions, P_1 of $[a,c]$ and P_2 of $[c,b]$, such that

$$\sum_{i=1}^{k}\big|f(x_i) - f(x_{i-1})\big| \ge V_a^c(f) - \frac{\varepsilon}{2}$$

and

$$\sum_{j=1}^{l} \big| f(x_j) - f(x_{j-1}) \big| \geq V_c^b(f) - \frac{\varepsilon}{2} \,.$$

Consider $P = P_1 \cup P_2 = \{x_0, x_1, \ldots, x_n\}$ and get

$$\sum_{i=1}^{n} \big| f(x_i) - f(x_{i-1}) \big| \geq V_a^c(f) + V_c^b - \varepsilon \,.$$

Therefore,

$$V_a^b(f) \geq V_a^c(f) + V_c^b - \varepsilon \,,\ \forall\, \varepsilon > 0 \,,$$

which implies that

$$V_a^b(f) \geq V_a^c(f) + V_c^b(f) \,.$$

$\square$

Theorem 5.40.
(1) *Any monotone function $f : [a,b] \to \mathbb{R}$ is of bounded variation and*

$$V_a^b(f) = \Big| f(b) - f(a) \Big| \,.$$

(2) *Any function which is absolutely continuous on $[a,b]$, it is of bounded variation on $[a,b]$.*

Proof. (1) Consider f to be monotone decreasing and let P be a partition of $[a,b]$. Then

$$\sum_{i=1}^{n} \big| f(x_i) - f(x_{i-1}) \big| = \sum_{i=1}^{n} f(x_{i-1}) - f(x_i) = f(a) - f(b) = |f(b) - f(a)| \,.$$

The proof for a monotone increasing function is identical.

(2) Let $\varepsilon > 0$. By the definition of absolute continuity, there exists a $\delta = \delta(\varepsilon) > 0$ such that for any collection (finite or countable) of pairwise disjoint intervals $(a_i, b_i) \subset [a,b]$ with

$$\sum_i (b_i - a_i) < \delta \,,$$

it follows

$$\sum_i \big| f(b_i) - f(a_i) \big| < \varepsilon \,.$$

Therefore, for any $[c,d] \subset [a,b]$ with $d - c \leq \frac{\delta}{2}$ we have

$$V_c^d(f) \leq \varepsilon \,.$$

We can cover $[a,b]$ by finitely many, say N, intervals fo length less or equal to $\frac{\delta}{2}$, so by Theorem 5.39 (3) we get

$$V_a^b(f) \leq \varepsilon N \,.$$

$\square$

The proof of the next lemma depends only on the simple observation that the sum from (5.6) increases if we enlarge the interval.

Lemma 5.41. *Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation. Consider the function $\alpha : [a,b] \to \mathbb{R}$ defined by*

$$\alpha(x) = V_a^x(f)\,,\ \forall\, a \le x \le b\,.$$

Then, α is monotone increasing on $[a,b]$ and for any $a \le x < y \le b$ we have

$$\Big|f(y) - f(x)\Big| \le \alpha(y) - \alpha(x)\,.$$

Theorem 5.42. *Let $f : [a,b] \to \mathbb{R}$ be a function. Then f is of bounded variation if and only if there exist two monotone increasing functions $\alpha, \beta : [a,b] \to \mathbb{R}$ such that $f = \alpha - \beta$.*

Proof. Let α be the function defined in Lemma 5.41 and

$$\beta(x) = \alpha(x) - f(x)\,.$$

Let $a \le x < y \le b$. Then, Lemma 5.41 implies that

$$\beta(y) - \beta(x) = \Big(\alpha(y) - \alpha(x)\Big) + \Big(f(y) - f(x)\Big) \ge 0\,,$$

which shows that β is monotone increasing.

Based on Theorems 5.39 and 5.40, the reverse implication is starightforward.

$\square$

Let us return to the Riemann-Stieltjes interval. If β is monotone increasing function, then $-\beta$ is monotone decreasing and we can define

$$\int_a^b f(x)\, d(-\beta(x)) = -\int_a^b f(x)\, d\beta(x)\,.$$

Definition 5.43. If $f, g : [a,b] \to \mathbb{R}$ and g is of bounded variation, we can define the Riemann-Stieltjes integral of f with respect to g on [a,b] as

$$\int_a^b f(x)\, dg(x) = \int_a^b f(x)\, d\alpha(x) - \int_a^b f(x) d\beta(x)\,,$$

provided that $g = \alpha - \beta$ as in Theorem 5.42 and f is Riemann-Stieltjes integrable with respect to both α and β on $[a,b]$.

Remark 5.44. All the results regarding the Riemann-Stieltjes integral with respect to monotone increasing functions can be reformulated with respect to functions of bounded variation.

To highlight the importance of functions of bounded variations, we state the following theorem. In terms of notations, $C([a,b])$ denotes the vector space of all continuous functions defined on $[a,b]$. (See also Chapter 7.)

Theorem 5.45. *(F. Riesz, 1909)*
Every continuous linear functional $L : C([a,b]) \to \mathbb{R}$ can be represented as

$$L(f) = \int_a^b f(x)\, d\alpha(x)\,,\ \forall\ f \in C([a,b])\,,$$

where $\alpha : [a,b] \to \mathbb{R}$ is a function of bounded variation on $[a,b]$.

Exercises

Exercise 5.19. Let

$$\alpha(x) = \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 4 & \text{if } 1 \le x \le 2 \end{cases}$$

and

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 3 & \text{if } 1 < x \le 2 \end{cases}$$

Use the definition to prove that f is Riemann-Stieltjes integrable with respect to α on $[0,2]$, and find $\int_0^2 f(x)\, d\alpha(x)$.

Exercise 5.20. Let

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

and $f : [-2,2] \to \mathbb{R}$ be a bounded function. Show that if f is Riemann-Stieltjes integrable with respect to α on $[-2,2]$, then f is continuous at 0.

Exercise 5.21. Find

$$\int_0^{\frac{\pi}{2}} \cos x\, d(\sin x),$$

and compare it with

$$\int_0^{\frac{\pi}{2}} \cos x\, dx.$$

Exercise 5.22. Let

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{I} \end{cases} .$$

Is f Riemann-Stieltjes integrable with respect to the Heaviside function $H(x)$ on $[-1, 1]$?

Exercise 5.23. Let $f : [a, b] \to \mathbb{R}$ be differentiable on $[a, b]$ and assume that $|f'(x)| \leq M$ for all $x \in [a, b]$. Show that f is a bounded variation and

$$V_a^b(f) \leq M(b - a) .$$

Exercise 5.24. Let $f : [a, b] \to \mathbb{R}$ be differentiable on $[a, b]$ and assume that f' is Riemann integrable on $[a, b]$. Show that f is a bounded variation and

$$V_a^b(f) = \int_a^b |f'(x)| \, dx .$$

Exercise 5.25. Let $g(x) = \sin x$ defined on the interval $[0, \pi]$. Use Theorem 5.42 to find the monotone increasing functions α and β such that $g = \alpha - \beta$.
Calculate

$$\int_0^\pi x \, dg(x) .$$

Exercise 5.26. Let $f : [a, b] \to \mathbb{R}$ be a function of bounded variation. Show that:

(1) f cannot have discontinuities of second kind.

(2) f is continuous on $[a, b]$, except maybe a set, which is at most countable.

(3) f is differentiable almost everywhere.

(4) Give an example of a continuous function on a compact interval, which is not of bounded variation.

Exercise 5.27. Use the definition to show that any Lipschitz function is of bounded variation.

Exercise 5.28. Let $p > 0$ and

$$f(x) = \begin{cases} x^p \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0\,. \end{cases}$$

Show that f is of bounded variation if and only if $p > 1$.

Exercise 5.29. Let $f : [a, b] \to \mathbb{R}$ be a function of bounded variation on $[a, b]$ and let $\alpha(x) = V_a^x(f)$. If $x_0 \in [a, b]$, show that:

(1) If f is continuous from left at x_0, then so is α.

(2) If f is continuous from right at x_0, then so is α.

Chapter 6

Series of numbers

6.1 Definitions

Definition 6.1. Let $\{x_n\}$ be a sequence of numbers and define

$$S_k = \sum_{n=1}^{k} x_n \,, \ \forall\, k \in \mathbb{N} \,,$$

which we call partial sums. The series

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

is defined as the ordered pair

$$\Big(\{x_n\}, \{S_k\}\Big) \,,$$

where $\{x_n\}$ is the sequence of terms and $\{S_k\}$ is the sequence of partial sums.

Definition 6.2. We say that the series $\sum_{n=1}^{\infty} x_n$ is convergent if the sequence of partial sums $\{S_k\}$ is convergent.

If $\{S_k\} \to S$, then we use the notation

$$S = \sum_{n=1}^{\infty} x_n \,,$$

and call it the sum of the series.

Otherwise, we say that the series is divergent.

Remark 6.3. If the series $\sum_{n=1}^{\infty} x_n$ is convergent, then we use the same notation for the series and for the sum of the series. This might seem confusing at first, but it is customary in mathematical writing.

Theorem 6.4. *The series* $\sum_{n=1}^{\infty} x_n$ *is convergent if and only if for all* $\varepsilon > 0$ *there exists* $K = K(\varepsilon) > 0$ *such that*

$$\left| \sum_{n=k}^{m} x_n \right| < \varepsilon \,,\ \forall\, m \geq k > K \,,$$

Proof. The proof follows from the observations that the sequence of partial sums $\{S_k\}$ is convergent if and only if a Cauchy sequence.

□

Theorem 6.5. *If the series* $\sum_{n=1}^{\infty} x_n$ *is convergent, then* $\{x_n\} \to 0$.

Proof. The proof follows from Theorem 6.4, by considering $k = m > K$.

□

The contrapositive form of Theorem 6.5 immediately gives the Test for Divergence.

Theorem 6.6. *(The Test for Divergence)*
If $\{x_n\}$ *does not converge to* 0*, then the series* $\sum_{n=1}^{\infty} x_n$ *is divergent.*

The next two series are very important, because we will decide the convergence of the majority of series by comparison with them.

Definition 6.7. Let a and r be two non-zero numbers. By geometric series we mean the series

$$\sum_{n=0}^{\infty} ar^n \,.$$

Definition 6.8. Let $p \in \mathbb{R}$. By a p-series we mean the series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \,.$$

Theorem 6.9. *The geometric series* $\sum_{n=0}^{\infty} ar^n$ *is convergent if and only if* $|r| < 1$*. In this case,*

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r} \,.$$

Proof. Consider the partial sums:

$$S_k = \sum_{n=0}^{k} ar^n .$$

Observe that if $r = 1$, then $S_k = k + 1$ and therefore the sequence of partial sums diverges. If $r \neq 1$, then

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

which shows that S_k converges if and only if $|r| < 1$. In this case,

$$\{r^{n+1}\} \to 0 ,$$

so

$$\{S_k\} \to \frac{a}{1 - r} .$$

□

Example 6.10. The following geometric series is often used to prove that countable sets have measure zero. Let $\varepsilon > 0$ and consider the geometric series

$$\sum_{n=0}^{\infty} \frac{\varepsilon}{2^{n+1}} .$$

Note that $a = \frac{\varepsilon}{2}$ and $r = \frac{1}{2}$, so the series is convergent. Therefore,

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

To prove the convergence of the p-series, we need some some preparation.

Definition 6.11. We say the the series $\sum_{n=1}^{\infty} x_n$ is a series with non-negative terms if $x_n \geq 0$ for all $n \in \mathbb{N}$.

Theorem 6.12. *Let $\sum_{n=1}^{\infty} x_n$ be series with non-negative terms. Then the series is convergent if and only if the sequence of partial sums is bounded from above.*

Proof. Note that, by the fact that the terms of the series are greater or equal to zero, the sequence of partial sums is monotonically increasing. Therefore, by Theorem 2.27, $\{S_k\}$ is convergent, if and only if is bounded from above.

□

Theorem 6.13. *(The Integral Test)*
Let $f : [1, +\infty) \to \mathbb{R}$ be a monotone decreasing function. Let us assume that $f(t) \geq 0$ for all $t \geq 1$. Define the sequence $\{x_n\}$ by $x_n = f(n)$ for all $n \in \mathbb{N}$. Then,

$$\sum_{n=1}^{\infty} x_n \ \textit{ is convergent}$$

if and only if

$$\int_1^{\infty} f(t)\, dt \ \textit{ is convergent}.$$

In case of convergence,

$$\sum_{n=1}^{\infty} x_n - x_1 \leq \int_1^{\infty} f(t)\, dt \leq \sum_{n=1}^{\infty} x_n\,.$$

Proof. First consider the interval $[1, 2]$. The function f is monotone decreasing, hence Riemann integrable on $[1, 2]$. Also, $f(2) \leq f(t) \leq f(1)$ for all $1 < t < 2$. Therefore by the Mean Value Theorem for integrals

$$x_2 = f(2) \leq \int_1^2 f(t)\, dt \leq f(1) = x_1\,.$$

We can repeat this argument for all intervals $[n, n+1]$ and add the inequalities. Therefore, for any $k \geq 2$ we have

$$x_2 + \cdots + x_k \leq \int_1^k f(t)\, dt \leq x_1 + \cdots + x_{k-1}\,.$$

Therefore,

$$S_k - x_1 \leq \int_1^k f(t)\, dt \leq S_{k-1}\,,\ \forall\, k \geq 2\,. \tag{6.1}$$

If the series $\sum_{n=1}^{\infty} x_n$ is convergent, then the sequence $\{S_k\}$ is bounded and hence the sequence

$$\left\{ \int_1^k f(t)\, dt \right\}$$

is bounded. Also, by the fact that $f(t) \geq 0$ for all $t \geq 1$, it is monotone increasing, and hence convergent. Moreover, the inequality

$$\int_1^k f(t)\, dt \leq \int_1^b f(t)\, dt \leq \int_1^{k+1} f(t)\, dt\,,\ \forall\, k < b < k+1$$

implies that

$$\lim_{\substack{k \to \infty \\ k \in \mathbb{N}}} \int_1^k f(t)\, dt = \lim_{\substack{b \to \infty \\ b \in \mathbb{R}}} \int_1^b f(t)\, dt\,.$$

In conclusion, $\int_1^\infty f(t)\,dt$ is convergent.

