Math 330 - Additional Material

Skeletal version: proofs omitted

Michael Fochler
Department of Mathematics
Binghamton University

This skeletal version of the document is meant to serve as a help to review the material for upcoming exams. It not only omits the proofs, but also many motivational paragraphs and examples, and even some propositions and theorems. The references are out of sync with the full student edition, so

Do NOT use this edition to find an item referenced, e.g., in your homework assignment!

Last update: October 4, 2025

Contents

1 Before You Start

1.1 About This Document

1.2 How to Properly Write a Proof

Transitivity of equality means that if A = B and B = C then A = C.

1.3 Blank Page after Ch.1

This page is intentionally left blank!

2 Preliminaries about Sets, Numbers and Functions

2.1 Sets and Basic Set Operations

Definition 2.1 (Sets)**.** A **set** is a collection of stuff called **members** or **elements** which satisfies the following rules:

- The order in which the elements are written does not matter.
- If an element is listed two or more times, then **it only counts once**!

We write a set by enclosing within curly braces the elements of the set. This can be done by listing all those elements or giving instructions that describe those elements. □

For example, to denote by X the set of all integer numbers between 18 and 24 we can write either of the following:

$$X := \{18, 19, 20, 21, 22, 23, 24\} \quad \text{or} \quad X := \{n : n \text{ is an integer and } 18 \leq n \leq 24\}$$

Both formulas clearly define the same collection of all integers between 18 and 24. On the left the elements of X are given by a complete list, on the right we use instead **setbuilder notation**, i.e., instructions that specify what belongs to the set.

It is customary to denote sets by capital letters and their elements by small letters but this is not a hard and fast rule. You will see many exceptions to this rule in this document.

We write $x_1 \in X$ to denote that an item x_1 is an element of the set X and $x_2 \notin X$ to denote that an item x_2 is not an element of the set X

For the above example we have $20 \in X$, $27 - 6 \in X$, $38 \notin X$, 'Jimmy' $\notin X$.

Definition 2.2 (empty set)**.** The **empty set** is the set that does not contain any elements. It is uniquely determined by this property.

The symbols $\emptyset$ and $\{\,\}$ are both in use to denote this set. However, we **STRONGLY DISCOURAGE** the use of $\{\,\}$, since since this makes expressions with nested braces hard to read. □

Definition 2.3 (subsets, supersets and equality of sets)**.**

(a) We say that a set A is a **subset** of the set B and we write $A \subseteq B$ if each element of A also belongs to B. Equivalently we say that B is a **superset** of the set A and we write $B \supseteq A$. We also say that B includes A or A is included by B. Note that $A \subseteq A$ and $\emptyset \subseteq A$ is true for all sets A.

(b) If $A \subseteq B$ but $A \neq B$, i.e., there is at least one $x \in B$ such that $x \notin A$, then we say that A is a **strict subset** or a **proper subset** of B. We write "$A \subsetneqq B$" or "$A \subset B$". Alternatively, we say that B is a **strict superset** or a **proper superset** of A and we write "$B \supsetneqq A$" or "$B \supset A$".

(c) We say that two sets A and B are **equal** and we write $A = B$, if both $A \subseteq B$ and $B \subseteq A$ $\square$

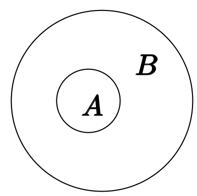


Figure 2.1: Set inclusion: $A \subseteq B,\ B \supseteq A$

Definition 2.4 (Unions and intersections of two sets)**.** Given are two arbitrary sets A and B. No assumption is made that either one is contained in the other or that either one is not empty!

(a) The **union** $A \cup B$ (pronounced "A union B") is defined as the set of all elements which belong to A or B or both.

(b) The **intersection** $A \cap B$ (pronounced "A intersection B") is defined as the set of all elements which belong to both A and B. $\square$

Definition 2.5 (Unions and intersections of n sets)**.** Let $A_1, A_2, \dots, A_n$ be arbitrary sets.

(a) The **union** $\bigcup\limits_{j=1}^{n} A_j := A_1 \cup A_2 \cup \dots \cup A_n$ is defined as the set of all those items which belong to at least one of the sets, i.e.,

$$x \in \bigcup_{j=1}^{n} A_j \Leftrightarrow x \in A_j \text{ for at least one index } j. \tag{2.1}$$

(b) The **intersection** $\bigcap\limits_{j=1}^{n} A_j := A_1 \cap A_2 \cap \dots \cap A_n$ is defined as the set of all those items which belong to each and everyone of the sets, i.e.,

$$x \in \bigcap_{j=1}^{n} A_j \Leftrightarrow x \in A_j \text{ for each index } j. \ \square \tag{2.2}$$

Definition 2.6 (Disjoint unions)**.** We call two sets A and B **disjoint** , also **mutually disjoint**, , if $A \cap B = \emptyset$. More generally, we say that a collection of sets $A_1, A_2, \dots, A_n$ is (mutually) disjoint if each pair A_i, A_j for different indices i and j is disjoint. We often write "$\uplus$" (pronounced "disjoint union") rather than "$\cup$" to remind the reader that we are dealing with unions of disjoint sets, i.e., we write

$$A \uplus B \qquad A_1 \uplus A_2 \uplus \cdots \uplus A_n\,, \qquad \biguplus_{j=1}^{n} A_j\,,$$

rather than $A \cup B,\ A_1 \cup A_2 \cup \cdots \cup A_n,\ \bigcup\limits_{j=1}^{n} A_j.$ $\square$

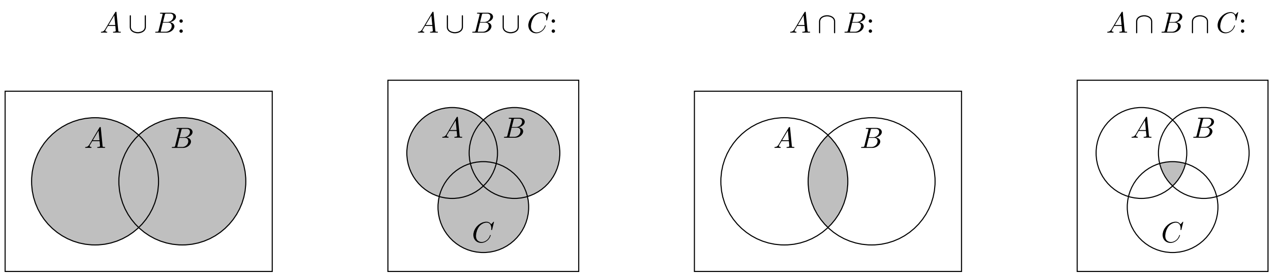


Figure 2.2: Union and intersection of sets

Definition 2.7 (Set differences and symmetric differences)**.** Given are two sets A and B. No assumption is made that either one is contained in the other or that either one is not empty!

The **difference set** or **set difference** $A \setminus B$ (pronounced "A minus B") is defined as the set of all elements which belong to A but not to B:

$$A \setminus B := \{x \in A : x \notin B\} \tag{2.3}$$

The **symmetric difference** $A \triangle B$ (pronounced "A delta B") is defined as the set of all elements which belong to either A or B but not to both A and B:

$$A \triangle B := (A \cup B)\ \setminus\ (A \cap B) \quad \square \tag{2.4}$$

Definition 2.8 (Universal set)**.**
There usually is a big set Ω that contains everything we are interested in, and we then deal with all kinds of subsets $A \subseteq \Omega$. Such a set is called a **"universal" set**. □

Definition 2.9 (Complement of a set)**.** Let Ω be a universal set. The **complement** $A^{\complement}$. of a set $A \subseteq \Omega$ consists of all elements of Ω which do not belong to A. In other words:

$$A^{\complement} == \Omega \setminus A = \{\omega \in \Omega : \omega \notin A\}. \quad \square \tag{2.5}$$

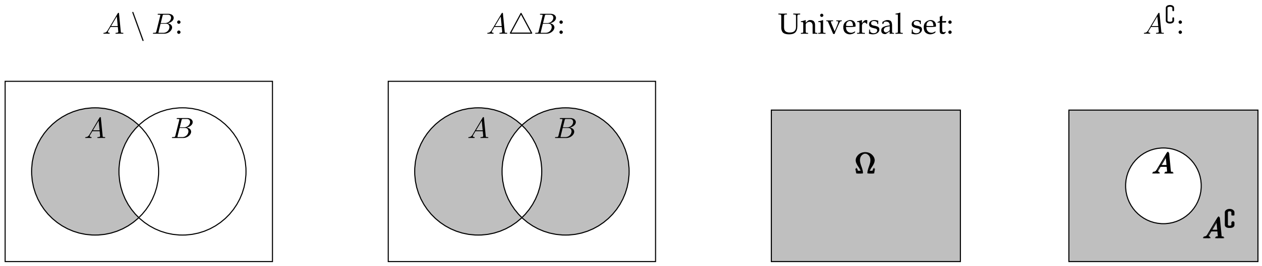


Figure 2.3: Difference, symmetric difference, universal set, complement

Remark 2.1. Note the following: If Ω is a universal set then

$$\Omega^{\complement} = \emptyset, \qquad \emptyset^{\complement} = \Omega. \quad \square \tag{2.6}$$

Proposition 2.1. *Let A, B, X be subsets of a universal set Ω and assume $A \subseteq X$. Then*

$$A \cup \emptyset = A; \qquad A \cap \emptyset = \emptyset \tag{2.7a}$$
$$A \cup \Omega = \Omega; \qquad A \cap \Omega = A \tag{2.7b}$$
$$A \cup A^{\complement} = \Omega; \qquad A \cap A^{\complement} = \emptyset \tag{2.7c}$$
$$A \triangle B = (A \setminus B) \uplus (B \setminus A) \tag{2.7d}$$
$$A \setminus A = \emptyset \tag{2.7e}$$
$$A \triangle \emptyset = A; \qquad A \triangle A = \emptyset \tag{2.7f}$$
$$X \triangle A = X \setminus A \tag{2.7g}$$
$$A \cup B = (A \triangle B) \uplus (A \cap B) \tag{2.7h}$$
$$A \cap B = (A \cup B) \setminus (A \triangle B) \tag{2.7i}$$
$$A \triangle B = \emptyset \text{ if and only if } B = A \tag{2.7j}$$

Proposition 2.2 (Distributivity of unions and intersections for two sets)**.**
Let A, B, C be sets. Then

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \tag{2.8}$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C). \tag{2.9}$$

Proposition 2.3 (De Morgan's Law for two sets)**.**
Let $A, B \subseteq \Omega$. Then the complement of the union is the intersection of the complements, and the complement of the intersection is the union of the complements:

(2.10) ***(a)*** $(A \cup B)^{\complement} = A^{\complement} \cap B^{\complement}$ ***(b)*** $(A \cap B)^{\complement} = A^{\complement} \cup B^{\complement}$

Proposition 2.4. *Let A, B, C, Ω be sets such that $A, B, C \subseteq \Omega$. Then*

- ***(a)*** $(A \triangle B) \triangle C = A \triangle (B \triangle C)$
- ***(b)*** $A \triangle \emptyset = \emptyset \triangle A = A$
- ***(c)*** $A \triangle A = \emptyset$
- ***(d)*** $A \triangle B = B \triangle A$

Further we have the following for the intersection operation:

- ***(e)*** $(A \cap B) \cap C = A \cap (B \cap C)$
- ***(f)*** $A \cap \Omega = \Omega \cap A = A$
- ***(g)*** $A \cap B = B \cap A$

And we have the following interrelationship between $\triangle$ and $\cap$:

- ***(h)*** $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$

Definition 2.10 (Power set)**.** The **power set**

$$2^{\Omega} := \{A : A \subseteq \Omega\}$$

of a set Ω is the set of all its subsets. Note that many older texts also use the notation $\mathfrak{P}(\Omega)$ for the power set. □

Remark 2.2. Note that $\emptyset \in 2^{\Omega}$ for all sets Ω, even if $\Omega = \emptyset$, since $2^{\emptyset} = \{\emptyset\}$. In particular, the power set of the empty set is not empty. □

Definition 2.11 (Partition)**.** Let Ω be a set and $\mathcal{A} \subseteq 2^{\Omega}$, i.e., the elements of $\mathcal{A}$ are subsets of Ω. We call $\mathcal{A}$ a **partition** or a **partitioning** of Ω if

(a) If $A, B \in \mathcal{A}$ such that $A \neq B$ then $A \cap B = \emptyset$. In other words, $\mathcal{A}$ consists of mutually disjoint subsets of Ω (see Definition 2.6),

(b) Each $x \in \Omega$ is an element of some $A \in \mathcal{A}$. □

Definition 2.12 (Size of a set (preliminary))**.**

(a) Let X be a finite set, i.e., a set which only contains finitely many elements. We write $|X|$ for the number of its elements, and we call $|X|$ the **size** of the set X.

(b) For infinite, i.e., not finite sets Y, we define $|Y| := \infty$. □

2.2 The Proper Use of Language in Mathematics: Any vs All, etc

2.3 Numbers

Definition 2.13 (Integers and decimal numerals)**.**

A **digit** or **decimal digit** Is one of the numbers $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$.

We call numbers that can be expressed as a finite string of digits, possibly preceded by a minus sign, **integers**. In particular we demand that an integer can be written without a decimal point.

A **decimal** or **decimal numeral** is a finite or infinite list of digits, possibly preceded by a minus sign, which is separated into two parts by a point, the **decimal point**. □

Definition 2.14 (Real numbers)**.** We call any kind of number which can be represented as a decimal numeral, a **real number**. We write $\mathbb{R}$ for the set of all real numbers. It follows from what was remarked at the end of Definition 2.13 that integers, in particular natural numbers, are real numbers. Thus we have the following set relations:

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{R}. \quad \square \tag{2.11}$$

Definition 2.15 (Rational numbers)**.** A number that is an integer or can be written as a fraction of integers, i.e., as $\frac{m}{n}$ where $m, n \in \mathbb{Z}$ and $n \neq 0$, is called a **rational number**. We write $\mathbb{Q}$ for the set of all rational numbers. □

Definition 2.16 (Irrational numbers)**.** We call real numbers that are not rational **irrational numbers**. $\square$

Definition 2.17 (Types of numbers)**.** $\mathbb{N} := \{1, 2, 3, \dots\}$ denotes the set of **natural numbers**.

$\mathbb{Z} := \{0, \pm 1, \pm 2, \pm 3, \dots\}$ denotes the set of all **integers**.

$\mathbb{Q} := \{n/d : n \in \mathbb{Z}, d \in \mathbb{N}\}$ denotes the set of all **rational numbers**.

$\mathbb{R} :=$ {all integers or decimal numbers with finitely or infinitely many decimal digits} denotes the set of all **real numbers**.

$\mathbb{R} \setminus \mathbb{Q}$ = {all real numbers which cannot be written as fractions of integers} denotes the set of all **irrational numbers**. There is no special symbol for irrational numbers. Example: $\sqrt{2}$ and π are irrational. $\square$

$\mathbb{N}_0 := \mathbb{Z}_+ := \mathbb{Z}_{\geq 0} := \{0, 1, 2, 3, \dots\}$ denotes the set of nonnegative integers,
$\mathbb{R}_+ := \mathbb{R}_{\geq 0} := \{x \in \mathbb{R} : x \geq 0\}$ denotes the set of all nonnegative real numbers,
$\mathbb{R}^+ := \mathbb{R}_{>0} := \{x \in \mathbb{R} : x > 0\}$ denotes the set of all positive real numbers,
$\mathbb{R}^\star := \mathbb{R}_{\neq 0} := \{x \in \mathbb{R} : x \neq 0\}$. $\square$

Definition 2.18 (Translation and dilation of sets of numbers)**.** For a set of numbers A and numbers λ and b, we define

$$\lambda A + b := \{\lambda a + b : a \in A\}. \tag{2.12}$$

In particular, for $\lambda = \pm 1$, we obtain

$$A + b = \{a + b : a \in A\}, \tag{2.13}$$

$$-A = \{-a : a \in A\}. \quad \square \tag{2.14}$$

Definition 2.19 (Intervals of Numbers)**.** For $a, b \in \mathbb{R}$ we have the following intervals.

- $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$ is the **closed interval** with endpoints a and b.
- $]a, b[:= \{x \in \mathbb{R} : a < x < b\}$ is the **open interval** with endpoints a and b.
- $[a, b[:= \{x \in \mathbb{R} : a \leq x < b\}$ and $]a, b] := \{x \in \mathbb{R} : a < x \leq b\}$ are **half-open intervals** with endpoints a and b.

(2.15)
$$
\begin{aligned}
]-\infty, a] &:= \{x \in \mathbb{R} : x \leq a\}, \quad]-\infty, a[\, := \{x \in \mathbb{R} : x < a\},\\
]a, \infty[\, &:= \{x \in \mathbb{R} : x > a\}, \quad [a, \infty[\, := \{x \in \mathbb{R} : x \geq a\}, \quad]-\infty, \infty[\, := \mathbb{R}
\end{aligned}
$$

- $[a, a] = \{a\}$; $[a, a[\, = \,]a, a[\, = \,]a, a] = \emptyset$
- $[a, b] = [a, b[\, = \,]a, b[\, = \,]a, b] = \emptyset$ for $a \geq b$ □

Notation 2.1 (Notation Alert for intervals of integers or rational numbers)**.**
It is at times convenient to also use the notation $[\dots],]\dots[, [\dots[,]\dots]$, for intervals of integers or rational numbers. We will subscript them with $\mathbb{Z}$ or $\mathbb{Q}$. For example,

$$
\begin{aligned}
[\,3, n\,]_{\mathbb{Z}} &= [\,3,\ n] \cap \mathbb{Z} = \{k \in \mathbb{Z} : 3 \leq k \leq n\},\\
]-\infty, 7\,]_{\mathbb{Z}} &=]-\infty, 7\,] \cap \mathbb{Z} = \{k \in \mathbb{Z} : k \leq 7\} = \mathbb{Z}_{\leq 7},\\
]a, b[_{\mathbb{Q}} &=]a, b[\, \cap \mathbb{Q} = \{q \in \mathbb{Q} : a < q < b\}.
\end{aligned}
$$

An interval which is not subscripted always means an interval of real numbers, but we will occasionally write, e.g., $[a, b]_{\mathbb{R}}$ rather than $[a, b]$, if the focus is on integers or rational numbers and an explicit subscript helps to avoid confusion. □

Definition 2.20 (Absolute value)**.** For a real number x we define its **absolute value** as

$$
|x| = \begin{cases} x & \text{if } x \geq 0,\\ -x & \text{if } x < 0. \end{cases} \quad \square
$$

Assumption 2.1 (Square roots are always assumed nonnegative)**.** We will always assume that "$\sqrt{b}$" is the **positive** value unless the opposite is explicitly stated. □

Proposition 2.5 (The Triangle Inequality for real numbers)**.** *The following inequality is used all the time in mathematical analysis to show that the size of a certain expression is limited from above:*

(2.16) $$\textit{Triangle Inequality}: \quad |\boldsymbol{a} + \boldsymbol{b}| \leq |\boldsymbol{a}| + |\boldsymbol{b}|$$

This inequality is true for any two real numbers a and b.

Definition 2.21 (Kronecker symbol). ⋆

For $i, j \in \mathbb{N}$, the **Kronecker symbol** δ_{ij}, also called the **Kronecker delta**, is defined as follows.

$$\delta_{ij} := \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \quad \square$$

2.4 A First Look at Functions, Sequences and Families

Definition 2.22 (Preliminary definition of a function). A **function** f consists of two nonempty sets X and Y and an assignment rule $x \mapsto f(x)$ which assigns any $x \in X$ uniquely to some $y \in Y$. We write $f(x)$ for this assigned value and call it the **function value** of the **argument** x. X is called the **domain** and Y is called the **codomain** of f. We write

(2.17) $$f : X \to Y, \qquad x \mapsto f(x).$$

We read "$a \mapsto b$" as "a is assigned to b" or "a maps to b" and refer to $\mapsto$ as the **maps to operator** or **assignment operator**. The **graph** of such a function is the collection of pairs

(2.18) $$\Gamma_f := \{(x, f(x)) : x \in X\}. \quad \square$$

Definition 2.23 (Preliminary definition of the inverse function). Given are two nonempty sets X and Y and a function $f : X \to Y$ with domain X and codomain Y. We say that f has an **inverse function** if it satisfies all of the following conditions which uniquely determine this inverse function, so that we are justified to give it the symbol f^{-1}:

- **(a)** $f^{-1} : Y \to X$, i.e., f^{-1} has domain Y and codomain X.
- **(b)** $f^{-1}(f(x)) = x$ for all $x \in X$, and $f(f^{-1}(y)) = y$ for all $y \in Y$. $\square$

Definition 2.24. Let $n_\star$ be an integer and let there be a uniquely determined item x_j for each integer $j \geq n_\star$. Such an item can be, e.g., a number or a set (the only items we are looking at for now).