If $\int_1^\infty f(t)\,dt$ is convergent, then, by (6.1), the sequence

$$\left\{\int_1^k f(t)\,dt\right\}$$

is bounded, and hence the sequence of partial sums $\{S_k\}$ is bounded. Therefore, by Theorem 6.12, the series $\sum_{n=1}^\infty x_n$ is convergent.

□

Theorem 6.14. *(The p-Series Test) Let $p \in \mathbb{R}$. Then the p-series*

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

is convergent if and only if $p > 1$.

Proof. If $p \leq 0$, then the sequence $\left\{\frac{1}{n^p}\right\}$ does not converge to 0, hence the p-series is divergent by the Test for Divergence.

If $p > 0$, consider the monotone decreasing function, $f : [1, +\infty) \to \mathbb{R}$,

$$f(t) = \frac{1}{t^p}\,, \;\forall\, t \geq 1\,.$$

If $p = 1$, then

$$\int_1^\infty \frac{1}{t}\,dt = \lim_{b\to\infty}\int_1^b \frac{1}{t}\,dt = \lim_{b\to\infty} \ln b = +\infty\,.$$

Then, if $p \neq 1$ then

$$\int_1^\infty \frac{1}{t^p}\,dt = \lim_{b\to\infty}\int_1^b \frac{1}{t^p}\,dt = \lim_{b\to\infty}\left(\frac{1}{p-1} - \frac{1}{(p-1)b^{p-1}}\right)$$

is finite if and only of $p > 1$. In this case,

$$\int_1^\infty \frac{1}{t^p}\,dt = \frac{1}{p-1}\,.$$

By the Integral Test, Theorem 6.13, the p-series is convergent if and only if $p > 1$.

□

Remark 6.15. Note that, the Integral Test doesn't say that the improper integral and the series have the same value, only that they converge at the same time. Indeed,

$$\int_1^\infty \frac{1}{t^2}\,dt = 1\,,$$

b ut

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

This sum can be obtained using Fourier series, which is the material for a Functional Analysis course.

Definition 6.16. If $\sum_{n=1}^{\infty} x_n$ is convergent, then for every $k \in \mathbb{N}$, the expression

$$R_k = \sum_{n=1}^{\infty} x_n - S_k = \sum_{n=k+1}^{\infty} x_n$$

is called the remainder, or error, of the approximation of the infinite series by the partial sum S_k.

Theorem 6.17. *Under the assumptions of Theorem 6.13, if $\sum_{n=1}^{\infty} x_n$ is convergent, then*

$$\int_{k+1}^{\infty} f(t)\,dt \leq R_k \leq \int_{k}^{\infty} f(t)\,dt\,,\ \forall\, k \in \mathbb{N}\,.$$

Proof. The proof follows immediately from the following consequence of the Mean Value Theorem for integrals:

$$\int_{n+1}^{n+2} f(t)\,dt \leq x_{n+1} \leq \int_{n}^{n+1} f(t)\,dt\,,\ \forall\, n \in \mathbb{N}\,.$$

□

Example 6.18. The series $\sum_{n=1}^{\infty} \frac{1}{n^3}$ is convergent and its sum is an unknown number, called Apéry's constant after the French mathematician Roger Apéry, who proved in 1978 that it is an irrational number. Let's calculate the integral:

$$\int_{n}^{\infty} \frac{1}{t^3}\,dt = \frac{1}{2n^2}\,.$$

This shows that if, for $\sum_{n=1}^{\infty} \frac{1}{n^3}$, we want the remainder $R_n \leq \frac{1}{200}$, then we need $n \geq 10$. We have

$$\frac{1}{2 \cdot 11^2} \leq R_{10} \leq \frac{1}{2 \cdot 10^2}\,.$$

Adding $S_{10} \approx 1.19753$ to each of the three sides, gives

$$1.20166 \leq \sum_{n=1}^{\infty} \frac{1}{n^3} \leq 1.20253\,.$$

Exercises

Exercise 6.1. Show that $\sum_{n=1}^{\infty}(x_n - x_{n+1})$ converges if and only if $\{x_n\}$ converges.

Exercise 6.2. Let $|a| < 1$. Prove that $\sum_{n=1}^{\infty} \frac{1}{1+a^n}$ diverges.

Exercise 6.3. Let $\sum_{n=1}^{\infty} x_n$ be a series with non-negative terms. Show that if $\sum_{n=1}^{\infty} x_n$ converges, then $\sum_{n=1}^{\infty} \frac{\sqrt{x_n}}{n}$ converges.

Exercise 6.4. Let $\sum_{n=1}^{\infty} x_n$ be a series with non-negative terms. Show that if $\sum_{n=1}^{\infty} x_n$ converges, then $\sum_{n=1}^{\infty} \sqrt{x_n\, x_{n+1}}$ converges.

Exercise 6.5. Let $\sum_{n=1}^{\infty} x_n$ be a divergent series with positive terms and $\{S_k\}$ be the sequence of partial sums.

(1) Prove that $\sum_{n=1}^{\infty} \frac{x_n}{S_n}$ is divergent.

(2) Prove that $\sum_{n=1}^{\infty} \frac{x_n}{(S_n)^2}$ is convergent.

Exercise 6.6. Let $p > 1$. Show that:

$$\sum_{n=1}^{\infty} \frac{1}{n^p} < \frac{2^p}{2^p - 2}\,.$$

Exercise 6.7. Let $a_n \geq 0$ and $a_{n+1} \leq a_n$, for all $n \in \mathbb{N}$. Prove that if $\sum_{n=1}^{\infty} a_n$ is convergent, then $\lim_{n\to\infty} n\, a_n = 0$.

Exercise 6.8. For which numbers $p \in \mathbb{R}$ are the following series convergent?

$$(1) \quad \sum_{n=2}^{\infty} \frac{1}{n(\ln n^2)^p}\,.$$

$$(2) \quad \sum_{n=3}^{\infty} \frac{1}{n(\ln n)(\ln(\ln n))^p}\,.$$

Exercise 6.9. Evaluate $\sum_{n=1}^{\infty} \frac{1}{n^4}$ with an error less than 0.001.

Exercise 6.10. Prove that if $\sum_{n=1}^{\infty} x_n$ is convergent then

$$\lim_{k\to\infty} \frac{x_1 + 2x_2 + \cdots + kx_k}{k} = 0\,.$$

6.2 Comparison theorems for series with non-negative terms

Theorem 6.19. *(Term-by-term comparison test)*
Let $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ be two series with non-negative terms and assume that there exists $N \in \mathbb{N}$ such that $x_n \leq y_n$ for all $n > N$. Then:

(1) *If $\sum_{n=1}^{\infty} y_n$ is convergent, then $\sum_{n=1}^{\infty} x_n$ is convergent.*

(2) *If $\sum_{n=1}^{\infty} x_n$ is divergent, then $\sum_{n=1}^{\infty} y_n$ is divergent.*

Proof. Let us denote $S_k = \sum_{n=1}^{k} x_n$ and $T_k = \sum_{n=1}^{k} y_n$. Then, for $k > N$ we have

$$S_k \leq T_k + \sum_{n=1}^{N} (x_n - y_n) \,.$$

This inequality is the key for the proof.

(1) If $\sum_{n=1}^{\infty} y_n$ is convergent, then $\{T_k\}$ is bounded, which implies that $\{S_k\}$ is bounded, and hence $\sum_{n=1}^{\infty} x_n$ is convergent.

(2) If $\sum_{n=1}^{\infty} x_n$ is divergent, then $\{S_k\}$ is unbounded, which implies that $\{T_k\}$ is unbounded and hence If $\sum_{n=1}^{\infty} y_n$ is divergent.

□

Theorem 6.20. *(The limit comparison test)*
Let $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ be two series with positive terms and assume that

$$\lim_{n\to\infty} \frac{x_n}{y_n} = L \,.$$

(1) *If $\sum_{n=1}^{\infty} y_n$ is convergent and $0 \leq L < \infty$, then $\sum_{n=1}^{\infty} x_n$ is convergent.*

(2) *If $\sum_{n=1}^{\infty} y_n$ is divergent and $0 < L \leq +\infty$, then $\sum_{n=1}^{\infty} x_n$ is divergent.*

Proof. (1) If $L = 0$, then there exists $N \in \mathbb{N}$ such that

$$\frac{x_n}{y_n} \leq 1 \,,\ \forall\, n > N \,.$$

Then, $\sum_{n=1}^{\infty} x_n$ is convergent by Theorem 6.19.

If $0 < L < \infty$, then for $\varepsilon = \frac{L}{2}$ there exists $N \in \mathbb{N}$ such that

$$\frac{L}{2} \leq \frac{x_n}{y_n} \leq \frac{3L}{2} \,,\ \forall\, n > N \,. \tag{6.2}$$

Observing that, if $\sum_{n=1}^{\infty} y_n$ is convergent, then $\sum_{n=1}^{\infty} \frac{3L}{2} y_n$ is convergent, the conclusion follows again from Theorem 6.19.

(2) If $0 < L < \infty$ the proof follows again from inequality (6.2) and Theorem 6.19.

If $L = \infty$, then there exists $N \in \mathbb{N}$ such that

$$\frac{x_n}{y_n} \geq 1 \,, \ \forall \, n > N \,.$$

Therefore, the series $\sum_{n=1}^{\infty} x_n$ is divergent by Theorem 6.19.

□

Theorem 6.21. *(The Root Test)*
Let $\sum_{n=1}^{\infty} x_n$ be a series with non-negative terms and let

$$L = \limsup \sqrt[n]{x_n} \,.$$

Then, the following statements hold:

(1) *If $L < 1$, then $\sum_{n=1}^{\infty} x_n$ converges.*

(2) *If $L > 1$, then $\sum_{n=1}^{\infty} x_n$ diverges.*

(3) *If $L = 1$, then the test is inconclusive.*

Proof. (1) Assume that $L < 1$. Then, for $\varepsilon = \frac{1-L}{2}$, there exists $N \in \mathbb{N}$ such that

$$\sqrt[n]{x_n} < 1 - \frac{1-L}{2} = \frac{1+L}{2} = a < 1 \,, \ \forall \, n > N \,.$$

Hence, $x_n < a^n$ for all $n > N$, and by comparison with the convergent geometric series $\sum_{n=1}^{\infty} a^n$, we conclude that $\sum_{n=1}^{\infty} x_n$ converges.

(2) If $L > 1$, then there are infinitely many terms with $\sqrt[n]{x_n} > 1$, which means that $\{x_n\}$ cannot converge to 0. Hence, by the Test for Divergence, $\sum_{n=1}^{\infty} x_n$ diverges.

(3) If we consider the series $\sum_{n=1}^{\infty} \frac{1}{n}$, then $L = 1$ and the series is divergent. But, for $\sum_{n=1}^{\infty} \frac{1}{n^2}$, we still have $L = 1$, and the series is convergent.

□

Remark 6.22. As we can see from the proof, the Root Test is a comparison test with a geometric series, and therefore cannot be used for comparison with the p-series.

Lemma 6.23. *Let $\{x_n\}$ be a sequence with positive terms. Then*

$$\liminf \frac{x_{n+1}}{x_n} \leq \liminf \sqrt[n]{x_n} \leq \limsup \sqrt[n]{x_n} \leq \limsup \frac{x_{n+1}}{x_n} \,.$$

Proof. We will prove the first inequality, the second is obvious and we leave the third one as an exercise.

Let $\alpha = \liminf \frac{x_{n+1}}{x_n}$. If $\alpha = -\infty$, then the result is evident. If $-\infty < \alpha$, then for all $-\infty < \beta < \alpha$, there exists $N = N(\beta) \in \mathbb{N}$ such that

$$\beta < \frac{x_{n+1}}{x_n}\,,\ \forall\, n \geq N\,.$$

The inequalities

$$x_N < \frac{1}{\beta} x_{N+1} < \frac{1}{\beta^2} x_{N+2} < \ldots$$

imply that

$$\beta^k x_N < x_{N+k}\,,\ \forall\, k \in \mathbb{N}\,.$$

Therefore,

$$\sqrt[N+k]{\beta^k x_N} < \sqrt[N+k]{x_{N+k}}\,,\ \forall\, k \in \mathbb{N}\,.$$

Observing that

$$\lim_{k\to\infty} \sqrt[N+k]{\beta^k x_N} = \beta\,,$$

we obtain

$$\beta \leq \liminf \sqrt[n]{x_n}\,.$$

In conclusion, we obtained that for all $-\infty < \beta < \alpha$ we have $\beta \leq \liminf \sqrt[n]{x_n}$. This implies that

$$\alpha \leq \liminf \sqrt[n]{x_n}\,.$$

□

Theorem 6.24. *(The Ratio Test)*
Let $\sum_{n=1}^{\infty} x_n$ be a series with positive terms and let

$$l = \liminf \frac{x_{n+1}}{x_n} \ \text{ and } \ L = \limsup \frac{x_{n+1}}{x_n}\,.$$

Then, the following statements hold:

(1) *If $L < 1$, then $\sum_{n=1}^{\infty} x_n$ converges.*

(2) *If $l > 1$, then $\sum_{n=1}^{\infty} x_n$ diverges.*

(3) *If $l \leq 1 \leq L$, then the test is inconclusive.*

Proof. (1) If $L < 1$, then by Lemma 6.23 we have $\limsup \sqrt[n]{x_n} < 1$, and $\sum_{n=1}^{\infty} x_n$ converges by the Root test.

(2) If $l > 1$, then by Lemma 6.23 we have $\limsup \sqrt[n]{x_n} > 1$, and $\sum_{n=1}^{\infty} x_n$ diverges by the Root test.

(3) To show that in this case the test is inconclusive, we can use the same series as in case (3) of the Root Test.

□

Remark 6.25. The proof of the Ratio Test shows that the Root Test is stronger, which means that if the Ratio Test works, then the Root Test gives the same result. However, there are series for which we can use the Root Test, but the Ratio Test is inconclusive.

Consider the series

$$\sum_{n=1}^{\infty} \frac{1}{\Big(3 + (-1)^n\Big)^n} .$$

Notice that

$$\sqrt[2n]{x_{2n}} = \frac{1}{2} \quad \text{and} \quad \sqrt[2n+1]{x_{2n+1}} = \frac{1}{\sqrt{2}} .$$

Therefore, the series is convergent by the Root Test, because

$$\limsup \sqrt[n]{x_n} = \frac{1}{\sqrt{2}} < 1 .$$

However,

$$\liminf \frac{x_{n+1}}{x_n} = \lim_{n\to\infty} \frac{x_{2n}}{x_{2n+1}} = 0$$

and

$$\limsup \frac{x_{n+1}}{x_n} = \lim_{n\to\infty} \frac{x_{2n+1}}{x_{2n}} = +\infty ,$$

which shows that the Ratio Test is inconclusive.

Example 6.26. Consider the series $\sum_{n=1}^{\infty} \frac{1}{n!}$. The Ratio test shows that this series is convergent. Denote its sum with

$$e = \sum_{n=1}^{\infty} \frac{1}{n!} ,$$

which is the usual notation for the Euler's number $e \approx 2.7182$.

Theorem 6.27. *The Euler's number e is irrational.*

Proof. Let us denote the partial sum by

$$S_k = \sum_{n=1}^{k} \frac{1}{n!} , \ \forall\, k \in \mathbb{N} .$$

Observe that

$$\begin{aligned}
0 < e - S_k &= \frac{1}{(k+1)!} + \frac{1}{(k+2)!} + \cdots \\
&< \frac{1}{(k+1)!} + \frac{1}{(k+1)!(k+1)} + \cdots \\
&= \frac{1}{(k+1)!}\left(1 + \frac{1}{k+1} + \frac{1}{(k+1)^2} + \cdots\right) = \frac{1}{k\,k!}\,.
\end{aligned}$$

Hence,

$$0 < e - S_k < \frac{1}{k\,k!}\,,\ \forall\, k \in \mathbb{N}\,.$$

Assume now that e is rational, which means that $e = \frac{m}{k}$ for some $m, k \in \mathbb{N}$. Then,

$$0 < k!\left(\frac{m}{k} - S_k\right) < \frac{1}{k}\,.$$

We can check that the number $k!\left(\frac{m}{k} - S_k\right)$ is an integer, which is impossible. Therefore, the assumption that e is rational is false.

□

Exercises

Exercise 6.11. Determine the convergence or divergence of the following series:

(1) $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n}\,.$

(2) $\displaystyle\sum_{n=2}^{\infty} \frac{\ln n}{n^{1+\varepsilon}}\,,\ \varepsilon > 0\,.$

(3) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{(\ln n)^{\ln n}}\,.$

(4) $\displaystyle\sum_{n=1}^{\infty} \ln\left(\frac{n+1}{n}\right)\,.$

(5) $\displaystyle\sum_{n=1}^{\infty} \sin\frac{1}{n}\,.$

$$(6) \quad \sum_{n=2}^{\infty} \frac{1}{\sqrt{n^3 - n}}\,.$$

$$(7) \quad \sum_{n=1}^{\infty} \left(\sqrt[n]{n} - 1\right)^n\,.$$

Exercise 6.12. Let $\sum_{n=1}^{\infty} x_n$ be series with positive terms. Prove the following:

(1) If $\sum_{n=1}^{\infty} x_n$ is divergent then $\sum_{n=1}^{\infty} \frac{x_n}{1+x_n}$ is divergent.

(2) If $\sum_{n=1}^{\infty} x_n$ is convergent then $\sum_{n=1}^{\infty} \frac{x_n}{1+x_n}$ is convergent.

(3) If $\sum_{n=1}^{\infty} x_n$ is convergent, and $x_n \neq 1$ for all $n \in \mathbb{N}$, then $\sum_{n=1}^{\infty} \frac{x_n}{1-x_n}$ is convergent.

(4) If $\sum_{n=1}^{\infty} x_n$ is convergent, then $\sum_{n=1}^{\infty} \ln(1 + x_n)$ is convergent.

Exercise 6.13. Let $\sum_{n=1}^{\infty} x_n$ be a sequence with non-negative terms and assume that $x_{n+1} \leq x_n$ for all $n \in \mathbb{N}$. Prove that $\sum_{n=1}^{\infty} x_n$ converges if and only if $\sum_{k=0}^{\infty} 2^k x_{2^k}$ converges.

Exercise 6.14. Let $\sum_{n=1}^{\infty} x_n$ be a series with positive terms . Denote

$$r_n = \frac{x_{n+1}}{x_n}\,,$$

and assume that

$$\lim_{n\to\infty} r_n = L < 1\,.$$

Show that for $k \in \mathbb{N}$ sufficiently large

$$R_k \leq \frac{2x_{k+1}}{1-L}\,.$$

Exercise 6.15. Use the first 10 terms to approximate

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

What is the error in the approximation?

6.3 Absolute convergence

For series which have infinitely many positive and negative terms, the comparison test from the previous section cannot be directly applied. However, no matter what terms the series $\sum_{n=1}^{\infty} x_n$ has, the series $\sum_{n=1}^{\infty} |x_n|$ is a series with non-negative terms.

Definition 6.28. We say that the series $\sum_{n=1}^{\infty} x_n$ is absolutely convergent if the series $\sum_{n=1}^{\infty} |x_n|$ is convergent.

Remark 6.29. For series with non-negative terms, the convergence and absolute convergence are equivalent.

Theorem 6.30. *If a series is absolute convergent, then it is convergent.*

Proof. The main idea of the proof is that for all $k < l \in \mathbb{N}$ we have

$$\left| \sum_{k}^{l} x_n \right| \leq \sum_{k}^{l} |x_n|$$

which shows that if Theorem 6.4 is applied to $\sum_{n=1}^{\infty} |x_n|$, then it can be applied to $\sum_{n=1}^{\infty} x_n$, too.

$\square$

Remark 6.31. The converse of Theorem 6.30 is not true. We will prove later in this section that the alternating series $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ is convergent. However, it is not absolutely convergent, because the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent.

Definition 6.32. We say that a series is conditionally convergent if it is convergent, but not absolutely convergent.

Example 6.33. Consider the series

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

We use the inequality

$$\left| \frac{\sin n}{n^2} \right| \leq \frac{1}{n^2} ,$$

the convergence of the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ and the Term-by-Term Comparison Test to conclude that $\sum_{n=1}^{\infty} \frac{\sin n}{n^2}$ is absolutely convergent, hence convergent.

However, the same idea cannot be applied to the series

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

because $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent. To prove that it converges, we need the following results.

Theorem 6.34. *Consider a series $\sum_{n=1}^{\infty} a_n x_n$ with the following properties:*

(1) *The sequence of partial sums $\{A_k\}$ of $\sum_{n=1}^{\infty} a_n$ is bounded.*

(2) *$x_n \geq x_{n+1} \geq 0$ for all $n \in \mathbb{N}$.*

(3) *$\lim_{n\to\infty} x_n = 0$.*

Then, the series $\sum_{n=1}^{\infty} a_n x_n$ is convergent.

Proof. First, we need the following formula. For any $2 \leq k \leq l$ we have

$$\begin{aligned}
\sum_{n=k}^{l} a_n x_n &= \sum_{n=k}^{l} (A_n - A_{n-1}) x_n = \sum_{n=k}^{l} A_n x_n - \sum_{n=k}^{l} A_{n-1} x_n \\
&= \sum_{n=k}^{l} A_n x_n - \sum_{n=k-1}^{l-1} A_n x_{n+1} \\
&= \sum_{n=k}^{l-1} A_n (x_n - x_{n+1}) + A_l x_l - A_{k-1} x_k \,.
\end{aligned}$$

Note that in the case $k = l$, by convention, the sum $\sum_{n=k}^{k-1} A_n x_n = 0$.

By assumption (1), there exists $M > 0$ such that $|A_k| \leq M$ for all $k \in \mathbb{N}$.

Consider any $\varepsilon > 0$. By (2) and (3), choose $N = N(\varepsilon) \in \mathbb{N}$, such that $0 \leq x_n < \frac{\varepsilon}{2M}$ for all $n > N$. Therefore, for all $N < k \leq l$ we have

$$\begin{aligned}
\left| \sum_{n=k}^{l} a_n x_n \right| &= \left| \sum_{n=k}^{l-1} A_n (x_n - x_{n+1}) + A_l x_l - A_{k-1} x_k \right| \\
&\leq M \left| \sum_{n=k}^{l-1} (x_n - x_{n+1}) + x_l + x_k \right| \\
&= M 2 x_k < M \frac{\varepsilon}{2M} < \varepsilon \,.
\end{aligned}$$

By the Cauchy criterion stated in Theorem 6.4, we conclude that $\sum_{n=1}^{\infty} a_n x_n$ is convergent. $\square$

Example 6.35. Let us return to the series

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

Denoting $a_n = \sin n$ and $x_n = \frac{1}{n}$, in order to apply Theorem 6.34, we have to check that the partial sums

$$\sum_{n=1}^{k} \sin n$$

are bounded. We use the trigonometric formula

$$\cos(n+1) - \cos(n-1) = -2 \sin n \, \sin 1 ,$$

and do summation for n from 1 to k:

$$\cos(k+1) + \cos k - \cos 1 - 1 = -2 S_k \, \sin 1 .$$

Therefore,

$$S_k = \frac{1 + \cos 1 - \cos k - \cos(k+1)}{2 \sin 1} ,$$

which shows that we have the following bound:

$$|S_k| \leq \frac{2}{\sin 1} , \ \forall \, k \in \mathbb{N} .$$

Therefore, by Theorem 6.34, we conclude that

$$\sum_{n=1}^{\infty} \frac{\sin n}{n} \ \text{ is convergent} .$$

A corollary of the Theorem 6.34 is the following:

Theorem 6.36. *(The Alternating Series Test)*
Consider a series $\sum_{n=1}^{\infty} (-1)^n x_n$ with the following properties:

(1) $x_n \geq x_{n+1} \geq 0$ *for all* $n \in \mathbb{N}$.

(2) $\lim_{n\to\infty} x_n = 0$.

Then the series $\sum_{n=1}^{\infty} (-1)^n x_n$ is convergent.
Moreover, for the error estimation we have the following formula:

$$|R_k| \leq x_{k+1} , \ \forall \, k \in \mathbb{N} .$$

Proof. The proof of the convergence is straightforward if we set $a_n = (-1)^n$ in Theorem 6.34.

For the error estimation, we can write

$$|R_k| = |(-1)^{k+1}x_{k+1} + (-1)^{k+2}x_{k+2} + \cdots| = |x_{k+1} - (x_{k+2} - x_{k+3}) - \cdots| \leq x_{k+1}\,.$$

$\square$

Example 6.37. Let us consider again the series $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n}$. The alternating series test implies that it is convergent. As it was mentioned earlier, it is conditionally convergent, because it is not absolutely convergent.

Exercises

Exercise 6.16. Decide whether the following series converge absolutely, conditionally or diverge.

$$(1)\quad \sum_{n=1}^{\infty} \frac{(-1)^n n}{n+2}\,.$$

$$(2)\quad \sum_{n=2}^{\infty} \frac{(-1)^n}{\ln n}\,.$$

$$(3)\quad \sum_{n=1}^{\infty} (-1)^n \sin\frac{1}{n}\,.$$

$$(4)\quad \sum_{n=1}^{\infty} (-1)^n \frac{n^n}{n!}\,.$$

$$(5)\quad \sum_{n=1}^{\infty} \frac{n\cos(n\pi)}{2^n}\,.$$

Exercise 6.17. Let $\sum_{n=1}^{\infty} x_n$ be a series and define the terms

$$x_n^+ = \max\{x_n, 0\} \quad \text{and} \quad x_n^- = \max\{-x_n, 0\}\,.$$

Show that:

(1) $x_n = x_n^+ - x_n^-$ for all $n \in \mathbb{N}$.

(2) If $\sum_{n=1}^{\infty} x_n$ is absolutely convergent , then both of the series $\sum_{n=1}^{\infty} x_n^+$ and $\sum_{n=1}^{\infty} x_n^-$ are convergent.

(3) If $\sum_{n=1}^{\infty} x_n$ is conditionally convergent, then both of the series $\sum_{n=1}^{\infty} x_n^+$ and $\sum_{n=1}^{\infty} x_n^-$ are divergent.

Exercise 6.18. Suppose that the series $\sum_{n=1}^{\infty} x_n$ is conditionally convergent. Show that $\sum_{n=1}^{\infty} n^2 x_n$ is divergent. What can we say about $\sum_{n=1}^{\infty} n x_n$?

Exercise 6.19. Let $\sum_{n=1}^{\infty} x_n$ be a series and assume that $\sum_{n=1}^{\infty} a_n x_n$ converges for any bounded sequence $\{a_n\}$. Show that $\sum_{n=1}^{\infty} x_n$ is absolutely convergent.

Exercise 6.20. Suppose that $\{x_n\} \to 0$. Show that there is a subsequence $\{x_{n_k}\}$ such that $\sum_{k=1}^{\infty} x_{n_k}$ converges absolutely.

Exercise 6.21. Prove that if $\sum_{n=1}^{\infty} x_n$ converges absolutely, then $\sum_{n=1}^{\infty} (x_n)^2$ converges.

Exercise 6.22. Suppose that $\{x_n\} \to x \neq 0$ and $x_n \neq 0$ for all $n \in \mathbb{N}$. Prove that:

$$\sum_{n=1}^{\infty} |x_{n+1} - x_n| \text{ converges} \iff \sum_{n=1}^{\infty} \left| \frac{1}{x_{n+1}} - \frac{1}{x_n} \right| \text{ converges}\,.$$

Chapter 7

Sequences and Series of functions

7.1 Sequences of functions

7.1.1 Definitions and properties

Definition 7.1. Let X be a metric space and $A \subset X$. Consider a sequence of functions $\{f_n\}$, where $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$.