In other words, assume that a unique item x_j is assigned to each $j \in [n_\star, \infty[_\mathbb{Z}$. We write

$$(x_j)_{j \geq n_\star} \quad \text{or} \quad (x_j)_{j \in [n_\star, \infty[_\mathbb{Z}} \quad \text{or} \quad (x_j)_{j=n_\star}^{\infty} \quad \text{or} \quad x_{n_\star}, x_{n_\star+1}, x_{n_\star+2}, \dots$$

for such a collection of items, and we call it a **sequence** with **start index** $n_\star$. We call the set $[n_\star, \infty[_\mathbb{Z}$ of indices the **index set** of the sequence.

The symbol j is a dummy variable, same as the name x of the argument of a function $f(x)$. See Remark **??** on p.**??**. $\square$

Definition 2.25. We occasionally admit an "ending index" $n^\star$ instead of ∞, i.e., there will be an indexed item x_j, for each $j \in [n_\star, n^\star]_{\mathbb{Z}}$. We then talk of a **finite sequence**, and we write

$$(x_n)_{n_\star \leq n \leq n^\star} \quad \text{or} \quad (x_j)_{j=n_\star}^{n^\star} \quad \text{or} \quad x_{n_\star},\, x_{n_\star+1},\, \dots,\, x_{n^\star}$$

for such a finite collection of items. If we refer to a sequence $(x_n)_n$ without qualifying it as finite then we imply that we deal with an **infinite sequence**, $(x_n)_{n=n_\star}^{\infty}$.
If one pares down the full set of indices $\{n_\star,\ n_\star+1,\ n_\star+2, \dots\}$ to a subset

$$\{n_1,\ n_2,\ n_3, \dots\} \qquad \text{such that} \qquad n_\star \leq n_1 < n_2 < n_3 < \dots$$

then we call the corresponding "thinned out" sequence $(x_{n_j})_{j \in \mathbb{N}}$ a **subsequence** $(x_n)_{n \geq n_\star}$.
If this subset of indices is finite, i.e., we have

$$n_\star \leq n_1 < n_2 < \dots < n_K \qquad \text{for some suitable } K \in \mathbb{N}\,,$$

then we call $(x_{n_j})_{j=1}^{K}$ a **finite subsequence** of the original sequence. □

Definition 2.26 (Indexed items)**.** Given is an expression of the form

$$a_i\,.$$

We say that a_i is **indexed by** or **subscripted by** or **tagged by** i. We call i the **index** or **subsript** of a_i, and we call a_i an **indexed item** . □

Definition 2.27 (Indexed families)**.** Let J and X be nonempty sets such that

each $i \in J$ is associated with exactly one indexed item $x_i \in X$.

We write $(x_i)_{i \in J}$ for this collection of indexed items and call it an **indexed family** or **family** in X with **index set** J. The indexed items $x_\jmath$ are called the **members of the family**. □

A family $(x_i)_{i \in J}$ can be interpreted as the function

$$x(\cdot) : J \longrightarrow X\,; \quad i \mapsto x(i) := x_i\,.$$

Families in X are functions with domain = index set = J and codomain X .

2.5 Cartesian Products

Definition 2.28 (Preliminary definition: Cartesian Product)**.** Let X and Y be two sets The set

$$X \times Y := \{(x,y) : x \in X, y \in Y\} \tag{2.19}$$

is called the **cartesian product** of X and Y.
Note that the order is important: (x,y) and (y,x) are different unless $x = y$.
We write X^2 as an abbreviation for$X \times X$.
This definition generalizes to more than two sets as follows: Let $X_1, X_2, \dots, X_n$ be sets. The set

$$X_1 \times X_2 \cdots \times X_n := \{(x_1, x_2, \dots, x_n) : x_j \in X_j \text{ for each } j = 1, 2, \dots n\} \tag{2.20}$$

is called the cartesian product of $X_1, X_2, \dots, X_n$.
We write X^n as an abbreviation for$X \times X \times \cdots \times X$. □

2.6 Arbitrary Unions and Intersections

Definition 2.29 (Arbitrary unions and intersections)**.** **(A)** For a (nonempty) set of sets $\mathscr{A}$, let

$$\bigcup_{B \in \mathscr{A}} B := \bigcup \left[B : B \in \mathscr{A}\right] := \{x : x \in B \text{ for at least one } B \in \mathscr{A}\}, \tag{2.21}$$

$$\bigcap_{B \in \mathscr{A}} B := \bigcap \left[B : B \in \mathscr{A}\right] := \{x : x \in B \text{ for each } B \in \mathscr{A}\}. \tag{2.22}$$

We call $\bigcup_{B \in \mathscr{A}} B$ the **union** and $\bigcap_{B \in \mathscr{A}} B$ the **intersection** of the members of $\mathscr{A}$

(B) For a family $(A_i)_{i \in I}$ of sets A_i, let

$$\bigcup_{i \in I} A_i := \bigcup \left[A_i : i \in I\right] := \{x : x \in A_i \text{ for at least one } i \in I\}, \tag{2.23}$$

$$\bigcap_{i \in I} A_i := \bigcap \left[A_i : i \in I\right] := \{x : x \in A_i \text{ for each } i \in I\}. \tag{2.24}$$

We call $\bigcup_{i \in I} A_i$ the **union** and $\bigcap_{i \in I} A_i$ the **intersection** of the family $(A_i)_{i \in I}$.

(C) Let $\mathscr{A}$ be a nonempty set of sets, let $(A_i)_{i\in I}$ be a family of sets.

We call the members of $\mathscr{A}$ **disjoint** , also **mutually disjoint**, if $A, A' \in \mathscr{A}$ and $A \neq A'$ implies $A \cap A' = \emptyset$. We call the family $(A_i)_{i\in I}$ **disjoint** , also **mutually disjoint**, if $A_i \cap A_j = \emptyset$ for all $i, j \in J$ such that $i \neq j$.

As done previously, we allow the use of $\biguplus$ instead of $\bigcup$ to indicate disjoint unions:

(2.25)
$$\biguplus_{B\in\mathscr{A}} B := \bigcup_{B\in\mathscr{A}} B\,, \qquad \biguplus_{i\in I} A_i := \bigcup_{i\in I} A_i\,.$$

(D) Assume that there is Ω, $\mathscr{A}$ such that $\mathscr{A} \subseteq \Omega$ and the members of $\mathscr{A}$ are disjoint.

If $\Omega = \biguplus_{B\in\mathscr{A}} B$, then we call $\mathscr{A}$ a **partition** of Ω.

Assume that there is Ω, $(A_i)_{i\in I}$ such that $A_j \subseteq \Omega$ for all $j \in J$ is a disjoint family.

If $\Omega = \biguplus_{i\in I} A_i$, then we call $(A_i)_{i\in I}$ a **partition** of Ω.

Note that being a partition means that each $x \in \Omega$ belongs to exactly one member of $\mathscr{A}$ (of $(A_i)_{i\in I}$ in case of a family).

Since sequences are special kinds of families with index sets

$$[n_*, \infty[_{\mathbb{Z}} = \{ n_*, n_* + 1, n_* + 2, \dots \},$$

it is natural to write

(2.26)
$$\bigcup_{i=n_*}^{\infty} A_i := \bigcup_{i\in[n_*,\infty[_{\mathbb{Z}}} A_i\,, \qquad \bigcap_{i=n_*}^{\infty} A_i := \bigcap_{i\in[n_*,\infty[_{\mathbb{Z}}} A_i\,, \quad \square$$

2.7 Proofs by Induction and Definitions by Recursion

Remark 2.3.

Principle of Mathematical Induction

Assume that for each integer $k \geq k_0$ there is an associated statement $P(k)$ such that the following is valid:

A. Base case. The statement $P(k_0)$ is true.

B. Induction Step. Assuming that $P(k)$ is true ("**Induction Assumption**"), it can be shown that $P(k+1)$ also is true.

It then follows that $P(k)$ is true for **each** $k \geq k_0$.

Proposition 2.6 (Distributivity of unions and intersections for finitely many sets)**.** *Let $A_1, A_2, \dots$ and B be sets. If $n \in \mathbb{N}$ then*

$$\left(\bigcup_{j=1}^{n} A_j\right) \cap B = \bigcup_{j=1}^{n}(A_j \cap B), \tag{2.27}$$

$$\left(\bigcap_{j=1}^{n} A_j\right) \cup B = \bigcap_{j=1}^{n}(A_j \cup B). \tag{2.28}$$

Proposition 2.7 (The Triangle Inequality for n real numbers)**.** *Let $n \in \mathbb{N}$ such that $n \geq 2$. Let $a_1, a_2, \dots, a_n \in \mathbb{N}$. Then*

$$|a_1 + a_2 + \ldots + a_n| \leq |a_1| + |a_2| + \ldots + |a_n| \tag{2.29}$$

2.8 Some Preliminaries From Calculus

3 The Axiomatic Method

3.1 Semigroups and Groups

Definition 3.1 (Semigroups and monoids). ⋆

Given is a nonempty set S with a binary operation $\diamond$,
i.e. an "assignment rule" $(s,t) \mapsto s \diamond t$ which assigns to any two elements $s,t \in S$ a third element $u := s \diamond t \in S$. [1] The pair $(S, \diamond)$ is called a **semigroup** if the operation $\diamond$ satisfies

(3.1) **associativity:** $(s \diamond t) \diamond u = s \diamond (t \diamond u)$ for all $s,t,u \in S$.

A semigroup for which there exists in addition a **neutral element** with respect to the operation$(s,t) \mapsto s \diamond t$, i.e., some $e \in S$ such that

(3.2) $$s \diamond e = e \diamond s = s \text{ for all } s \in S$$

is called a **monoid**.
We can write S instead of $(S, \diamond)$ if it is clear which binary operation on S is represented by $\diamond$.
□

Proposition 3.1. *Let A be a nonempty set and let $S := \{f : f$ is a function $A \to A\}$.*
We define a binary operation $\circ$ on S as follows.

$$(f,g) \mapsto g \circ f$$

assigns to two functions $f,g : A \to A$ the function

$$g \circ f : A \to A; \qquad x \mapsto g \circ f(x) := g\big(f(x)\big).$$

$(S, \circ)$ is a monoid.

Theorem 3.1 (Uniqueness of the neutral element in monoids).
Let $(S, \diamond)$ be a monoid and let $e, e' \in S$ such that both

(3.3) $$s \diamond e = e \diamond s = s$$
(3.4) $$s \diamond e' = e' \diamond s = s$$

for all $s \in S$. Then $e = e'$.