(1) We say that $\{f_n\}$ converges pointwise on A to a function $f : A \to \mathbb{R}$ if for all $x \in A$ and $\varepsilon > 0$ there exists $N = N(x, \varepsilon) > 0$ such that

$$\Big| f_n(x) - f(x) \Big| < \varepsilon \,,\ \forall\, n > N \,.$$

(2) We say that $\{f_n\}$ converges uniformly on A to a function $f : A \to \mathbb{R}$ if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that

$$\Big| f_n(x) - f(x) \Big| < \varepsilon \,,\ \forall\, n > N \,,\ x \in A \,.$$

(3) If $X = \mathbb{R}$, we say that $\{f_n\}$ converges almost everywhere on A to a function $f : A \to Y$, if it converges pointwise on A except a set of measure zero.

The following two propositions are immediate consequences of the definitions and their proofs are left as exercises.

Proposition 7.2. *Let X be a metric space and $A \subset X$. Consider a sequence of functions $\{f_n\}$, where $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$. If $\{f_n\}$ converges uniformly on A, then it converges pointwise on A.*

Proposition 7.3. *Let X be a metric space and $A \subset X$ be a finite set. Consider a sequence of functions $\{f_n\}$, where $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$. If $\{f_n\}$ converges pointwise on A, then it converges uniformly on A.*

The following theorem gives two characterizations of the uniform convergence.

Theorem 7.4. *Let X be a metric space and $A \subset X$. Consider a function $f : A \to \mathbb{R}$ and a sequence of functions $\{f_n\}$, where $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$. Let*

$$M_n = \sup_{x \in A} \Big| f_n(x) - f(x) \Big| \,, \ \forall\, n \in \mathbb{N} \,.$$

Then the following hold:

(1) *$\{f_n\}$ converges uniformly on A if and only if*

$$\lim_{n \to \infty} M_n = 0 \,.$$

(2) *$\{f_n\}$ converges uniformly to f on A if and only if for all $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that*

$$\Big| f_n(x) - f_m(x) \Big| < \varepsilon \,, \ \forall\, n, m > N \,, x \in A \,.$$

Proof. (1) Assume that $\{f_n\}$ converges uniformly to f on A and let $\varepsilon > 0$. Then, by definition, there exists $N = N(\varepsilon) > 0$ such that

$$\Big| f_n(x) - f(x) \Big| < \frac{\varepsilon}{2} \,, \ \forall\, n > N \,, \ x \in A \,.$$

The inequality is true for all $x \in A$, hence we can take the supremum on the left side and get

$$0 \leq M_n \leq \frac{\varepsilon}{2} < \varepsilon \,, \ \forall\, n > N \,,$$

which implies that $\{M_n\} \to 0$.

Assume now that $\{M_n\} \to 0$ and left $\varepsilon > 0$. Then, by definition, there exists $N = N(\varepsilon) > 0$ such that

$$0 \leq M_n < \varepsilon \,, \ \forall\, n > N \,.$$

The inequality is true for the supremum, which implies that it holds for each $x \in A$. Hence,

$$\Big| f_n(x) - f(x) \Big| < \varepsilon \,, \ \forall\, n > N \,, \ x \in A \,,$$

which gives the uniform convergence.

(2) Assume that $\{f_n\}$ converges uniformly to f. Then, by definition for all $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that

$$\Big| f_n(x) - f(x) \Big| < \frac{\varepsilon}{2}\,,\ \forall\, n > N\,,\ x \in A\,.$$

Therefore, for all $n > m > N$ and $x \in A$ we have

$$\Big| f_n(x) - f_m(x) \Big| \le \Big| f_n(x) - f(x) \Big| + \Big| f(x) - f_m(x) \Big| < \varepsilon\,.$$

Assume now that for all $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that

$$\Big| f_n(x) - f_m(x) \Big| < \varepsilon\,,\ \forall\, n, m > N\,, x \in A\,.$$

Fix any $x \in A$. From the above inequality follows that the sequence $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{R}$. Therefore, $\{f_n(x)\}$ is convergent and denote its limit by $f(x)$. Note that the function f is built by finding, for each x, the limit of the sequence $\{f_n(x)\}$.
In the inequality

$$\Big| f_n(x) - f_m(x) \Big| < \varepsilon\,,\ \forall\, n, m > N\,,\ x \in A$$

we let $m \to \infty$ and get

$$\Big| f_n(x) - f(x) \Big| \le \varepsilon\,,\ \forall\, n > N\,,\ x \in A\,.$$

This last inequality implies that $\{f_n\}$ converges uniformly to f.

□

Example 7.5. For each $n \in \mathbb{N}$ consider the functions $f_n : [0, 1] \to \mathbb{R}$ defined by $f_n(x) = x^n$. It is easy to check that if $0 \le x < 1$, then $\{f_n(x)\} \to 0$ and $\{f_n(1)\} \to 1$. Therefore, $\{f_n\}$ converges pointwise on $[0, 1]$ to the function

$$f(x) = \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } x = 1 \end{cases}\,.$$

To evaluate the uniform convergence let's write

$$f_n(x) - f(x) = \begin{cases} x^n & \text{if } 0 \le x < 1 \\ 0 & \text{if } x = 1 \end{cases}\,.$$

Fix any $n \in \mathbb{N}$ and calculate

$$M_n = \sup_{0 \le x < 1} x^n = 1\,.$$

This shows that $\{M_n\}$ is the constant sequence 1, which means that it cannot converge to 0. Hence, the convergence is not uniform.

Can we change the domain $[0,1]$ to obtain uniform convergence. Yes, if we fix $0 < \delta < 1$ and $f_n : [0, 1-\delta] \to \mathbb{R}$, then

$$M_n = \sup_{0 \le x \le 1-\delta} x^n = (1-\delta)^n .$$

In this case $\{M_n\} \to 0$, which shows that we have uniform convergence on the interval $[0, 1-\delta]$ to the constant zero function.

Example 7.6. Consider $f_n : [0,1] \to \mathbb{R}$ defined by

$$f_n(x) = \frac{x}{1+nx^2}\,,\ \forall\, 0 \le x \le 1 .$$

Fix any $x \in [0,1]$. The limit

$$\lim_{n\to\infty} \frac{x}{1+nx^2} = 0$$

shows that $\{f_n\}$ converges pointwise to the constant 0 function.

Fix $n \in \mathbb{N}$. We want to find the maximum of f_n, so let's calculate

$$f_n'(x) = \frac{1-nx^2}{(1+nx^2)^2} .$$

The only critical point inside $[0,1]$ is $x = \frac{1}{\sqrt{n}}$. We compare the values of f_n at the endpoints and at the critical points:

$$f_n(0) = 0\,,\ \ f_n\left(\frac{1}{\sqrt{n}}\right) = \frac{1}{2\sqrt{n}}\,,\ f_n(1) = \frac{1}{1+n} .$$

Therefore,

$$M_n = \sup_{0\le x\le 1} f_n(x) = \frac{1}{2\sqrt{n}}\,,$$

which shows that $\{M_n\} \to 0$. We conclude that $\{f_n\}$ converges uniformly to the constant zero function on $[0,1]$.

The next theorem gives another methods to prove uniform convergence. We will state it for a decreasing sequence of functions, but it is also true for an increasing sequence.

Theorem 7.7. *(Dini's Theorem)*
Suppose that X is a metric space, $K \subset X$ is a compact subset and $f_n : K \to \mathbb{R}$ for all $n \in \mathbb{N}$. We assume that

(1) *For all $n \in \mathbb{N}$, f_n is continuous on K.*

(2) *$\{f_n\}$ converges pointwise to a continuous function f on K.*

(3) *$f_n(x) \ge f_{n+1}(x)$ for all $x \in K$ and $n \in \mathbb{N}$.*

Then $\{f_n\}$ *converges uniformly to* f *on* K.

Proof. Define $g_n : K \to \mathbb{R}$ by $g_n(x) = f_n(x) - f(x)$. Then each g_n is continuous on K, the sequence $\{g_n\}$ converges pointwise to 0 and also satisfies assumption (3).

Let $\varepsilon > 0$ and for each $n \in \mathbb{N}$ define

$$K_n = \{x \in K \mid g_n(x) \geq \varepsilon\}.$$

The sets K_n are closed, because of the continuity of g_n, and hence compact, as closed subsets of a compact set. The property (3) implies that $K_{n+1} \subset K_n$ for all $n \in \mathbb{N}$.

If each K_n were nonempty, then by Corollary 2.61, their intersection would be nonempty. However, for any $x \in K$ there exists $n \in \mathbb{N}$ such that $g_n(x) < \varepsilon$, since $\{g_n(x)\} \to 0$. This means that $x \notin K_n$, so the intersection of the sets K_n is empty.

Therefore, there must exist $N = N(\varepsilon) \in \mathbb{N}$ such that $K_N = \emptyset$. Then $K_n = \emptyset$ for all $n > N$, from which it follows

$$0 \leq g_n(x) < \varepsilon, \ \forall \ n > N, \ x \in K.$$

This proves that $\{g_n\}$ converges uniformly to 0 on K and hence $\{f_n\}$ converges uniformly to f on K.

□

Remark 7.8. The compactness of K is necessary. For example, consider $f_n : (0,1) \to \mathbb{R}$ defined by

$$f_n(x) = \frac{1}{nx+1}, \ \forall \ 0 < x < 1.$$

The sequence $\{f_n\}$ converges pointwise to 0, because for every $0 < x < 1$ we have

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

Moreover, the sequence $\{f_n\}$ satisfies (1)-(3) of Theorem 7.7, but the convergence is not uniform, because

$$M_n = \sup_{0<x<1} \frac{1}{nx+1} = 1, \ \forall \ n \in \mathbb{N}.$$

7.1.2 Properties preserved by uniform convergence

The following theorems will describe the conditions under which continuity, differentiability and integrability are transmitted to the limit function.

Lemma 7.9. *Let* X *be a metric space,* $A \subset X$ *and* x_0 *a limit point of* A. *Assume that* $f, f_n : A \setminus \{x_0\} \to \mathbb{R}$ *for all* $n \in \mathbb{N}$ *and*

$$\lim_{x\to x_0} f_n(x) = y_n, \ \forall \ n \in \mathbb{N}. \tag{7.1}$$

If $\{f_n\}$ *converges uniformly to* f *on* $A \setminus \{x_0\}$, *then*

$$\lim_{x\to x_0} f(x) = \lim_{n\to\infty} y_n. \tag{7.2}$$

Proof. Let $\varepsilon > 0$. By the uniform convergence, there exists $N = N(\varepsilon) > 0$ such that

$$\Big| f_n(x) - f_m(x) \Big| < \frac{\varepsilon}{3}\,,\ \forall\ n, m > N,\ x \in A\,.$$

Letting $x \to x_0$ gives

$$\Big| y_n - y_m \Big| < \frac{\varepsilon}{3}\,,\ \forall\ n, m > N\,,$$

which means that $\{y_n\}$ is a Cauchy sequence in $\mathbb{R}$, hence convergent to a $y \in \mathbb{R}$.

Letting $m \to \infty$ gives

$$\Big| y_n - y \Big| \leq \frac{\varepsilon}{3}\,,\ \forall\ n > N\,.$$

Fix $n_0 > N$. By (7.1), there exists $\delta = \delta(\varepsilon) > 0$ such that for all $x \in A$ with $0 < d(x, x_0) < \delta$ we have

$$\Big| f_{n_0}(x) - y_{n_0} \Big| < \frac{\varepsilon}{3}\,.$$

Therefore, for any $x \in A$ with $0 < d(x, x_0) < \delta$ holds

$$\Big| f(x) - y \Big| \leq \Big| f(x) - f_{n_0}(x) \Big| + \Big| f_{n_0}(x) - y_{n_0} \Big| + \Big| y_{n_0} - y \Big| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon\,.$$

Therefore,

$$\lim_{x \to x_0} f(x) = y\,.$$

□

Remark 7.10. Note that (7.2) is equivalent to

$$\lim_{n \to \infty} \lim_{x \to x_0} f_n(x) = \lim_{x \to x_0} \lim_{n \to \infty} f_n(x)\,.$$

Theorem 7.11. *Let X be a metric space, $A \subset X$ and $f_n : A \to \mathbb{R}$ be continuous on A for all $n \in \mathbb{N}$. If $\{f_n\}$ converges uniformly to $f : A \to \mathbb{R}$, then f is continuous on A.*

Proof. Let us fix any $x_0 \in A$. If x_0 is an isolated point of A, then f is continuous at x_0. If x_0 is a limit point of A, then we apply Lemma 7.9 and get

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \lim_{x \to x_0} f_n(x) = \lim_{n \to \infty} f_n(x_0) = f(x_0)\,.$$

This shows that f is continuous at x_0 and because x_0 was chosen arbitrarily, f is continuous on A.

□

Remark 7.12. Without uniform convergence the theorem is not true. See Example 7.5, where $f_n(x) = x^n$ is continuous on $[0,1]$, but the limit function

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

is not continuous on $[0,1]$.

However, there are cases when the convergence is not uniform, still the limit function is continuous. For an example, consider the $f_n(x) = x^n$ functions defined on $[0,1)$.

Theorem 7.13. *Let $\alpha : [a,b] \to \mathbb{R}$ be monotone increasing and $f, f_n : [a,b] \to \mathbb{R}$ for all $n \in \mathbb{N}$. Suppose that:*

(1) *For all $n \in \mathbb{N}$, f_n is Riemann-Stieltjes integrable with respect to α on $[a,b]$.*

(2) *$\{f_n\}$ converges uniformly to f on $[a,b]$.*

Then, f is Riemann-Stieltjes integrable with respect to α on $[a,b]$ and

$$\int_a^b f(x)\, d\alpha(x) = \lim_{n\to\infty} \int_a^b f_n(x)\, d\alpha(x)\,.$$

Proof. For every $n \in \mathbb{N}$ let

$$M_n = \sup_{x\in[a,b]} \Big| f_n(x) - f(x) \Big|\,.$$

The uniform convergence of $\{f_n\}$ implies that $\{M_n\} \to 0$.

Using the Riemann-Stieltjes integrability of f_n, the inequality

$$f_n(x) - M_n \leq f(x) \leq f_n(x) + M_n \tag{7.3}$$

implies that

$$\int_a^b \Big(f_n(x) - M_n\Big)\, d\alpha(x) \leq \underline{S}(f,\alpha) \leq \overline{S}(f,\alpha) \leq \int_a^b \Big(f_n(x) + M_n\Big)\, d\alpha(x)\,.$$

Therefore,

$$0 \leq \overline{S}(f,\alpha) - \underline{S}(f,\alpha) \leq \int_a^b 2M_n\, d\alpha(x) = 2M_n(\alpha(b) - \alpha(a))\,.$$

Letting $n \to \infty$ gives $\overline{S}(f,\alpha) = \underline{S}(f,\alpha)$, which implies the Riemann-Stieltjes integrability of f with respect to α.

Now we can go back to (7.3) and, by integration, get

$$\int_a^b f_n(x)\, d\alpha(x) - M_n(\alpha(b) - \alpha(a)) \leq \int_a^b f(x)\, d\alpha(x) \leq \int_a^b f_n(x)\, d\alpha(x) - M_n(\alpha(b) - \alpha(a))\,.$$

This implies that

$$\left|\int_a^b f(x)\,d\alpha(x) - \int_a^b f_n(x)\,d\alpha(x)\right| \le M_n(\alpha(b)-\alpha(a))\,,$$

which, taking into account that $\{M_n\} \to 0$, leads to

$$\int_a^b f(x)\,d\alpha(x) = \lim_{n\to\infty}\int_a^b f_n(x)\,d\alpha(x)\,.$$

$\square$

Example 7.14. The uniform convergence is a necessary assumption in Theorem 7.13. For every $n \in \mathbb{N}$ consider $f_n : [0,1] \to \mathbb{R}$ defined by

$$f_n(x) = nx(1-x^2)^n\,.$$

For every $x \in [0,1]$ we have

$$\lim_{n\to\infty} nx(1-x^2)^n = 0\,,$$

hence $\{f_n\}$ converges pointwise to 0. Integration by substitution leads to

$$\int_0^1 nx(1-x^2)^n\,dx = \frac{n}{2n+2}\,,$$

so

$$\lim_{n\to\infty}\int_0^1 f_n(x)\,dx = \frac{1}{2} \neq 0 = \int_0^1 0dx\,.$$

We can argue that the contra-postive form of Theorem 7.7 leads to the conclusion that the convergence is not uniform. However, some not so simple calculations lead to $M_n = \frac{1}{\sqrt{2}}$, which shows also that we don't have uniform convergence.

The next Theorem will be extensively used for power series later this chapter.

Theorem 7.15. *Let $\{f_n\}$ be a sequence of functions defined on $[a,b]$ with the following properties:*

(1) *Each f_n is differentiable on $[a,b]$.*

(2) *$\{f_n'\}$ converges uniformly to a function g on $[a,b]$.*

(3) *There exists $x_0 \in [a,b]$ such that $\{f_n(x_0)\}$ is convergent.*

Then, $\{f_n\}$ converges uniformly to a function f on $[a,b]$. Moreover, f is differentiable on $[a,b]$ and $f' = g = \lim_{n\to\infty} f_n'$.

Proof. Let $\varepsilon > 0$. By assumption (3) let us choose $N_1 = N_1(\varepsilon) > 0$ such that for all $n, m > N_1$ we have

$$\Big| f_n(x_0) - f_m(x_0) \Big| < \frac{\varepsilon}{2} \,. \tag{7.4}$$

Assumption (2) implies that there exists $N_2 = N_2(\varepsilon) > 0$ such that for all $n, m > N_2$ we have

$$\Big| f_n'(x) - f_m'(x) \Big| < \frac{\varepsilon}{2(b-a)} \,, \ \forall \ \ a \le x \le b \,. \tag{7.5}$$

Let $N = \max\{N_1, N_2\}$. For $n, m > N$ and $f_n - f_m$ we apply Theorem 4.14 and (7.5) to get

$$\Big| f_n(x) - f_m(x) - \big(f_n(y) - f_m(y)\big) \Big| < \frac{\varepsilon}{2(b-a)} \big| x - y \big| < \frac{\varepsilon}{2} \,, \forall \, x, y \in [a, b] \,. \tag{7.6}$$

For any $x \in [a, b]$ we write

$$\Big| f_n(x) - f_m(x) \Big| \le \Big| f_n(x) - f_m(x) - \big(f_n(x_0) - f_m(x_0)\big) \Big| + \Big| f_n(x_0) - f_m(x_0) \Big| \,,$$

and by using (7.4) and (7.6) we obtain

$$\Big| f_n(x) - f_m(x) \Big| < \varepsilon \,, \ \forall \ \ x \in [a, b] \,, \ n, m > N \,.$$

Therefore, by Theorem 7.4, $\{f_n\}$ converges uniformly on $[a, b]$ and let us denote

$$f(x) = \lim_{n \to \infty} f_n(x) \,, \ \forall \ \ x \in [a, b] \,.$$

Fix any $x \in [a, b]$ and define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x} \,, \ \ \phi(t) = \frac{f(t) - f(x)}{t - x} \,, \ \forall \ \ t \in [a, b] \setminus \{x\} \,.$$

Inequality (7.6) implies that $\{\phi_n\}$ converges uniformly on $[a, b] \setminus \{x\}$. Also,

$$\lim_{n \to \infty} \phi_n(t) = \phi(t)$$

and

$$\lim_{t \to x} \phi_n(t) = f_n'(x) \,.$$

We can apply now Lemma 7.9 to obtain that $f'(x)$ exists and

$$f'(x) = \lim_{t \to x} \phi(t) = \lim_{n \to \infty} f_n'(x) = g(x) \,.$$

$\square$

Example 7.16. This example show that the uniform convergence of $\{f_n'\}$ is essential, even if $\{f_n\}$ converges uniformly. Let

$$f_n(x) = \frac{\sin(nx)}{\sqrt{n}}\,,\ x \in \mathbb{R}\,,\ n \in \mathbb{N}\,.$$

We see that

$$\lim_{n\to\infty} f_n(x) = 0\,,\ \forall\ \ x \in \mathbb{R}\,.$$

Moreover,

$$M_n = \sup_{x\in\mathbb{R}} \Big| f_n(x) \Big| = \frac{1}{\sqrt{n}}$$

show that $\{f_n\}$ converges uniformly to the constant function $f = 0$. However,

$$f_n'(x) = \sqrt{n}\cos(nx)\,,$$

which shows that $\{f_n'\}$ doesn't converge. For example, $f_n'(0) = \sqrt{n}$ and $f_n(\pi) = (-1)^n\sqrt{n}$.

7.1.3 Equicontinuity

The notion of equicontinuity is related to compactness. Let us remind that in Theorem 2.62 we gave several characterization of compactness in $\mathbb{R}^n$. For example, a set is compact if and only if it is closed and bounded or if every sequence has a convergent subsequence.

These characterizations are not true when we dealing with infinite dimensional spaces. Among them is the set of continuous functions defined on an interval $[a, b]$.

The compactness is essential in proving that many equations of applied mathematics, like differential or partial differential equations, have solutions. These equations are defined on infinite dimensional spaces.

Definition 7.17. Let X be a metric space, $K \subset X$ be a compact set. We define

$$C(K) = \Big\{ f : K \to \mathbb{R} \ \Big|\ f \ \text{ is continuous on } K \Big\}\,.$$

By the fact that the sums and multiplications by numbers preserve continuity, the $C(K)$ becomes a vector space, which is infinite dimensional. On $C(X)$ we can introduce the norm

$$||f|| = \max_{x\in K}\Big|f(x)\Big|$$

and the distance

$$d(f, g) = ||f - g||\,.$$

Therefore, we can refer to $C(K)$ as a normed space or a metric space.

Remark 7.18. Note that in $C(K)$ the convergence with respect to the metric, just introduced earlier, is the uniform convergence. If we say $\{f_n\} \to f$ in $C(K)$, it means that $\{f_n\}$ converges uniformly to f on K.

Example 7.19. Consider $C([0,1])$ and $A = B_1[0]$, which is the closed ball centered at 0 and radius 1. This is a closed and bounded set, which is not compact. Let

$$f_n(x) = \frac{x^2}{x^2 + (1-nx)^2}\,,\ \forall\ 0 \le x \le 1\,,\ n \in \mathbb{N}\,.$$

It is evident that

$$0 \le f_n(x) \le 1\,,\ \forall\ 0 \le x \le 1\,,$$

so $||f_n|| \le 1$, which means that $f_n \in B_1[0]$. Note that $\{f_n\}$ converges pointwise to 0, but because of

$$f_n\left(\frac{1}{n}\right) = 1\,,\ \forall\ n \in \mathbb{N}\,,$$

no subsequence of $\{f_n\}$ can converge uniformly. This shows that $B_1[0]$ is not compact.

Definition 7.20. Let X be a metric space and $A \subset X$. Let $\Phi \subset C(A)$.

(1) We say that Φ is pointwise bounded if for all $x \in A$ there exists $M(x) > 0$ such that

$$\Big|f(x)\Big| \le M(x)\,,\ \forall\ f \in \Phi\,.$$

(2) We say that Φ is uniformly bounded if there exists $M > 0$ such that

$$\Big|f(x)\Big| \le M\,,\ \forall\ x \in A\,,\ f \in \Phi\,.$$

(3) We say that Φ is equicontinuous if for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all $x, y \in A$ with $d(x,y) < \delta$ we have

$$\Big|f(x) - f(y)\Big| < \varepsilon\,,\ \forall\, f \in \Phi\,.$$

Remark 7.21. We can see from definition that each function from an equicontinuous set is uniformly continuous and the δ is the same for each function. So, we can think about an equicontinuous set of functions as a uniformly-uniformly continuous set.
It is evident that the above definition apply to sequences, too. In this case $\Phi = \{f_n \mid n \in \mathbb{N}\}$.

The proof of the following proposition is left as an exercise.

Proposition 7.22. *Let X be a metric space and $\Phi \subset C(A)$ be a finite set. If every $f \in \Phi$ is uniformly continuous, then Φ is equicontinuous.*

Proposition 7.23. *Let $f_n \in C([a,b])$ for all $n \in \mathbb{N}$. Assume that each f_n is differentiable on $[a,b]$ and $\{f_n'\}$ is uniformly bounded. Then, $\{f_n\}$ is equicontinuous.*

Proof. The uniform boundedness of $\{f_n'\}$ implies that there exists $M > 0$ such that

$$\left|f_n'(x)\right| \leq M \,,\ \forall\ \ x \in [a,b],\ n \in \mathbb{N}\,.$$

Theorem 4.14 implies that for all $a \leq x < y \leq b$ there exists $x < t < y$ such that

$$\left|f_n(y) - f_n(x)\right| = \left|f_n'(t)\right|(y-x) \leq M(y-x)\,.$$

This shows that in the definition of equicontinuity, for $\varepsilon > 0$ we can choose $\delta = \frac{\varepsilon}{M}$.

□

Theorem 7.24. *Let X be a metric space and $K \subset X$ be a compact set. If $f_n \in C(K)$ for all $n \in \mathbb{N}$ and $\{f_n\}$ converges uniformly, then $\{f_n\}$ is equicontinuous.*

Proof. Let $\varepsilon > 0$. Since $\{f_n\}$ is uniformly convergent, then there exists $N = N(\varepsilon) \in \mathbb{N}$ such that for all $n, m > N$ we have

$$\left|f_n(x) - f_m(x)\right| < \frac{\varepsilon}{3}\,,\ \forall\ \ x \in K\,.$$

From the fact that continuous functions defined on compact sets are uniformly continuous, for all $1 \leq i \leq N+1$ there exists $\delta_i = \delta_i(\varepsilon) > 0$ such that if $x, y \in K$ and $d(x,y) < \delta_i$, then

$$\left|f_i(x) - f_i(y)\right| < \frac{\varepsilon}{3}\,.$$

Let $\delta = \min\{\delta_1, \ldots, \delta_{N+1}\}$. If $n > N+1$ and $d(x,y) < \delta$, then

$$\left|f_n(x) - f_n(y)\right| \leq \left|f_n(x) - f_{N+1}(x)\right| + \left|f_{N+1}(x) - f_{N+1}(y)\right| + \left|f_{N+1}(y) - f_n(y)\right| < \varepsilon\,.$$

Therefore, $\{f_n\}$ is equicontinuous.

□

Lemma 7.25. *Let X be a metric space, $A \subset X$ be a countable set and $f_n : A \to \mathbb{R}$ for all $n \in \mathbb{N}$. It $\{f_n\}$ is pointwise bounded on A, then $\{f_n\}$ contains a subsequence converging pointwise on A.*

Proof. Let $A = \{x_i \,\big|\, i \in \mathbb{N}\}$. The sequence $\{f_n(x_1)\}$ is bounded in $\mathbb{R}$, so it has a convergent subsequence. Let us denote it by $f_{1,k}$.

The sequence $\{f_{1,k}(x_2)\}$ is bounded, too, so it contains a convergent subsequence. Denote it by $\{f_{2,k}\}$.

Continue this process until you go over all elements of A. The subsequence, which is pointwise convergent on A is selected as:

$$f_{1,1},\ f_{2,2}, \ldots, f_{k,k},\ \ldots .$$

This sequence is a subsequence of all $\{f_{m,k}\}$ subsequences, hence converges at every point of A.

$\square$

Theorem 7.26. *(The Arzelà-Ascoli Theorem)*
Let X be a metric space, $K \subset X$ be a compact set and $f_n \in C(K)$ for all $n \in \mathbb{N}$. If $\{f_n\}$ is pointwise bounded and equicontinuous, then $\{f_n\}$ contains a uniformly convergent subsequence.

Proof. Let $\varepsilon > 0$. Since $\{f_n\}$ is equicontinuous, there exist $\delta = \delta(\varepsilon) > 0$ such that for all $x, y \in K$ with $d(x,y) < \delta$ and for all $n \in \mathbb{N}$ we have

$$\Big|f_n(x) - f_n(y)\Big| < \frac{\varepsilon}{3}\,.$$

From the compactness of K it follows that there is a countable set $A \subset K$, which is dense in K.

Since A is countable, by Lemma 7.25, there exists a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for any $x \in A$.

The density of A implies that

$$K \subset \bigcup_{x \in A} B_\delta(x)$$

is an open cover of K. Hence, we can select a finite subcover:

$$K \subset \bigcup_{i=1}^{m} B_\delta(x_i)\,,$$

where $x_i \in A$ for all $1 \le i \le m$.