Definition 3.2 (Groups and Abelian groups)**.** Let $(G, \diamond)$ be a monoid with neutral element e which satisfies the following: For each $g \in G$ there exists some $g' \in G$ such that

$$g \diamond g' = g' \diamond g = e \quad \text{for all } g \in G. \tag{3.5}$$

We call such a g' an **inverse element**. of g, and we then call $(G, \diamond)$ a **group**.
Assume moreover that the operation $\diamond$ satisfies

$$\textbf{commutativity:} \quad g \diamond h = h \diamond g \quad \text{for all } g, h \in G. \tag{3.6}$$

Then G is called a **commutative group** or **abelian group**. We write G instead of $(G, \diamond)$ if it is clear which binary operation on G is represented by $\diamond$. □

Groups $(G, \diamond)$ are characterized as follows.

(a)	If $g, h \in G$ then $g \diamond h \in G$	**binary operation**
(b)	If $g, h, k \in G$ then $(g \diamond h) \diamond k = g \diamond (h \diamond k)$	**associativity**
(c)	There exists $e \in G$ such that $g \diamond e = e \diamond g = g$ for all $g \in G$	**neutral element**
(d)	For each $g \in G$ there exists $g' \in G$ such that $g \diamond g' = g' \diamond g = e$	**inverse element**

G is a **commutative group (abelian group)** if, in addition,

(e)	$g \diamond h = h \diamond g$ for all $g, h \in G$	**commutativity**

Theorem 3.2 (Uniqueness of the inverse in groups)**.** *Let $(G, \diamond)$ be a group and let $g \in G$. Assume that there exists besides g' another $g'' \in G$ which satisfies (3.5). Then $g'' = g'$.*

Definition 3.3 (inverse element g^{-1})**.** It is customary to write g^{-1} for the unique element of G that is associated with the given $g \in G$ by means of (3.5). We call g^{-1} the inverse element of g rather than an inverse element of g. □

Proposition 3.2. *Let $(G, \diamond)$ be a group with neutral element e. Let $g, h \in G$. Then*

$$\left(g^{-1}\right)^{-1} = g, \tag{3.7}$$
$$(h \diamond g)^{-1} = g^{-1} \diamond h^{-1}. \tag{3.8}$$

Proposition 3.3. ⋆

Let $(G, \diamond)$ be a group. Let $g, h \in G$. Then

$$h \diamond g^{-1} = (g \diamond h^{-1})^{-1}. \tag{3.9}$$

Proposition 3.4 (B/G prop.1.9 and B/G prop.8.10)**.** *Let $g, h, h' \in (G, \diamond)$. If $g \diamond h = g \diamond h'$ then $h = h'$.*

Proposition 3.5. *Let G be the set of all polynomials of degree 1. In other words,*

$$G = \{f : \mathbb{R} \to \mathbb{R} : f(x) = ax + b \text{ for some } a, b \in \mathbb{R} \text{ where } a \neq 0\}$$

*This is the set of functions whose graph is a straight line in the x, y–plane, which is parallel neither to the x–axis, nor to the y–axis. As in example **??**, let $(f, g) \mapsto g \circ f$ be defined as $g \circ f(x) = g(f(x))$. Then $(G, \circ)$ is a group.*

Definition 3.4 (Linear functions on $\mathbb{R}$)**.** ⋆
A function $f : \mathbb{R} \to \mathbb{R}$ is a **linear function on** $\mathbb{R}$ if the following is true for all $x, y, \lambda \in \mathbb{R}$:

$$f(x + y) = f(x) + f(y) \quad \textbf{(additivity)}, \tag{3.10}$$
$$f(\lambda x) = \lambda f(x) \quad \textbf{(homogeneity)}. \tag{3.11}$$

□

Theorem 3.3.
Let $f : \mathbb{R} \to \mathbb{R}$. Then f is linear if and only if there exists $a \in \mathbb{R}$ such that $f(x) = ax$ for all $x \in \mathbb{R}$.

Definition 3.5 (Subgroup)**.** ⋆ Let $(G, \diamond)$ be a group and $H \subseteq G$.
We call $(H, \diamond)$ a **subgroup** of G if the following is true:

(3.12) H is not empty,
(3.13) if $h, h' \in H$ then $h \diamond h' \in H$,
(3.14) if $h \in H$ then its inverse element h^{-1} (in G!) belongs to H.

We also write H for $(H, \diamond)$, if there is no confusion about the nature of "$\diamond$". □

Proposition 3.6. *Subgroups are groups.*

Proposition 3.7. *Let $(G, \circ)$ be the set of all polynomials of degree 1 with function composition, i.e.,*

$$G = \{\mathbb{R} \xrightarrow{f} \mathbb{R} : f(x) = ax + b, \text{ for some } a, b \in \mathbb{R} \text{ such that } a \neq 0\},$$
$$g \circ f : x \mapsto g \circ f(x) = g(f(x)).$$

Further, let

$$H := \{\mathbb{R} \xrightarrow{f} \mathbb{R} : f(x) = ax, \text{ for some nonzero } a \in \mathbb{R}\}.$$

Then $(H, \circ)$ is a subgroup of $(G, \circ)$.

Proposition 3.8. *The intersection of an arbitrary collection of subgroups is a subgroup.*

Definition 3.6 (Homomorphisms and isomorphisms)**.** Let $(G, \diamond)$ and $(H, \bullet)$ be groups with neutral elements e_G and e_H and let us write g^{-1} and h^{-1} for the inverses (in the sense of def. 3.3 on p.21.
Let $\varphi : (G, \diamond) \to (H, \bullet)$ be a function which satisfies the following:

$$\varphi(g_1 \diamond g_2) = \varphi(g_1) \bullet \varphi(g_2). \tag{3.15}$$

Then we call φ a **homomorphism**, more specifically, a **group homomorphism**, from the group $(G, \diamond)$ to the group $(H, \bullet)$.
Let $\psi : (H, \bullet) \to (G, \diamond)$ be a group homomorphism from $(H, \bullet)$ to $(G, \diamond)$ such that φ and ψ are inverse to each other. We call such a bijective homomorphism an **isomorphism**, and we call the groups $(G, \diamond)$ and $(H, \bullet)$ **isomorphic**.

For bijectivity, see Definition 5.12 on p.42). □

Theorem 3.4. *Let $(G, \diamond)$ and $(H, \bullet)$ be two groups and let $\varphi : (G, \diamond) \to (H, \bullet)$ be a homomorphism. Let e_G be the neutral element of G and e_H be the neutral element of H. Then*

- ***(a)*** $\varphi(e_G) = e_H$,
- ***(b)*** *Let $g \in G$. Then* $\varphi(g^{-1}) = \big(\varphi(g)\big)^{-1}$,
- ***(c)*** *Direct images of subgroups of G are subgroups of H.*
- ***(d)*** *Preimages of subgroups of G are subgroups of H.*

Theorem 3.5. ⋆ *Let $(G, \diamond)$ and $(H, \bullet)$ be two groups and let $\varphi : (G, \diamond) \to (H, \bullet)$ be a homomorphism which possesses an inverse.*
Then $\varphi^{-1} : H \to G$ also is a homomorphism and thus φ is an isomorphism

3.2 Commutative Rings and Integral Domains

Definition 3.7 (Commutative rings with unit)**.** ⋆ Let R be a nonempty set with two binary operations

$$\oplus : (a, b) \mapsto a \oplus b, \text{ called } \textbf{addition}, \quad \text{and} \quad \odot : (a, b) \mapsto a \odot b, \text{ called } \textbf{multiplication},$$

which assign to any two elements $a, b \in R$ uniquely determined $a \oplus b \in R$ and $a \odot b \in R$ such that the following holds:

- **(a)** $(R, \oplus)$ is an abelian (i.e., commutative) group; we denote the neutral element for addition by 0 and the inverse element of $a \in R$ for addition by $\ominus a$.
- **(b)** $(R, \odot)$ is a commutative monoid, i.e., a monoid for which $a \odot b = b \odot a$ for all $a, b \in R$. We denote the neutral element with respect to multiplication by 1.
- **(c)** Multiplication is **distributive** over addition:

  (3.16) $$a \odot (b \oplus c) \; = \; (a \odot b) \oplus (a \odot c) \quad \text{for all } a, b, c \in R.$$

- **(d)** $1 \neq 0$.

The triplet $(R, \oplus, \odot)$ is called a **commutative ring with unit**. We may write R instead of $(R, \oplus, \odot)$ if it is clear which binary operations on R are represented by $\oplus$ and by $\odot$. □

Notation 3.1 (Notation Alert for Commutative Rings With Unit)**.**

- **(a)** It is customary to write ab instead of $a \odot b$ if this does not give rise to confusion.
- **(b)** Multiplication has precedence over (binds stronger than) addition: $a \odot b \oplus c$ means $(a \odot b) \oplus c$, not $a \odot (b \oplus c)$.
- **(c)** Let $a, b, \in R$. Recall from thm.3.1 and thm.3.2 that not only the neutral elements 0 and 1 but also the additive inverse $\ominus b$ are uniquely determined. Accordingly, we can define another binary operation, $\ominus$, on $(R, \oplus, \odot)$ as follows:

  (3.17) $$a \ominus b \; := \; a \oplus (\ominus b).$$

  We call $a \ominus b$ the **difference** of a and b. □

Definition 3.8 (Translation and dilation of sets). ⋆
Let $R = (R, \oplus, \odot)$ be a commutative ring with unit and $A \subseteq R$. and $\alpha, b \in R$. We define

$$\lambda A \oplus b := \{\lambda a \oplus b : a \in A\}. \tag{3.18}$$

In particular, for $\lambda = \pm 1$, we obtain

$$A \oplus b = \{a \oplus b : a \in A\}, \tag{3.19}$$
$$\ominus A = \{\ominus a : a \in A\}. \quad \square \tag{3.20}$$

Proposition 3.9. *Let $(R, \oplus, \odot)$ be a nonempty set with two binary operations $\oplus$ and $\odot$ which satisfies **(a)**, **(b)**, **(c)** of Definition 3.7, i.e., R satisfies all conditions for a commutative ring with unit except that 1 and 0 need not be different elements of R. Then*

(a) $a \ominus a = 0$ *for all* $a \in R$,
(b) $a \odot 0 = 0$ *for all* $a \in R$.

Proposition 3.10.

(a) *The set $R := \{0\}$ satisfies conditions **(a)**, **(b)**, **(c)** of Definition 3.7,*
(b) *Let $(R, \oplus, \odot)$ be a nonempty set with two binary operations $\oplus$ and $\odot$ which satisfies **(a)**, **(b)**, **(c)** of Definition 3.7. Then the following is true: $1 = 0$ if and only if $R = \{0\}$*
.

Definition 3.9 (Zero Divisors and Cancellation Rule). Let $(R, \oplus, \odot)$ be a commutative ring with unit.