By Proposition 7.3 the subsequence $\{f_{n_k}\}$ converges uniformly of $\{x_1, \ldots x_m\}$, so there exists $K = K(\varepsilon) > 0$ such that

$$\Big|f_{n_k}(x_i) - f_{n_l}(x_i)\Big| < \frac{\varepsilon}{3}\,,\ \ \forall\ \ k,l > K\,,\ 1 \le i \le m\,.$$

Let us fix an arbitrary $x \in K$. Then there exists $1 \le i \le m$ such that $x \in B_\delta(x_i)$. For any $k, l > K$ we have

$$\Big|f_{n_k}(x) - f_{n_l}(x)\Big| \le \Big|f_{n_k}(x) - f_{n_k}(x_i)\Big| + \Big|f_{n_k}(x_i) - f_{n_l}(x_i)\Big| + \Big|f_{n_l}(x_i) - f_{n_l}(x)\Big| < \varepsilon\,.$$

Since this last inequality is true for any $x \in K$, Theorem 7.4 implies that $\{f_{n_k}\}$ is uniformly convergent on K.

□

The Arzelà-Ascoli Theorem implies the following characterizations of compact sets in spaces of continuous functions.

Theorem 7.27. *Let X be a metric space, $K \subset X$ be a compact set and $\Phi \subset C(K)$. Then, Φ is compact if and only if it is closed, bounded and equicontinuous.*

7.1.4 The Stone-Weierstrass Theorem

The following theorem is one of the most used in applied mathematics. Polynomials are those functions which can be easily coded in computer programs, so any function that can be approximated by polynomials can be handled as well.

Theorem 7.28. *(The Stone-Weierstrass Theorem)*
Let $f \in C([a,b])$. Then, there exists a sequence $\{P_n\}$ of polynomials such that $\{P_n\}$ converges uniformly to f on $[a,b]$.

Proof. Without loss of generality we can assume that $[a,b] = [0,1]$, because the function $\phi : [0,1] \to [a,b]$, $\phi(x) = a + x(b-a)$, when composed with continuous functions and polynomials, preserves these properties.

Also, we can assume $f(0) = f(1)$, because otherwise we replace f by

$$\hat{f}(x) = f(x) - f(0) - x\big(f(1) - f(0)\big) .$$

So, we assume $f \in C([0,1])$ and $f(0) = f(1)$. Moreover, we extend f outside of $[0,1]$ by constantly 0 values. In this way $f \in C(\mathbb{R})$ and $f(x) = 0$ if $x \leq 0$ or $1 \leq x$.

Define

$$c_n = \frac{1}{\int_{-1}^{1} (1-x^2)^n \, dx}$$

and

$$Q_n(x) = c_n (1-x^2)^n .$$

With these choices we have

$$\int_{-1}^{1} Q_n(x) \, dx = 1 .$$

Although, there are formulas for c_n, we need some simpler estimates:

$$\int_{-1}^{1} (1-x^2)^n \, dx = 2 \int_{0}^{1} (1-x^2)^n \, dx \geq 2 \int_{0}^{\frac{1}{\sqrt{n}}} (1-x^2)^n \, dx$$

$$\geq 2 \int_0^{\frac{1}{\sqrt{n}}} (1 - nx^2)\, dx = \frac{4}{3\sqrt{n}} > \frac{1}{\sqrt{n}}\,.$$

Hence,

$$c_n < \sqrt{n}\,,$$

and therefore for any $0 < \delta < 1$ we have

$$Q_n(x) \leq \sqrt{n}(1 - \delta^2)^n\,, \text{ if } \delta \leq |x| \leq 1\,.$$

Because of

$$\lim_{n\to\infty} \sqrt{n}(1 - \delta^2)^n = 0\,,$$

we obtain that $\{Q_n\}$ converges uniformly to 0 on $[-1, -\delta] \cup [\delta, 1]$. Define

$$P_n(x) = \int_{-1}^{1} f(x+t) Q_n(t)\, dt\,,\ \forall\, 0 \leq x \leq 1\,.$$

Note that because f has zero values outside of $[0, 1]$, we can write

$$P_n(x) = \int_{-x}^{1-x} f(x+t) Q_n(t)\, dt\,,\ \forall\, 0 \leq x \leq 1\,.$$

To realize that P_n is a polynomial, we can use a change of variables $\tau = x + t$ to get

$$P_n(x) = \int_0^1 f(\tau) Q_n(\tau - x)\, d\tau\,.$$

As a continuous function initially defined [0,1] and extended to $\mathbb{R}$ by constant 0 values, f is uniformly continuous on $\mathbb{R}$. Therefore, for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that if $\big|x - y\big| < \delta$, then $\big|f(x) - f(y)\big| < \frac{\varepsilon}{2}$. Let

$$M = \sup_{x\in\mathbb{R}} \big|f(x)\big|\,.$$

Then,

$$\begin{aligned}
\Big|P_n(x) - f(x)\Big| &= \Big|\int_{-1}^{1} f(x+t) Q_n(t)\, dt - f(x) \int_{-1}^{1} Q_n(t)\, dt\Big| \\
&= \Big|\int_{-1}^{1} \big(f(x+t) - f(x)\big) Q_n(t)\, dt\Big| \\
&\leq \int_{-1}^{1} \big|f(x+t) - f(x)\big| Q_n(t)\, dt \\
&\leq 2M \int_{-1}^{-\delta} Q_n(t)\, dt + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} Q_n(t)\, dt + 2M \int_{\delta}^{1} Q_n(t)\, dt
\end{aligned}$$

$$\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2}\,.$$

As $\{\sqrt{n}(1-\delta^2)^n\} \to 0$, we can choose $N = N(\varepsilon) > 0$ such that for all $n > N$ we have

$$\left|P_n(x) - f(x)\right| < \varepsilon\,,\ \forall\, x \in [0,1]\,,$$

which means that $\{P_n\}$ converges uniformly to f on $[0,1]$.

$\square$

Corollary 7.29. *The set of polynomials is dense in $C([a,b])$.*

Exercises

Exercise 7.1. What can we say about the pointwise and uniform convergences of the following sequences of functions:

$$(a)\quad f_n(x) = \frac{\sin(nx)}{nx}\,,\ x \in [1,+\infty).$$

$$(b)\quad f_n(x) = x\,e^{-nx}\,,\ x \in [0,+\infty).$$

$$(c)\quad f_n(x) = x\,e^{-nx} + \frac{n+1}{n}\sin x\,,\ x \in [0,+\infty).$$

$$(d)\quad f_n(x) = \frac{x}{1+nx}\,,\ x \in (0,1).$$

$$(e)\quad f_n(x) = \frac{nx}{1+n^2x^2}\,,\ x \in [0,+\infty).$$

$$(f)\quad f_n(x) = \frac{\ln(1+nx)}{n}\,,\ x \in [0,1].$$

$$(g)\quad f_n(x) = x(1-x)^n\,,\ x \in [0,1].$$

$$(h)\quad f_n(x) = nx(1-x)^n\,,\ x \in [0,1].$$

Exercise 7.2. For $n \geq 2$ define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} n^2x & \text{if } 0 \leq x \leq \frac{1}{n} \\ 2n - n^2x & \text{if } \frac{1}{n} < x \leq \frac{2}{n} \\ 0 & \text{if } \frac{2}{n} < x \leq 1 \end{cases}\,.$$

(1) Plot f_n.

(2) Show that $\{f_n\}$ converges pointwise to 0.

(3) Is the convergence uniform?

(4) Is $\{f_n\}$ uniformly bounded?

(5) Is $\{f_n\}$ equicontinuous?

Exercise 7.3. Let $f, f_n : [a, b] \to \mathbb{R}$ for all $n \in \mathbb{N}$. Assume that $\{f_n\}$ converges uniformly to f on $[a, b]$. Let $x \in [a, b]$ and a sequence $\{x_n\}$ from $[a, b]$ such that $\{x_n\} \to x$.

(1) Prove that
$$\lim_{n\to\infty} f_n(x_n) = f(x)\,.$$

(2) Is (1) true if we assume only pointwise convergence?

Exercise 7.4. Suppose that $\{f_n\}$ converges uniformly on (a, b) and also that $\{f_n(a)\}$ and $\{f_n(b)\}$ converge. Prove that $\{f_n\}$ converges uniformly on $[a, b]$.

Exercise 7.5. Let $f_n : [a, b] \to \mathbb{R}$ be bounded functions for all $n \in \mathbb{N}$.Assume that $\{f_n\}$ converges uniformly to f on $[a, b]$.

(1) Show that $\{f_n\}$ is uniformly bounded on $[a, b]$.

(2) Show that f is bounded on $[a, b]$.

Exercise 7.6. Let $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = x + \frac{1}{n}$ for all $n \in \mathbb{N}$.

(1) Show that $\{f_n\}$ converges uniformly on $\mathbb{R}$.

(2) Prove that $\{(f_n)^2\}$ converges pointwise, but not uniformly, on $\mathbb{R}$.

Exercise 7.7. Let $A \subset \mathbb{R}$ and $f_n, g_n : A \to \mathbb{R}$ be bounded functions for all $n \in \mathbb{N}$.
(a) Prove that if $\{f_n\}$ and $\{g_n\}$ are uniformly convergent on A, then $\{f_n g_n\}$ is uniformly convergent on A.
(b) What happens if we remove the "bounded" assumption?

Exercise 7.8. Let $f : [0, 1] \to \mathbb{R}$ be a continuous function and assume that $f(1) = 0$. Prove that $\{x^n f(x)\}$ converges uniformly on $[0, 1]$.

Exercise 7.9. For each $n \in \mathbb{N}$ define

$$f_n(x) = \frac{1}{n} e^{-n^2x^2}\,.$$

(a) Prove that $\{f_n\}$ converges uniformly on $\mathbb{R}$.

(b) Evaluate the convergence of $\{f_n'\}$.

Exercise 7.10. Prove or disprove: If $\{f_n'\}$ converges uniformly on (a, b), then $\{f_n\}$ converges uniformly on (a, b).

Exercise 7.11. Suppose that $\{f_n\}$ is a sequence of continuous functions defined on $[0, 1]$ and $\{f_n\}$ converges uniformly to f on $[0, 1]$. Show that:

$$\lim_{n\to\infty} \int_0^{1-\frac{1}{n}} f_n(x)\, dx = \int_0^1 f(x)\, dx\,.$$

Exercise 7.12. Suppose that $f : [0, 1] \to \mathbb{R}$ is continuous and define $g_n(x) = f(x^n)$ for all $n \in \mathbb{N}$. Let $a_n = \int_0^1 g_n(x)\, dx$. Prove that $\{a_n\} \to f(0)$.

Exercise 7.13. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuous function and define $f_n(x) = f(nx)$ for all $n \in \mathbb{N}$. Show that if $\{f_n\}$ is equicontinuous on $[0, 1]$, then $f =$ constant.

Exercise 7.14. Suppose that $\{f_n\}$ is an equicontinuous sequence of functions defined on a compact set K, which converges pointwise to f on K. Prove that $\{f_n\}$ converges uniformly to f.

Exercise 7.15. Let $\{f_n\}$ be a uniformly bounded sequence of Riemann integrable functions on $[a, b]$. For every $n \in \mathbb{N}$ define

$$F_n(x) = \int_a^x f_n(t)\, dt\,,\ \forall\, x \in [a, b]\,.$$

Prove that there is a subsequence $\{F_{n_k}\}$ which converges uniformly on $[a, b]$.

Exercise 7.16. Let $f_n : [0, 1] \to \mathbb{R}$ for all $n \in \mathbb{N}$ and suppose that there exists $L > 0$ such that

$$\Big|f_n(x) - f_n(y)\Big| \le L\,\big|x - y\big|\,,\ \forall\, x, y \in [0, 1]\,,\ n \in \mathbb{N}\,.$$

Prove that if $\{f_n\}$ converges pointwise to f on $[0, 1]$, then $\{f_n\}$ converges uniformly to f on $[0, 1]$.

Exercise 7.17. Prove that $\{\sin(n\pi x)\}$ is not equicontinuous on $[0,1]$.

Exercise 7.18. Suppose that $\{f_n\}$ is an equicontinuous sequence which converges pointwise to f on $[a,b]$. Prove that f is continuous on $[a,b]$.

Exercise 7.19. Suppose that f is continuous on $[0,1]$ and assume that

$$\int_0^1 f(x)x^n\,dx = 0\,,\ \forall\, n = 0,1,2,\dots\,.$$

Show that $f(x) = 0$ for all $x \in [0,1]$.

Exercise 7.20. Let $\Phi \subset C([a,b])$ be a bounded set. Show that

$$\Psi = \left\{ F(x) = \int_a^x f(t)\,dt \;\middle|\; f \in \Phi \right\}$$

is compact in $C([a,b])$.

Exercise 7.21. Let $\{f_n\}$ be an equicontinuous sequence on $[a,b]$. Let

$$A = \big\{ x \in [a,b] \;\big|\; \{f_n(x)\} \text{ is convergent} \big\}\ .$$

Show that A is closed.

Exercise 7.22. Let $\{f_n\}$ be a sequence of nonnegative, continuous functions, which converges pointwise to 0 on $[a,b]$. For each $n \in \mathbb{N}$ define

$$g_n(x) = \min\Big\{ f_1(x), \dots, f_n(x) \Big\}\,.$$

Prove that $\{g_n\}$ converges uniformly to 0 on $[a,b]$.

7.2 Series of functions

7.2.1 Definitions and properties

Definition 7.30. Let $A \subset \mathbb{R}$ and $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$. For every $k \in \mathbb{N}$ define $S_k : A \to \mathbb{R}$ by

$$S_k(x) = \sum_{n=1}^{k} f_n(x)\,,\ \forall\, x \in A\,.$$

(1) We say that $\sum_{n=1}^{\infty} f_n$ converges pointwise on A if the sequence $\{S_k\}$ converges pointwise on A.

(2) We say that $\sum_{n=1}^{\infty} f_n$ converges uniformly on A if the sequence $\{S_k\}$ converges uniformly on A.

For simplifying the expressions, when we want to refer to the sum of the series, we will say that $\sum_{n=1}^{\infty} f_n = f$ pointwise (or uniform) on A.

Most of theorems in this sections are corollaries to the theorems regarding sequences of functions.

The following two propositions are immediate consequences of the definitions and their proofs are left as exercises.

Proposition 7.31. *(See Proposition 7.2)*
Let $A \subset \mathbb{R}$ and $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$. If $\sum_{n=1}^{\infty} f_n$ converges uniformly on A, then it converges pointwise on A.

Proposition 7.32. *(See Proposition 7.3)*
Let $A \subset \mathbb{R}$ be a finite set and $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$. If $\sum_{n=1}^{\infty} f_n$ converges pointwise on A, then it converges uniformly on A.

If we remove the assumption that A is a finite set, we have the following theorems.

Theorem 7.33. *(See Theorem 7.4 (2))*
Let $A \subset \mathbb{R}$ and $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$. Then $\sum_{n=1}^{\infty} f_n$ converges uniformly on A if and only if for all $\varepsilon > 0$ there exists $K = K(\varepsilon) > 0$ such that for all $l > k > N$ we have

$$\left|\sum_{n=k}^{l} f_n(x)\right| < \varepsilon\,,\ \forall\, x \in A\,.$$

Proof. Observing that

$$\left|S_l(x) - S_{k-1}(x)\right| = \left|\sum_{n=k}^{l} f_n(x)\right| ,$$

the proof is an immediate consequence of Theorem 7.4 (2).

□

Theorem 7.34. *(The Weierstrass M-Test) (See Theorem 7.4 (1))*
Let $A \subset \mathbb{R}$ and $f_n : A \to \mathbb{R}$, for all $n \in \mathbb{N}$. Define

$$M_n = \sup_{x \in A}\left|f_n(x)\right| , \ \forall\, n \in \mathbb{N} .$$

If $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n$ converges uniformly on A.

Proof. The convergence of $\sum_{n=1}^{\infty} M_n$ implies that for all $\varepsilon > 0$ there exists $K = K(\varepsilon) > 0$ such that for all $l > k > K$ we have

$$0 \leq \sum_{n=k}^{l} M_n < \varepsilon .$$

For all $x \in A$ we can write

$$\left|\sum_{n=k}^{l} f_n(x)\right| \leq \sum_{n=k}^{l}\left|f_n(x)\right| \leq \sum_{n=k}^{l} M_n < \varepsilon .$$

Therefore, by Theorem 7.33, $\sum_{n=1}^{\infty} f_n$ converges uniformly on A.

□

Remark 7.35. Observe that, contrary to Theorem 7.4 (1), the Weierstrass M-Test gives only a one-way implication. The reverse implication is not true in general. Define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} \frac{1}{n} & \text{if } \frac{1}{2^n} < x \leq \frac{1}{2^{n-1}} \\ \\ 0 & \text{otherwise} \end{cases} .$$

It is easy to see that $M_n = \frac{1}{n}$. As the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges, the Weierstrass M-Test cannot be applied. However, as for each $0 < x \leq 1$ there is only one $n \in \mathbb{N}$ such that $f_n(x) \neq 0$,

$$\left|\sum_{n=k}^{l} f_n(x)\right| \leq \frac{1}{k} .$$

Therefore, we can define $K = K(\varepsilon) = \frac{1}{\varepsilon}$ and this proves that $\sum_{n=1}^{\infty} f_n$ converges uniformly on $[0,1]$.

Theorem 7.36. *(Dini's Theorem for Series)*
Let $K \subset \mathbb{R}$ be a compact set and $f, f_n : K \to \mathbb{R}$ for all $n \in \mathbb{N}$. We assume that:

(1) For all $n \in \mathbb{N}$, f_n is continuous on K.

(2) $\sum_{n=1}^{\infty} f_n = f$ pointwise on K and f is continuous on K.

(3) $f_n(x) \geq 0$ for all $x \in K$ and $n \in \mathbb{N}$.

Then $\sum_{n=1}^{\infty} f_n = f$ uniformly on K.

Proof. Notice that the partial sums S_k satisfies the assumptions of Theorem 7.7, and therefore $\sum_{n=1}^{\infty} f_n = f$ uniformly on K.

□

7.2.2 Properties preserved by uniform convergence

Theorem 7.37. *(See Theorem 7.11)*
Let $A \subset \mathbb{R}$ and $f_n : A \to \mathbb{R}$ be continuous on A for all $n \in \mathbb{N}$. If $\sum_{n=1}^{\infty} f_n = f$ uniformly on A, then f is continuous on A.

Proof. We apply Theorem 7.11 to the sequence of partial sums $\{S_k\}$. Indeed, each S_k is continuous on A as a sum of finitely many continuous functions, and $\{S_k\}$ converges uniformly on A, because $\sum_{n=1}^{\infty} f_n = f$ uniformly on A.

Therefore, the limit of the sequence $\{S_k\}$, which is f, is a continuous function on A.

□

Theorem 7.38. *(See Theorem 7.13)*
Let $\alpha : [a, b] \to \mathbb{R}$ be monotone increasing and $f, f_n : [a, b] \to \mathbb{R}$ for all $n \in \mathbb{N}$. Suppose that:

(1) *For all $n \in \mathbb{N}$, f_n is Riemann-Stieltjes integrable with respect to α on $[a, b]$.*

(2) *$\sum_{n=1}^{\infty} f_n = f$ uniformly on $[a, b]$. Then, f is Riemann-Stieltjes integrable with respect to α on $[a, b]$ and*

$$\int_a^b \sum_{n=1}^{\infty} f_n(x)\, d\alpha(x) = \sum_{n=1}^{\infty} \int_a^b f_n(x)\, d\alpha(x)\,.$$

Proof. We apply Theorem 7.13 to the sequence of partial sums $\{S_K\}$. Observe that, each S_k is Riemann-Stieltjes integrable with respect to α on $[a, b]$, as a sum of finitely many such integrable functions.

Also, $\{S_k\}$ converges uniformly on $[a, b]$, because $\sum_{n=1}^{\infty} f_n = f$ uniformly on $[a, b]$.

As conclusion, f is Riemann-Stieltjes integrable with respect to α on $[a, b]$, and

$$\int_a^b f(x)\, d\alpha(x) = \lim_{k\to\infty} \int_a^b S_k(x)\, d\alpha(x)\, d\alpha(x)$$

$$
\begin{aligned}
&= \lim_{k\to\infty} \int_a^b \sum_{n=1}^{k} f_n(x)\, d\alpha(x)\, d\alpha(x) \\
&= \lim_{k\to\infty} \sum_{n=1}^{k} \int_a^b f_n(x)\, d\alpha(x)\, d\alpha(x) \\
&= \sum_{n=1}^{\infty} \int_a^b f_n(x)\, d\alpha(x)\, d\alpha(x)
\end{aligned}
$$

□

Theorem 7.39. *(See Theorem 7.15)*
Let $\{f_n\}$ be a sequence of functions defined on $[a, b]$ with the following properties:

(1) *Each f_n is differentiable on $[a, b]$.*

(2) *$\sum_{n=1}^{\infty} f_n' = g$ uniformly on $[a, b]$.*

(3) *There exists $x_0 \in [a, b]$ such that $\sum_{n=1}^{\infty} f_n(x_0)$ is convergent.*

Then, $\sum_{n=1}^{\infty} f_n = f$ uniformly on $[a, b]$, for some function $f : [a, b] \to \mathbb{R}$. Moreover, f is differentiable on $[a, b]$ and for all $x \in [a, b]$

$$
f'(x) = \left(\sum_{n=1}^{\infty} f_n(x) \right)' = \sum_{n=1}^{\infty} f_n'(x)\,.
$$

Proof. We apply Theorem 7.13 to the sequence of partial sums $\{S_K\}$. Indeed, each S_k is differentiable on $[a, b]$, as a sum of finitely many differentiable functions. Moreover, $\{S_k'\}$ converges uniformly on $[a, b]$, because $\sum_{n=1}^{\infty} f_n' = g$ uniformly on $[a, b]$. In conclusion,

$$
\begin{aligned}
f'(x) &= \lim_{k\to\infty} S_k'(x) \\
&= \lim_{k\to\infty} \left(\sum_{n=1}^{k} f_n(x) \right)' \\
&= \lim_{k\to\infty} \sum_{n=1}^{k} f_n'(x) \\
&= \sum_{n=1}^{\infty} f_n'(x)\,.
\end{aligned}
$$

□

7.2.3 Continuous, nowhere differentiable function

We use Theorem 7.37 to construct a function which is continuous, but nowhere differentiable.

Theorem 7.40. *There exists a function $f : \mathbb{R} \to \mathbb{R}$ which is continuous on $\mathbb{R}$, but there is no $x \in \mathbb{R}$ where $f'(x)$ exists.*

Proof. Define $\phi(x) = |x|$ on $[-1, 1]$ and extend it to $\mathbb{R}$ by requiring the periodicity condition $\phi(x + 2) = \phi(x)$ for all $x \in \mathbb{R}$.

With this definition, it is easy to see that

$$\Big|\phi(x) - \phi(y)\Big| \leq \Big|x - y\Big| \,,\ \forall\, x, y \in \mathbb{R}\,,$$

which shows that ϕ is Lipschitz-continuous on $\mathbb{R}$. Define,

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \phi\left(4^n x\right)\ ,\ \forall\, x \in \mathbb{R}\,.$$

Noticing that $\Big|\phi\left(4^n x\right)\Big| \leq 1$ for all $x \in \mathbb{R}$, the Weierstrass M-Test implies that the series of functions is uniformly convergent. Now, Theorem 7.37 implies that f is continuous.

To show that f is nowhere differentiable, fix any $x \in \mathbb{R}$.
For any $m \in \mathbb{N}$ choose $\delta_m = \pm\frac{1}{2}\, 4^{-m}$, where the sign is chosen in such a way that no integer lies between $4^m x$ and $4^m(x + \delta_m)$. For all $n \in \mathbb{N}$ define

$$\gamma_n = \frac{\phi\left(4^n(x + \delta_m)\right) - \phi\left(4^n x\right)}{\delta_m}\,.$$

If $n > m$, then $4^n \delta_m$ is an integer, so by the periodicity of ϕ, we have $\gamma_n = 0$.
If $0 \leq n < m$, then the Lipschitz-continuity of ϕ implies that $|\gamma_n| \leq 4^n$.
If $n = m$, then $|\gamma_m| = 4^m$. Therefore,

$$\begin{aligned}\left|\frac{f(x + \delta_m) - f(x)}{\delta_m}\right| &= \left|\sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n\right| \\ &\geq 3^m - \sum_{n=0}^{m-1} 3^n \\ &= \frac{1}{2}(3^m + 1)\,.\end{aligned}$$

The fact that

$$\lim_{n\to\infty} \frac{1}{2}(3^m + 1) = +\infty\,,$$

implies that f is not differentiable at x.

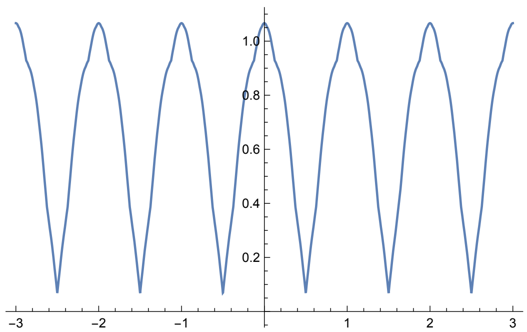
$\square$

Remark 7.41. To see the interplay between the amplitudes and frequences, three graphs will be shown. For easier coding, instead of the function ϕ we will use the $|\cos(x\pi)|$.

First, let

$$f_1(x) = \sum_{n=0}^{\infty} \left(\frac{1}{16}\right)^n \left|\cos(4^n x\pi)\right|.$$

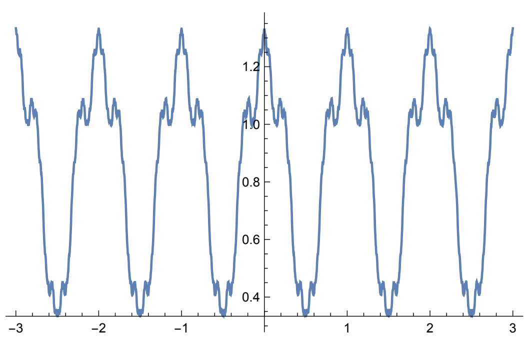
Its graph approximately looks as:



Continue with:

$$f_1(x) = \sum_{n=0}^{\infty} \left(\frac{1}{4}\right)^n \left|\cos(4^n x\pi)\right|.$$

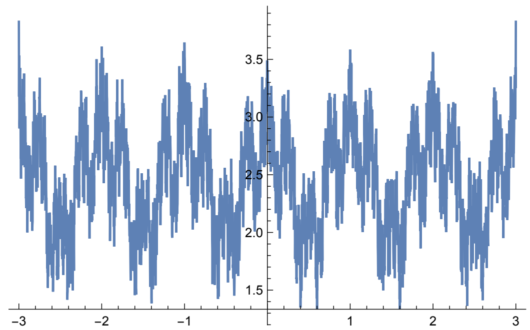
Its graph approximately looks as:



Now consider a nowhere differentiable function:

$$f_1(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \left|\cos(4^n x\pi)\right|.$$

Its graph approximately looks as:



We can notice a gradual loss of the differentiability properties, as the amplitude increases, but the frequency stays the same.

7.2.4 Power series

Definition 7.42. Let $\{c_0, c_1, \dots\}$ be a sequence of real numbers and $a \in \mathbb{R}$. We say that the series

$$\sum_{n=0}^{\infty} c_n (x-a)^n$$

is a power series with coefficients $\{c_n\}$ and center at a.

Definition 7.43. Let $\sum_{n=0}^{\infty} c_n (x-a)^n$ be a power series and

$$\lambda = \limsup \sqrt[n]{|c_n|}\,.$$

Define

$$R = \begin{cases} \frac{1}{\lambda} & \text{if } 0 < \lambda < \infty \\ 0 & \text{if } \lambda = \infty \\ \infty & \text{if } \lambda = 0 \end{cases} \,. \tag{7.7}$$

R is called the radius of convergence (See Theorem 7.46).