(a) If $a, b \in R$ such that $a \neq 0$ and $b \neq 0$ and $a \odot b = 0$ then we call a and b **zero divisors**.
(b) We say that the **cancellation rule** holds in R if the following is true
for all $a, b, c \in R$ such that $a \neq 0$:

$$\text{If } a \odot b = a \odot c \text{ then } b = c. \quad \square \tag{3.21}$$

Definition 3.10 (Integral domains). Let $(R, \oplus, \odot)$ be a commutative ring with unit which satisfies the

- **no zero divisors condition**: If $a, b \in R$ such that $a \odot b = 0$ then $a = 0$ or $b = 0$ (or both are zero).

The triplet $(R, \oplus, \odot)$ is called an **integral domain**. □

Remark 3.1. Integral domains $(R, \oplus, \odot)$ are characterized as follows.

(a)	If $a, b \in R$ then $a \oplus b \in R$ and $a \odot b \in R$	**binary operations**
(b)	If $a, b, c \in R$ then $(a \oplus b) \oplus c = a \oplus (b \oplus c)$	**associativity of** $\oplus$
(c)	If $a, b, c \in R$ then $(a \odot b) \odot c = a \odot (b \odot c)$	**associativity of** $\odot$
(d)	If $a, b \in R$ then $a \oplus b = b \oplus a$	**commutativity of** $\oplus$
(e)	If $a, b \in R$ then $a \odot b = b \odot a$	**commutativity of** $\odot$
(f)	If $a, b, c \in R$ then $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$	**distributivity**
(g)	There exists $0 \in R$ such that $a \oplus 0 = a$ for all $a \in R$	**neutral element f.** $\oplus$
(h)	There exists $1 \in R$ such that $1 \neq 0$ and $a \odot 1 = a$ for all $a \in R$	**neutral element f.** $\odot$
(i)	For each $a \in R$ there exists $a' \in R$ such that $a \oplus a' = 0$	**inverse element f.** $\oplus$
(j)	If $a, b \in R$ such that $a \neq 0$ and $b \neq 0$ then $a \odot b \neq 0$	**no zero divisors**

Proposition 3.11. *Let $(R, \oplus, \odot)$ be a commutative ring with unit. Then R satisfies the No zero divisors condition if and only if the cancellation rule holds in R.*

Corollary 3.1. *A commutative ring with unit is an integral domain $\Leftrightarrow$ the cancellation rule holds.*

Proposition 3.12. *Each of the following algebraic structures is an integral domain:*

(a) $(\mathbb{Z}, +, \cdot)$: *the integers with addition and multiplication,*
(b) $(\mathbb{Q}, +, \cdot)$: *the rational numbers with addition and multiplication,*
(c) $(\mathbb{R}, +, \cdot)$: *the real numbers with addition and multiplication.*
(d) [2] $(\mathbb{C}, +, \cdot)$: *the complex numbers with addition and multiplication.*

3.3 Arithmetic in Integral Domains

Proposition 3.13 (B/G prop.1.6 and B/G prop.8.8)**.** *Let $a, b, c \in R$. Then* $(a \oplus b) \odot c = a \odot c \oplus b \odot c$.

Proposition 3.14 (B/G prop.1.7 and B/G prop.8.9). *Let $a \in R$. Then $0 \oplus a = a$ and $1 \odot a = a$.*

Proposition 3.15 (B/G prop.1.8). *Let $a \in R$. Then $(\ominus a) \oplus a = 0$.*

Proposition 3.16 (B/G prop.1.10 and B/G prop.8.11). *Let $a, b_1, b_2 \in R$. If $a \oplus b_1 = 0$ and $a \oplus b_2 = 0$ then $b_1 = b_2$.*

Proposition 3.17 (B/G prop.1.11 and B/G prop.8.12). *Let $a, b, c, d \in R$. Then*

- ***(a)*** $(a \oplus b)(c \oplus d) = (ac \oplus bc) \oplus (ad \oplus bd)$,
- ***(b)*** $a \oplus \big(b \oplus (c \oplus d)\big) = (a \oplus b) \oplus (c \oplus d) = \big((a \oplus b) \oplus c\big) \oplus d$,
- ***(c)*** $a \oplus (b \oplus c) = (c \oplus a) \oplus b$,
- ***(d)*** $a(bc) = c(ab)$,
- ***(e)*** $a\big(b \oplus (c \oplus d)\big) = (ab \oplus ac) \oplus ad$,
- ***(f)*** $\big(a(b \oplus c)\big)d = (ab)d \oplus a(cd)$.

Proposition 3.18. ★ *Let $a, b \in R$. Then $b \ominus a = \ominus(a \ominus b)$.*

Proposition 3.19 (B/G prop.1.12 and B/G prop.8.13). *Let $x \in R$ satisfy the following: For each $a \in R$ it is true that $a \oplus x = a$. Then $x = 0$.*

Proposition 3.20 (B/G prop.1.13 and B/G prop.8.14). *Let $x \in R$ satisfy the following: There exists (at least one) $a \in R$ such that $a \oplus x = a$. Then $x = 0$.*

Proposition 3.21 (B/G prop.1.14 and B/G prop.8.15). *Let $a \in R$. Then $a \odot 0 = 0 = 0 \odot a$.*

Proposition 3.22 (B/G prop.1.18 and B/G prop.8.16). *Let $x \in R$ satisfy the following: For each $a \in R$ it is true that $a \odot x = a$. Then $x = 1$.*

Proposition 3.23 (B/G prop.1.19 and B/G prop.8.17). *Let $x \in R$ satisfy the following: There exists (at least one) nonzero $a \in R$ such that $a \odot x = a$. Then $x = 1$.*

Proposition 3.24 (B/G prop.1.20 and B/G prop.8.18). *Let $a, b \in R$. Then $(\ominus a)(\ominus b) = ab$.*

Corollary 3.2 (B/G cor.1.21). $(\ominus 1)(\ominus 1) = 1$.

Proposition 3.25 (B/G prop.1.22 and B/G prop.8.19).

(a) *If $a \in R$ then $\ominus(\ominus a) = a$.*
(b) $\ominus 0 = 0$.

Proposition 3.26 (Unique Solutions of Linear Equations). *Let $(R, \oplus, \odot)$ be an integral domain and $a, b, y \in R$ such that $a \neq 0$. The equation $y = a \odot x \oplus b$ possesses at most one solution $x \in R$.*

Proposition 3.27 (B/G prop.1.23 and B/G prop.8.20).
Let $a, b \in R$. Then there exists one and only one $x \in R$ such that $a \oplus x = b$.

Proposition 3.28 (B/G prop.1.24 and B/G prop.8.21).
Let $x \in R$. If $x \odot x = x$ then $x = 0$ or $x = 1$.

Proposition 3.29 (B/G prop.1.25 and B/G prop.8.22). *Let $a, b \in R$. Then*

(a) $\ominus(a \oplus b) \ = \ (\ominus a) \oplus (\ominus b)$,
(b) $\ominus a \ = \ (\ominus 1)a$,
(c) $(\ominus a)b \ = \ a(\ominus b) \ = \ \ominus(ab)$.

Proposition 3.30 (B/G prop.1.26 and B/G prop.8.23). *Let $a, b \in R$. If $ab = 0$ then $a = 0$ or $b = 0$.*

Proposition 3.31 (B/G prop.1.27 and B/G prop.8.24). *Let $a, b, c, d \in R$. Then*

(a) $(a \ominus b) \oplus (c \ominus d) \;=\; (a \oplus c) \ominus (b \oplus d)$,
(b) $(a \ominus b) \ominus (c \ominus d) \;=\; (a \oplus d) \ominus (b \oplus c)$,
(c) $(a \ominus b)(c \ominus d) \;=\; (ac \oplus bd) \ominus (ad \oplus bc)$,
(d) $a \ominus b \;= c \ominus d$ *if and only if* $a \oplus d \;=\; b \oplus c$,
(e) $(a \ominus b)c \;=\; ac \ominus bc$.

3.4 Order Relations in Integral Domains

Definition 3.11 (Ordered Integral Domains). **I.** Let $(R, \oplus, \odot)$ be an integral domain. Assume there exists $P \subseteq R$ which satisfies the following:

(a) If $p_1, p_2 \in P$ then $p_1 \oplus p_2 \in P$,
(b) If $p_1, p_2 \in P$ then $p_1 \odot p_2 \in P$,
(c) $0 \notin P$,
(d) Let $a \in R$. Then at least one of the following is true: $a \in P$, $\ominus a \in P$, $a = 0$.

We call P a **positive cone** on the integral domain R.

II. We use P to define on R an "order relation" a < b as follows: Let $a, b \in R$. We define

(3.22) $a \;< b$ if and only if $b \ominus a \in P$ ("a **is less than** b"),
(3.23) $a \;\leq b$ if and only if $a < b$ or $a = b$, ("a **is less than or equal** b"),
(3.24) $a \;> b$ if and only if $b < a$, ("a **is greater than** b"),
(3.25) $a \;\geq b$ if and only if $b \leq a$. ("a **is greater than or equal** b"),

We say that $<$ is the **order induced by** P, and we call the quadruple $(R, \oplus, \odot, P)$ an **ordered integral domain**. Let $a \in R$. If $a \in P$ then we call a a **positive** element of R, and if $\ominus a \in P$ then we call a a **negative** element of R. If a is positive or zero then we call a **nonnegative**, and if a is negative or zero then we call a **nonpositive**. □

Proposition 3.32. *Each of the following algebraic structures is an ordered integral domain:*

(a) $(\mathbb{Z}, +, \cdot, \mathbb{N})$: *The integers with addition and multiplication: The positive cone is the subset of all natural numbers.*
(b) $(\mathbb{Q}, +, \cdot, \mathbb{Q}_{>0})$: *The rational numbers with addition and multiplication: The positive cone $\mathbb{Q}_{>0}$ is the subset of all fractions $\frac{m}{n}$. where both m, n are positive integers.* [3]
(c) $(\mathbb{R}, +, \cdot, \mathbb{R}_{>0})$: *The real numbers with addition and multiplication. The positive cone here is* $]0, \infty[$.

Notation: In this entire chapter we assume that a fixed ordered integral domain $(R, \oplus, \odot, P)$ is given and phrases such as "let $a \in R$" refer to elements of that integral domain. We further assume that order relations such as "$a < b$" and "$a \geq b$" refer to the order induced by the positive cone P.

Definition 3.12 (Intervals in Ordered Integral Domains)**.**
(A) For the following let $a, b \in (R, \oplus, \odot, P)$.
$[a, b]_R := \{x \in R : a \leq x \leq b\}$ is called the **closed interval** with endpoints a and b.
$]a, b[_R := \{x \in R : a < x < b\}$ is called the **open interval** with endpoints a and b.
$[a, b[_R := \{x \in R : a \leq x < b\}$ and $]a, b]_R := \{x \in R : a < x \leq b\}$ are called **half-open intervals** with endpoints a and b.

(B) We generalize the symbol "∞" from real numbers (see Definition 2.19 on p.13) to arbitrary ordered integral domains as follows. The symbol "∞" stands for an object which itself is not an element of $(R, \oplus, \odot, P)$ but is larger than any of its elements, and the symbol "$\ominus\infty$" stands for an object which itself is not an element of $(R, \oplus, \odot, P)$ but is smaller than any of its elements. We thus have $\ominus\infty < x < \infty$ for any $x \in R$. We write $\overset{\oplus}{\ominus}\infty$ when we mean "either $\oplus\infty$ or $\ominus\infty$."
We now define

$$]\ominus\infty, a]_R := \{x \in R : x \leq a\} \qquad]\ominus\infty, a[_R := \{x \in R : x < a\}$$
$$]a, \infty[_R := \{x \in R : x > a\} \qquad [a, \infty[_R := \{x \in R : x \geq a\}. \quad \square$$

Proposition 3.33 (B/G prop.2.2 and B/G prop.8.27)**.** *Let $a \in R$. Then **either** $a \in P$ **or** $\ominus a \in P$ or $a = 0$.*

Proposition 3.34 (B/G prop.2.13 and B/G prop.8.38)**.** *If $(R, \oplus, \odot, P)$ is an ordered integral domain, then* $P = \{ x \in R : x > 0 \}$.

Proposition 3.35 (B/G prop.2.3 and B/G prop.8.28)**.** *The multiplicative unit 1 of R belongs to P.*

Proposition 3.36. *If $a \in R$ then $a \oplus 1 > a$.*

Corollary 3.3. $1 > 0$.

Proposition 3.37 (B/G prop.2.4 and B/G prop.8.29). *Let $a, b, c \in R$.*

(3.26) $$\text{If} \quad a < b \quad \text{and} \quad b < c, \quad \text{then} \quad a < c.$$

Proposition 3.38. *Let $a, b, c \in R$.*

(3.27) $$\text{If} \quad a \leq b \quad \text{and} \quad b \leq c, \quad \text{then} \quad a \leq c.$$

Proposition 3.39 (B/G prop.2.5 and B/G prop.8.30). *For each $a \in R$ there exists $p \in P$ such that $a \oplus p > a$.*

Proposition 3.40 (B/G prop.2.6 and B/G prop.8.31). *Let $a, b \in R$. If $a \leq b \leq a$ then $a = b$.*

Proposition 3.41 (B/G prop.2.7 and B/G prop.8.32). *Let $a, b, c, d \in R$. Then*

(a) *If $a < b$ then $a \oplus c < b \oplus c$.*
(b) *If $a < b$ and $(c < d)$ then $a \oplus c < b \oplus d$.*
(c) *If $0 < a < b$ and $0 < c \leq d$ then $ac < bd$.*
(d) *If $0 < a \leq b$ and $0 < c \leq d$ then $ac \leq bd$.*
(e) *If $a < b$ and $c < 0$ then $bc < ac$.*

Proposition 3.42 (B/G prop.2.8 and B/G prop.8.33). *Let $a, b \in R$. Then either $a < b$ or $a = b$ or $a > b$.*

Proposition 3.43. *Let $a, b \in R$. Then*

(a) *$ab > 0 \Leftrightarrow a, b > 0$ or $a, b < 0$,*
(b) *$ab < 0 \Leftrightarrow$ [either $a > 0$ and $b < 0$] or [$a < 0$ and $b > 0$]*
(c) *$ab = 0 \Leftrightarrow a = 0$ or $b = 0$*

Proposition 3.44 (B/G prop.2.9 and prop.8.34). *Let $a \in R$. If $a \neq 0$ then $a^2 \in P$.*

Proposition 3.45 (B/G prop.2.10 and B/G prop.8.35). *The equation $x^2 = \ominus 1$ has no solution (in R).*

Proposition 3.46 (B/G prop.2.11 and B/G prop.8.36). *Let $a \in R$ and $p \in P$. If $ap \in P$, then $a \in P$.*

Proposition 3.47 (B/G prop.2.12 and B/G prop.8.37). *Let $a, b, c \in R$. Then*

(a) $\ominus a < \ominus b$ *if and only if* $a > b$.
(b) *If* $c > 0$ *and* $ac < bc$ *then* $a < b$.
(c) *If* $c < 0$ *and* $ac < bc$ *then* $b < a$.
(d) *If* $a \leq b$ *and* $0 \leq c$ *then* $ac \leq bc$.

Definition 3.13 (Absolute value). For an element x of the ordered integral domain R, we define its **absolute value** as

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ \ominus x & \text{if } x < 0. \end{cases} \qquad \square$$

Proposition 3.48 (Generalization of B/G prop.10.5). *Let $x, y \in P \cup \{0\}$, i.e., $x, y \geq 0$. Then*

(a) $x \leq y$ *if and only if* $x^2 \leq y^2$,
(b) $x = y$ *if and only if* $x^2 = y^2$,
(c) $x < y$ *if and only if* $x^2 < y^2$.

Proposition 3.49 (B/G prop.10.6). *Let $a \in R$. Then $|a|^2 = a^2$.*

Proposition 3.50 (B/G prop.10.7). *Let $a, b \in R$. Then $|a| < |b| \;\Leftrightarrow\; a^2 < b^2$.*

Proposition 3.51 (B/G prop.10.8). *Let $a, b \in R$. Then the following holds:*

(a) $|a| = 0$ *if and only if* $a = 0$,
(b) $|ab| = |a| \odot |b|$,
(c) $\ominus|a| \leq a \leq |a|$,
(d) $|a \oplus b| \leq |a| \oplus |b|$,
(e) *if* $\ominus b < a < b$ *then* $|a| < b$, *in particular,* $b \geq 0$.

Proposition 3.52 (B/G prop.10.10)**.** *If $a, b, c \in R$, then*

(a) $|a \ominus b| = 0 \Leftrightarrow a = b$,
(b) $|a \ominus b| = |b \ominus a|$,
(c) $|a \ominus b| \leq |a \ominus c| \oplus |c \ominus b|$,
(d) $|a \ominus b| \geq \big||a| \ominus |b|\big|$.

Proposition 3.53. *This proposition is similar to prop.3.51**(e)**.*
*Let $a, b \in R$ such that both **#1)** $\ominus a \leq b$ and **#2)** $a \leq b$. Then $|a| \leq b$.*

3.5 Minima, Maxima, Infima and Suprema in Ordered Integral Domains

Definition 3.14 (Upper and lower bounds, maxima and minima)**.** Let $A \subseteq R$ and let $l, u \in R$.

(a) We call l a **lower bound** of A if $l \leq a$ for all $a \in A$.
(b) We call u an **upper bound** of A if $u \geq a$ for all $a \in A$.
(c) We call A **bounded above** if this set has an upper bound.
(d) We call A **bounded below** if A has a lower bound.
(e) We call A **bounded** if A is both bounded above and bounded below.
(f) A **minimum** (min) of A is a lower bound l of A such that $l \in A$.
(g) A **maximum** (max) of A is an upper bound u of A such that $u \in A$. □

Proposition 3.54. *Let $A \subseteq R$. If A has a maximum or a minimum, then it is unique.*

Definition 3.15. Let $A \subseteq R$. If A possesses a minimum, we write

$$\min(A) \text{ or } \min A$$

for this uniquely determined element of R. Likewise, if A possesses a maximum, we write

$$\max(A) \text{ or } \max A$$

for that uniquely determined element of R. □

Definition 3.16. ⋆ Let $A \subseteq R$. We define

$$\begin{aligned} A_{lowb} &:= \{l \in R : l \text{ is lower bound of } A\} \\ A_{uppb} &:= \{u \in R : u \text{ is upper bound of } A\}. \end{aligned} \tag{3.28}$$ □

Definition 3.17 (Infimum and supremum in an ordered integral domain)**.**
Let A be a nonempty subset of R.

(a) If $\max(A_{\ell owb})$ exists then it is unique by prop.3.54. We write $\inf(A)$ or g.l.b.(A) for $\max(A_{\ell owb})$ and call this number the **infimum** or **greatest lower bound** of A.
(b) If $\min(A_{uppb})$ exists then it is unique by prop.3.54. We write $\sup(A)$ or l.u.b.(A) for $\min(A_{uppb})$ and call this element of R the **supremum** or **least upper bound** of A. □

Notation 3.2. Notational conveniences:

(a) We may drop the parentheses in expressions like $\max(A)$, $\sup(\{f(x) : x \in B\})$ (here $f : X \to R$ is a function which takes values in an ordered integral domain R and where $B \subseteq X$), etc., if this does not lead to any confusion. We also can write the above as $\max A$ and $\sup\{f(x) : x \in B\}$.
(b) If A consists of two elements $x, y \in R$, i.e., $A = \{x, y\}$ then it is customary to write $\max(x, y),\ \min(x, y),\ \sup(x, y),$ and $\inf(x, y)$. □

Proposition 3.55. *Let $A \subseteq R$. If A has a maximum then it also has a supremum, and $\max(A) = \sup(A)$. Likewise, if A has a minimum then it also has an infimum, and $\min(A) = \inf(A)$.*

Proposition 3.56. *Let $\emptyset \neq A \subseteq B \subseteq R$.*

(a) *If both A and B possess an infimum (resp., supremum) then $\inf(A) \geq \inf(B)$ (resp., $\sup(A) \leq \sup(B)$).*
(b) *If both A and B possess a minimum (resp., maximum) then $\min(A) \geq \min(B)$ (resp., $\max(A) \leq \max(B)$).*
(c) *If both A and B possess a minimum (resp., maximum) and $\min(B) \notin A$ (resp., $\max(B) \notin A$) then $\min(A) > \min(B)$ (resp., $\max(A) < \max(B)$).*

Definition 3.18 (Supremum and Infimum of unbounded and empty sets). ⋆
Let $A \subseteq R$. If A is not bounded above, we define

$$\sup A = \infty \tag{3.29}$$

If A is not bounded below, we define

$$\inf A = \ominus\infty \tag{3.30}$$

Finally, we define

$$\sup \emptyset = \ominus\infty, \qquad \inf \emptyset = \oplus\infty. \tag{3.31}$$
□

Proposition 3.57. *Let $A \subseteq B \subseteq R$.*
(a) *If $\inf(A)$ and $\inf(B)$ both exist then $\inf(A) \geq \inf(B)$.*
(b) *If $\sup(A)$ and $\sup(B)$ both exist then $\sup(A) \leq \sup(B)$.*