Remark 7.44. By Lemma 6.23, if we use $\limsup \frac{|a_{n+1}|}{|a_n|}$ we might get a higher λ and therefore a smaller R. However, if $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$ exists, the two quantities are the same.

Definition 7.45. By $C^\infty((a,b))$ we denote the set of those functions $f : (a,b) \to \mathbb{R}$ which have derivatives of any order on (a,b). We call them C^∞ or smooth functions.

Theorem 7.46. *Let $\sum_{n=0}^{\infty} c_n (x-a)^n$ be a power series and R defined by (7.7). Then:*

(1) *If $R = 0$ then the power series converges only at $x = a$.*

(2) *If* $0 < R < \infty$, *then*

(2.1) *The power series converges pointwise on* $(a - R, a + R)$.

(2.2) *For any* $0 < \varepsilon < \frac{R}{2}$, *the power series converges uniformly on* $[a - R + \varepsilon, a + R - \varepsilon]$.

(2.3) *If* $f(x) = \sum_{n=1}^{\infty} c_n (x - a)^n$, *then* $f \in C^{\infty}((a - R, a + R))$ *and for all* $k \in \mathbb{N}$ *we have*

$$f^{(k)}(x) = \sum_{n=k}^{\infty} c_n n(n-1)\ldots(n-k+1)(x-a)^{n-k},$$

$$c_k = \frac{f^{(k)}(a)}{k!},$$

and for any $[\alpha, \beta] \subset (a - R, a + R)$

$$\int_{\alpha}^{\beta} f(x)\,dx = \sum_{n=0}^{\infty} c_n \int_{\alpha}^{\beta} (x-a)^n\,dx.$$

(3) *If* $R = +\infty$, *then*

(3.1) *The power series converges pointwise on* $\mathbb{R}$.

(3.2) *For any* $M > 0$, *the power series converges uniformly on* $[a - M, a + M]$.

(3.3) *If* $f(x) = \sum_{n=1}^{\infty} c_n (x - a)^n$, *then* $f \in C^{\infty}(\mathbb{R})$ *and the formulas from* (2.3) *hold.*

Proof. Fix any $x \in \mathbb{R}$ and use the Root Test to check the absolute convergence:

$$\limsup \sqrt[n]{|c_n(x-a)^n|} = |x-a| \limsup \sqrt[n]{|c_n|} = |x-a|\lambda < 1. \tag{7.8}$$

(1) If $\lambda = +\infty$, the only way to get

$$\limsup \sqrt[n]{|c_n(x-a)^n|} < 1,$$

is to have $x - a = 0$, which is $x = a$. In this case the only possibly nonzero term in the power series is c_0.

(2) If $0 < \lambda < \infty$, then (7.8) implies that

$$|x-a| < \frac{1}{\lambda} = R.$$

Therefore, for all $x \in (a - R, a + R)$ the power series converges, and this proves (2.1).

(2.2) If $x \in [a - R + \varepsilon, a + R - \varepsilon]$, then $|x - a| \leq R - \varepsilon$ and

$$\sum_{n=0}^{\infty} |c_n|(R-\varepsilon)^n$$

is convergent, because

$$\limsup \sqrt[n]{|c_n|}(R-\varepsilon) = \frac{R-\varepsilon}{R} < 1\,.$$

Hence, the Weierstrass M-Test implies the uniform convergence of the power series on the interval $[a-R+\varepsilon, A+R-\varepsilon]$.

(2.3) Observe, that for any $n \geq k$ we have

$$\limsup \sqrt[n]{c_n} = \limsup \sqrt[n]{c_n n(n-1)\ldots(n-k+1)}\,,$$

so, the differentiated power series have the same interval of convergence. Moreover, (2.2) applies to them, and therefore, by Theorem 7.39, f is differentiable, as many times we want, on $[a-R+\varepsilon, A+R-\varepsilon]$. However, if we fix any $x \in (a-R, a+R)$, then we can find $0 < \varepsilon < \frac{R}{2}$ such that $x \in [a-R+\varepsilon, A+R-\varepsilon]$. Hence, f is differentiable at x and the formulas from (2.3) follow from Theorem 7.39.

(3) If $\lambda = 0$, then (7.8) shows that for all $x \in \mathbb{R}$

$$\limsup \sqrt[n]{|c_n(x-a)^n|} = 0\,,$$

which implies the pointwise convergence on $\mathbb{R}$. In case of any $M > 0$, the series converges at $x = M + a$ and the Weierstrass M-Test implies the uniform convergence on $[a-M, a+M]$. The rest of the proof is similar to (2.3).

□

Theorem 7.47. *Let $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ be a power series and R defined by (7.7). Let us assume that:*

(1) $0 < R < \infty$.

(2) $\sum_{n=1}^{\infty} c_n$ *converges.*

Then, f can be continuously extended to $a+R$ and

$$\lim_{x\to a+R} f(x) = \sum_{n=1}^{\infty} c_n\,.$$

Remark 7.48. In a similar way, if $\sum_{n=1}^{\infty}(-1)^n c_n$ converges, then f can be continuously extended to $a-R$ and

$$\lim_{x\to a-R} f(x) = \sum_{n=1}^{\infty}(-1)^n c_n\,.$$

Proof. Without loos of generality we can consider $a = 0$ and $R = 1$. Let us define

$$S_k = \sum_{n=0}^{k} c_n\,,\ \forall\, k = 0,1,2,\ldots\,.$$

Then,

$$\begin{aligned}
\sum_{n=0}^{k} c_n x^n &= c_0 + \sum_{n=1}^{k} \big(S_n - S_{n-1}\big) x^n \\
&= c_0 + \sum_{n=1}^{k} S_n x^n - \sum_{n=1}^{k} S_{n-1} x^n \\
&= \sum_{n=0}^{k} S_n x^n - \sum_{n=0}^{k-1} S_n x^{n+1} \\
&= (1-x) \sum_{n=0}^{k-1} S_n x^n + S_k x^k \, .
\end{aligned}$$

□

If $|x| < 1$, then letting $k \to +\infty$ leads to

$$f(x) = (1-x) \sum_{n=0}^{\infty} S_n x^n \, .$$

Let us denote $S = \sum_{n=0}^{\infty} c_n$ and fix any $\varepsilon > 0$. By the convergence $\{S_k\} \to S$, we can choose $K = K(\varepsilon) \in \mathbb{N}$ such that

$$\Big| S - S_k \Big| < \frac{\varepsilon}{2} \, , \; \forall \, k > K \, .$$

Using the fact that

$$(1-x) \sum_{n=0}^{\infty} x^n = 1 \, , \text{ if } |x| < 1 \, ,$$

we obtain

$$\begin{aligned}
\Big| f(x) - S \Big| &= \Big| (1-x) \sum_{n=0}^{\infty} (S_n - S) x^n \\
&\leq (1-x) \sum_{n=0}^{K} \big| S_n - S \big| \, \big| x \big|^n + (1-x) \sum_{n=K+1}^{\infty} \big| S_n - S \big| \, \big| x \big|^n \\
&\leq (1-x) \sum_{n=0}^{K} \big| S_n - S \big| + \frac{\varepsilon}{2} \, .
\end{aligned}$$

Therefore, by choosing $\delta = \delta(\varepsilon) > 0$ such that for all $1 - \delta < x < 1$ we have

$$(1-x) \sum_{n=0}^{K} \big| S_n - S \big| < \frac{\varepsilon}{2} \, ,$$

we get

$$\Big| f(x) - S \Big| < \varepsilon\,.$$

In conclusion,

$$\lim_{x \to 1} f(x) = \sum_{n=1}^{\infty} c_n\,.$$

Example 7.49. Consider the power series

$$\sum_{n=0}^{\infty} x^n\,,$$

which is the geometric series studied earlier. We can observe that $R = 1$, hence

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}\,,\ \forall\, x \in (-1,1)\,.$$

By Theorem 7.46, we can differentiate it and integrate it term by term. So,

$$\sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}\,,\ \forall\, x \in (-1,1)\,.$$

We can multiply both sides by x and get

$$\sum_{n=1}^{\infty} n x^{n} = \frac{x}{(1-x)^2}\,,\ \forall\, x \in (-1,1)\,.$$

For $x = \frac{1}{2}$ we get

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

which cannot be obtained by the methods shown in Chapter 6.

Example 7.50. Consider the power series

$$\sum_{n=0}^{\infty} (-1)^n x^n\,.$$

It is easy to observe that $R = 1$, and hence we get

$$\sum_{n=0}^{\infty} (-1)^n x^n = \frac{1}{1+x}\,,\ \forall\, x \in (-1,1)\,.$$

Integrating this series on $[0,x]$ for any $0<x<1$ we get

$$\sum_{n=0}^{\infty}(-1)^n\int_0^x t^n\,dt=\int_0^x\frac{1}{1+t}\,dt\,.$$

Therefore,

$$\sum_{n=0}^{\infty}(-1)^n\frac{x^n}{n}=\ln(1+x)\,,\ \forall\,x\in(-1,1]\,.$$

Notice that, by Theorem 7.47, the interval of convergence was extended to the right endpoint. Hence, for $x=1$ we get

$$\sum_{n=0}^{\infty}(-1)^n\frac{1}{n}=\ln 2\,.$$

Exercises

Exercise 7.23. Show that, if $\sum_{n=1}^{\infty} f_n$ converges uniformly on $A\subset\mathbb{R}$, then $\{f_n\}$ converges uniformly to 0 on A.

Exercise 7.24. Study the pointwise and uniform convergences of the following series of functions:

$$(a)\quad \sum_{n=1}^{\infty}\frac{x^{2n}}{(x+n)^2}\,,\ \text{on } [0,1].$$

$$(b)\quad \sum_{n=0}^{\infty}\frac{x^n}{n!}\,,\ \text{on } [-M,M]\,,\ M>0.$$

$$(c)\quad \sum_{n=0}^{\infty}\frac{x^n}{n!}\,,\ \text{on } \mathbb{R}.$$

$$(d)\quad \sum_{n=1}^{\infty}n^{-x}\,,\ \text{on } [2,+\infty).$$

$$(e)\quad \sum_{n=1}^{\infty}\frac{x^2}{1+n^2x^2}\,,\ \text{on } \mathbb{R}.$$

$$(f)\quad \sum_{n=1}^{\infty}\frac{1}{1+n^2x^2}\,,\ \text{on } (0,1].$$

$$(g)\quad \sum_{n=1}^{\infty} \frac{1}{1+n^2x^2}\,, \text{ on } [1,+\infty).$$

$$(h)\quad \sum_{n=1}^{\infty} \frac{n^3}{x^n}\,, \text{ on } [2,+\infty).$$

$$(i)\quad \sum_{n=1}^{\infty} \left(\frac{\ln x}{x}\right)^n\,, \text{ on } [1,+\infty).$$

$$(j)\quad \sum_{n=1}^{\infty} 2^n \sin\left(\frac{1}{3^n x}\right)\,, \text{ on } [M,+\infty)\,,\ M>0.$$

$$(k)\quad \sum_{n=1}^{\infty} (-1)^n \frac{x^2+n}{n^2}\,, \text{ on } [0,1].$$

$$(l)\quad \sum_{n=1}^{\infty} \frac{x}{n(1+nx^2)}\,, \text{ on } \mathbb{R}.$$

$$(m)\quad \sum_{n=1}^{\infty} \sin^{2n} x\,, \text{ on } \left[0,\frac{\pi}{3}\right].$$

Exercise 7.25. Find the interval of convergence for the following power series:

$$(a)\quad \sum_{n=1}^{\infty} n^n\,(x+2)^n\,.$$

$$(b)\quad \sum_{n=0}^{\infty} \frac{(n!)^2}{(2n)!}\,(x-2)^n\,.$$

$$(c)\quad \sum_{n=1}^{\infty} \frac{(x+3)^n}{n2^n}\,.$$

$$(d) \quad \sum_{n=1}^{\infty} \frac{10^n}{n!}\, x^n\,.$$

$$(f) \quad \sum_{n=1}^{\infty} \frac{(2n+1)!}{2^n (n!)^2}\, x^n\,.$$

Exercise 7.26. Calculate the sum of the following series. Explain the validity of your answers.

$$(a) \quad \sum_{n=1}^{\infty} n^2 x^n\,, \qquad \sum_{n=1}^{\infty} \frac{n^2}{3^n}\,.$$

$$(b) \quad \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}\,, \qquad \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1}\,.$$

$$(c) \quad \sum_{n=1}^{\infty} \frac{x^n}{n 3^n}\,, \qquad \sum_{n=1}^{\infty} \frac{1}{n 3^n}\,.$$

Exercise 7.27. Let $\sum_{n=1}^{\infty} c_n$ be an absolutely convergent series. Show that $\sum_{n=1}^{\infty} c_n \cos(nx)$ converges uniformly on $\mathbb{R}$.

Exercise 7.28. Suppose that for all $n \in \mathbb{N}$ the functions $f_n, g_n : [a,b] \to \mathbb{R}$ satisfy the following conditions:

(1) $\sum_{n=1}^{\infty} f_n$ has uniformly bounded partial sums on $[a,b]$.

(2) $\{g_n\}$ converges uniformly to 0 on $[a,b]$.

(3) $g_n(x) \geq g_{n+1}(x)$ for all $n \in \mathbb{N}$ and $x \in [a,b]$.

Prove that $\sum_{n=1}^{\infty} f_n g_n$ converges uniformly on $[a,b]$.

Bibliography

[1] J. Dieudonné, Foundations of Modern Analysis, Volume 10-I, 1969, Academic Press, Inc.

[2] Russel A. Gordon, Real Analysis: a first course, 2002, Pearson Education, Inc.

[3] James R. Kirkwood, An Introduction to Analysis, 2nd edition, 1995, Waveland Press, Inc.

[4] A. N. Kolmogorov and S. V. Fomin, Introductory Real Analysis, 1975, Dover Publications, Inc.

[5] H. L. Royden, Real Analysis, 1964, The Macmillan Company, New York.

[6] Walter Rudin, Principles of Mathematical Analysis, 3rd edition, 1976, McGraw-Hill, Inc.

Index