Proposition 3.58. *Let $A \subseteq R$ and $x \in R$. Then*

$$x \leq a \text{ for all } a \in A \quad \Leftrightarrow \quad \ominus x \geq a' \text{ for all } a' \in \ominus A, \tag{3.32}$$
$$x \in A_{lowb} \quad \Leftrightarrow \quad \ominus x \in (\ominus A)_{uppb}, \tag{3.33}$$
$$\ominus A_{lowb} \quad = \quad (\ominus A)_{uppb}, \tag{3.34}$$
$$x \geq a \text{ for all } a \in A \quad \Leftrightarrow \quad \ominus x \leq a' \text{ for all } a' \in \ominus A, \tag{3.35}$$
$$x \in A_{uppb} \quad \Leftrightarrow \quad \ominus x \in (\ominus A)_{lowb}, \tag{3.36}$$
$$\ominus A_{uppb} \quad = \quad (\ominus A)_{lowb}, . \tag{3.37}$$

Proposition 3.59. *Let $\emptyset \neq A \subseteq R$. If the maximum of A_{lowb} exists, the following holds true: A has lower bounds, $\ominus A$ has lower bounds, the minimum of $(\ominus A)_{uppb}$ exists, and we have*

$$\ominus \max(A_{lowb}) \ = \ \min\big((\ominus A)_{uppb}\big), \tag{3.38}$$
$$\ominus \min(A_{uppb}) \ = \ \max\big((\ominus A)_{lowb}\big). \tag{3.39}$$

Corollary 3.4. *The following equations are to be understood in the sense that if the item on the left exists and vice versa, and both sides then are equal.*

$$\ominus \inf(A) \ = \ \sup(\ominus A), \tag{3.40}$$
$$\ominus \sup(A) \ = \ \inf(\ominus A), \tag{3.41}$$
$$\ominus \min(A) \ = \ \max(\ominus A). \tag{3.42}$$
$$\ominus \max(A) \ = \ \min(\ominus A), \tag{3.43}$$

Proposition 3.60. *Let a, b be nonnegative elements of R. Then*

$$|b \ominus a| \ \leq \ \max(a, b), \quad \text{i.e.,} \tag{3.44}$$
$$\ominus \max(a, b) \ \leq b \ominus a \ \leq \ \max(a, b). \tag{3.45}$$

Corollary 3.5. *Let $a, b, c \in R$ such that $0 \leq a, b < c$. Then*

$$\ominus c \; < b \ominus a \; < \; c. \tag{3.46}$$

4 Logic ⋆

5 Relations, Functions and Families

5.1 Cartesian Products and Relations

Definition 5.1 (Cartesian Product of Two Sets)**.** The **cartesian product** of two sets A and B is

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

i.e., it consists of all pairs (a, b) with $a \in A$ and $b \in B$.
Let $(a_1, b_1), (a_2, b_2) \in A \times B$. We say they are **equal**, and we write $(a_1, b_1) = (a_2, b_2)$ if and only if $a_1 = a_2$ and $b_1 = b_2$.
As a shorthand, we abbreviate $A^2 := A \times A$.

It follows from this definition of equality that the pairs (a, b) and (b, a) are different unless $a = b$. In other words, the order of a and b is important. We express this by saying that the cartesian product consists of **ordered pairs**. □

Definition 5.2 (Relation)**.** Let X and Y be two sets and $R \subseteq X \times Y$ a subset of their cartesian product $X \times Y$. We call R a **relation** on (X, Y). A relation on (X, X) is simply called a relation on X. If $(x, y) \in R$ we say that x **and** y **are related** and we usually write xRy instead of $(x, y) \in R$.
A relation on X is
- **(a)** **reflexive** if xRx for all $x \in X$,
- **(b)** **symmetric** if x_1Rx_2 implies x_2Rx_1 for all $x_1, x_2 \in X$,
- **(c)** **transitive** if x_1Rx_2 and x_2Rx_3 implies x_1Rx_3 for all $x_1, x_2, x_3 \in X$,
- **(d)** **antisymmetric** if x_1Rx_2 and x_2Rx_1 implies $x_1 = x_2$ for all $x_1, x_2 \in X$. □

Definition 5.3 (Equivalence relations and equivalence classes)**.** Let R be a relation on a set X.

(a) If R is • reflexive, • symmetric, • transitive, we call R an **equivalence relation** on X.

(b) For an equivalence relation R it is customary to write $x \sim x'$ rather than xRx' (or $(x, x') \in R$). We say in this case that x and x' are **equivalent**.

(c) Given is an equivalence relation "$\sim$" on a set X. For $x \in X$ let

(5.1) $$[x]_\sim := \{x' \in X : x' \sim x\} = \{ \text{ all items equivalent to } x \}.$$

We call $[x]_\sim$ the **equivalence class** of x. If it is clear from the context what equivalence relation is referred to then we can write $[x]$ instead of $[x]_\sim$. □

Proposition 5.1 (see [1] B/G prop.6.4 & B/G prop.6.5)**.** *Let "$\sim$" be an equivalence relation on a nonempty set X and $x, y \in X$. Then*

(a) $x \in [x]$,

(b) $x \sim y \Leftrightarrow [x] = [y]$,

(c) *either* $[x] = [y]$ *or* $[x] \cap [y] = \emptyset$.

Proposition 5.2 (see [1] B/G prop.6.6 for parts **(a)** and **(b)**)**.**

(a) *Let "$\sim$" be an equivalence relation on a nonempty set X and let $\mathscr{P}_\sim := \{[x] : x \in X\}$ be the set of all its equivalence classes. Then $\mathscr{P}_\sim$ is a partition of X.*

(b) *Conversely, let $\mathscr{P}$ be a partition of X and define a relation "$\sim_\mathscr{P}$" on X as follows: $x \sim_\mathscr{P} y \Leftrightarrow$ there is $P \in \mathscr{P}$ such that $x, y \in P$. Then $\sim_\mathscr{P}$ is an equivalence relation on X.*

(c) *Let "$\sim$" be an equivalence relation on X. Let $\mathscr{P}_\sim$ be the associated partition of its equivalence classes. Let "$\sim_{\mathscr{P}_\sim}$" be the equivalence relation associated with the partition $\mathscr{P}_\sim$. Then "$\sim_{\mathscr{P}_\sim}$" = "$\sim$" (i.e., both equivalence relations are equal as subsets of $X \times X$.*

(d) *Let $\mathscr{P}$ be a partition of X. Let $\sim_\mathscr{P}$ be the associated equivalence relation defined in part* **(b)**. *Let $\mathscr{P}_{\sim_\mathscr{P}}$ be the associated partition of its equivalence classes. Then $\mathscr{P}_{\sim_\mathscr{P}} = \mathscr{P}$.*

Definition 5.4 (Partial Order Relation)**.** Let R be a relation on a set X.

If R is reflexive, antisymmetric and transitive, it is called a **partial ordering** of X aka **partial order relation** on X. It is customary to write "$x \preceq y$" or "$y \succeq x$" rather than "xRy" for a partial ordering R. We say that "x **before** y" or "y **after** x".

We then call $(X, \preceq)$ a **partially ordered set** aka **POset**. □

Remark 5.1. The properties of a partial ordering can now be phrased as follows:

(5.2)	$x \preceq x$ for all $x \in X$	**reflexivity**
(5.3)	$x \preceq y$ and $y \preceq x \Rightarrow y = x$	**antisymmetry**
(5.4)	$x \preceq y$ and $y \preceq z \Rightarrow x \preceq z$	**transitivity** □

Definition 5.5 (Linear orderings). ⋆

(a) Let $(X, \preceq)$ be a nonempty POset, i.e., $\preceq$ is a partial ordering on X (see Definition 5.4 on p.37). We say that $\preceq$ is a **linear ordering** , also called a **total ordering** of X if and only if, for all x and $y \in X$ such that $x \neq y$, either $x \preceq y$ or $y \preceq x$. We call $(X, \preceq)$ a **linearly ordered set** or a **totally ordered set** set.

(b) Let $(X, \preceq)$ be a nonempty POset and $C \subseteq X$. C is a **chain** in X if $(C, \preceq)$ is linearly ordered (with the same ordering). □

Definition 5.6 (Inverse Relation). ⋆

Let X and Y be two sets and $R \subseteq X \times Y$ a relation on (X, Y). Let

$$R^{-1} := \{ (y, x) : (x, y) \in R \}.$$

Clearly R^{-1} is a subset of $Y \times X$ and hence a relation on (Y, X). We call R^{-1} the **inverse relation** of the relation R. □

5.2 Functions (Mappings) and Families

5.2.1 Some Preliminary Observations about Functions

5.2.2 Definition of a Function and Some Basic Properties

Definition 5.7 (Mappings (functions)). Given are two arbitrary nonempty sets X and Y and a relation Γ on (X, Y) (see 5.2 on p.36) which satisfies the following:

(5.5) for each $x \in X$ there exists exactly one $y \in Y$ such that $(x, y) \in \Gamma$.

We call the triplet $f(\cdot) := (X, Y, \Gamma)$ a **function** or **mapping** from X to Y. The set X is called the **domain** or **source** and Y is called the **codomain** or **target** of the mapping $f(\cdot)$. We will usually use the words "domain" and "codomain" in this document.

Usually mathematicians simply write f instead of $f(\cdot)$ We mostly follow that convention, but sometimes include the "$(\cdot)$" part to emphasize that a function rather than an "ordinary"' element of a set is involved. We write Γ_f or $\Gamma(f)$ if we want to stress that Γ is the relation associated with the function $f = (X, Y, \Gamma)$. Let $x \in X$. We write $f(x)$ for the uniquely determined $y \in Y$ such that $(x, y) \in \Gamma$. It is customary to write

$$f : X \to Y, \qquad x \mapsto f(x) \tag{5.6}$$

instead of $f = (X, Y, \Gamma)$ and we henceforth follow that convention. We abbreviate that to $f : X \to Y$ if it is clear or irrelevant how to compute $f(x)$ from x. We read the expression "$a \mapsto b$" as "a is assigned to b" or "a maps to b".

We call Γ the **graph** of the function f. Clearly

$$\Gamma = \Gamma_f = \Gamma(f) = \{(x, f(x)) : x \in X\}. \tag{5.7}$$

We refer to $\mapsto$ as the **maps to operator** or **assignment operator**.

Domain elements $x \in X$ are called **independent variables** or **arguments** and $f(x) \in Y$ is called the **function value** of x. The subset

$$f(X) := \{y \in Y : y = f(x) \text{ for some } x \in X\} = \{f(x) : x \in X\} \tag{5.8}$$

of Y is called the **range** or **image** of the function $f(\cdot)$.

We say "f maps X into Y" and "f maps the domain value x to the function value $f(x)$".

We say that two functions $f = (X, Y, \Gamma)$ and $f' = (X', Y', \Gamma')$ are **equal** if $X = X'$, $Y = Y'$, and $\Gamma = \Gamma'$. Note that $X = X'$ follows from $\Gamma = \Gamma'$ because

$$x \in X \;\Leftrightarrow\; (x, y) \in \Gamma \text{ for some (unique) } y \in Y \;\Leftrightarrow\; (x, y) \in \Gamma' \text{ for some } y \in Y \;\Leftrightarrow\; x \in X'. \quad \square$$

Figure 5.1 on p.40 illustrates the graph of a function as a subset of $X \times Y$.

Definition 5.8 (Function composition)**.** Given are three nonempty sets X, Y and Z and two functions $f : X \to Y$ and $g : Y \to Z$. Given $x \in X$ we know the meaning of the expression $g\big(f(x)\big)$:

$y := f(x)$ is the function value of x for the function f, i.e., the unique $y \in Y$ such that $(x, y) \in \Gamma_f$.

$z := g(y) = g\big(f(x)\big)$ is the function value of $f(x)$ for the function g, i.e., the unique $z \in Z$ such that $\big(f(x), z\big) = \big(f(x), g(f(x))\big) \in \Gamma_g$.

The set $\Gamma := \big\{\big(x, g(f(x)) : x \in X\big)\big\}$ is a relation on (X, Z) such that

$$\text{for each } x \in X \text{ there exists exactly one } z \in Z \text{, namely, } z = g\big(f(x)\big) \text{, such that } (x, z) \in \Gamma. \tag{5.9}$$

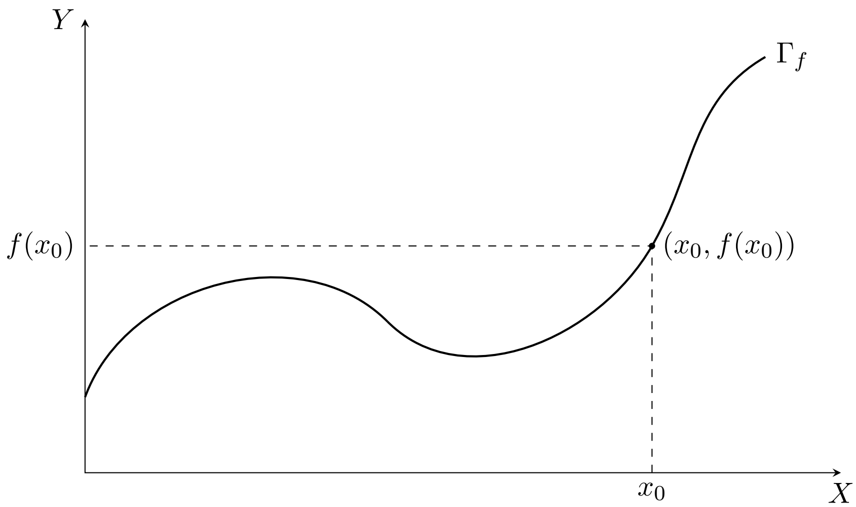
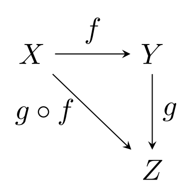


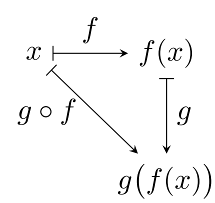
Figure 5.1: Graph of a function.

It follows that Γ is the graph of a function $h = (X, Z, \Gamma)$ with function values $h(x) = g\big(f(x)\big)$ for each $x \in X$. We call h the **composition** of f and g and we write $h = g \circ f$ ("g after f").

As far as notation is concerned it is OK to write either of $g \circ f(x)$ or $(g \circ f)(x)$. The additional parentheses may give a clearer presentation if f and/or g are defined by fairly complex formulas. □

(5.10) Function composition





Definition 5.9 (Constant functions)**.** Let Y be a nonempty set and $y_0 \in Y$. You can think of y_0 as a function from any nonempty set X to Y as follows:

$$y_0(\cdot) : X \to Y; \qquad x \mapsto y_0.$$

In other words, the function $y_0(\cdot)$ assigns to each $x \in X$ one and the same value y_0. We call such a function which only takes a single value a **constant function**.

The most important constant function is the **zero function** $0(\cdot)$ which maps any $x \in X$ to the number zero. We usually just write 0 for this function unless doing so would confuse the reader. □

Definition 5.10 (identity mapping)**.** Given any nonempty set X, we use the symbol id_X for the **identity** mapping defined as

$$id_X : X \to X; \qquad x \mapsto x.$$

We drop the subscript if it is clear what set is referred to. □

5.2.3 Examples of Functions

5.2.4 A First Look at Direct Images and Preimages of a Function

Definition 5.11. Let X, Y be two nonempty sets and $f : X \to Y$. We associate with f the functions

$$f : 2^X \to 2^Y; \qquad A \mapsto f(A) := \{f(a) : a \in A\}, \tag{5.11}$$

$$f^{-1} : 2^Y \to 2^X; \qquad B \mapsto f^{-1}(B) := \{x \in X : f(x) \in B\}. \tag{5.12}$$

We call $f : 2^X \to 2^Y$ the **direct image function** and $f^{-1} : 2^Y \to 2^X$ the **indirect image function** or **preimage function** associated with $f : X \to Y$.

For each $A \subseteq X$ we call $f(A)$ the **direct image** of A under f, and for each $B \subseteq Y$ we call $f^{-1}(B)$ the **indirect image** or **preimage** of B under f. □

Remark 5.2 (Notational conveniences II:)**.** In probability theory the following notation is also very common:
$\{f \in B\} := f^{-1}(B),\ \{f = y\} := f^{-1}\{y\}$.
Let $\mathscr{R}$ be either of $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$. Assume that the codomain of f is considered a subset of $\mathscr{R}$. Let $a, b \in \mathscr{R}$ such that $a < b$. We write $\{a \leq f \leq b\} := f^{-1}([a,b]_{\mathscr{R}}),\ \{a < f < b\} := f^{-1}(]a,b[_{\mathscr{R}}),$ $\{a \leq f < b\} := f^{-1}([a,b[_{\mathscr{R}}),\ \{a < f \leq b\} := f^{-1}(]a,b]_{\mathscr{R}}),\ \{f \leq b\} := f^{-1}(]-\infty,b]_{\mathscr{R}})$, etc. □

Proposition 5.3. *Some simple properties:*

$$f(\emptyset) = f^{-1}(\emptyset) = \emptyset \tag{5.13}$$

$$A_1 \subseteq A_2 \subseteq X \Rightarrow f(A_1) \subseteq f(A_2) \quad (\textit{\textbf{monotonicity of}}\ f\{\dots\}) \tag{5.14}$$

$$B_1 \subseteq B_2 \subseteq Y \Rightarrow f^{-1}(B_1) \subseteq f^{-1}(B_2) \quad (\textit{\textbf{monotonicity of}}\ f^{-1}\{\dots\}) \tag{5.15}$$

$$x \in X \Rightarrow f(\{x\}) = \{f(x)\} \tag{5.16}$$

$$f(X) = Y \quad \Leftrightarrow \quad f \text{ is “surjective” (see def.5.12 on p.42)} \tag{5.17}$$

$$f^{-1}(Y) = X \quad \textit{always!} \tag{5.18}$$

5.2.5 Injective, Surjective and Bijective functions

Definition 5.12 (Surjective, injective, bijective). Let $f : X \to Y$, with graph Γ_f.
a. Surjectivity: It need not be true that $f(X) = \{f(x) : x \in X\}$ equals the entire codomain Y, i.e., that

(5.19) for each $y \in Y$ there exists at least one $x \in X$ such that $(x, y) \in \Gamma_f$.

But if $f(X) = Y$, i.e., if (5.19) holds, we call f **surjective** aka **surjection**. aka **onto function**. We also say that f maps X **onto** Y.

b. Injectivity: It need not be true that if $y \in f(X)$, then $y = f(x)$ for a unique x, i.e., that if there is another $x_1 \in X$ such that also $y = f(x_1)$ then it follows that $x_1 = x$. But if this is the case, i.e., if

(5.20) for each $y \in Y$ there exists at most one $x \in X$such that $(x, y) \in \Gamma_f$.

then we call f **injective** aka **injection** aka **one to one** function.
We can express (5.20) also as follows: If $x, x_1 \in X$ and $y \in Y$ are such that $(x, y) \in \Gamma_f$ and $(x_1, y) \in \Gamma_f$ then it follows that $x_1 = x$.

c. Bijectivity: Let $f : X \to Y$ be both injective and surjective. Such a function is called **bijective**, aka **bijection**. We often write $f : X \xrightarrow{\sim} Y$ for a bijective function f.
It follows from (5.19) and (5.20) that f is bijective if and only if

(5.21) for each $y \in Y$ there exists exactly one $x \in X$such that $(x, y) \in \Gamma_f$.

We rewrite (5.21) by employing Γ_f's inverse relation $\Gamma_f^{-1} = \{(y, x) : (x.y) \in \Gamma\}$ (see def. 5.6 on p.38) and obtain

(5.22) for each $y \in Y$ there exists exactly one $x \in X$such that $(y, x) \in \Gamma_f^{-1}$.

But this implies, according to (5.5), that Γ_f^{-1} is the graph of a function $g := (Y, X, \Gamma_f^{-1})$ with domain Y and codomain X where, for a given $y \in Y$, $g(y)$ stands for the uniquely determined $x \in X$ such that $(y, x) \in \Gamma_f^{-1}$. Note that

(5.23) $$\Gamma_f^{-1} = \Gamma_g.$$

We call g the **inverse mapping** or **inverse function** of f and write f^{-1} instead of g. □

Notation 5.1. We will occasionally use special arrow symbols to give a visual clue about injectivity, surjectivity and bijectivity of a function.

a) $f : X \twoheadrightarrow Y$ and $X \overset{f}{\twoheadrightarrow} Y$ indicate that the function f is surjective,

b) $f : X \rightarrowtail Y$ and $X \overset{f}{\rightarrowtail} Y$ indicate that the function f is injective,

c) $f : X \xrightarrow{\sim} Y$ and $f : X \xrightarrow{\cong} Y$ indicate that the function f is bijective. $\square$

Moreover, $X \cong Y$ implies that there exists a bijection between the sets X and Y.

Remark 5.3.

(a) It follows from (5.23) that

$$\Gamma_f^{-1} = \Gamma_{f^{-1}}. \tag{5.24}$$

(b) Each $x \in X$ is mapped to $y = f(x)$ which is the only element of Y such that $f^{-1}(y) = x$,

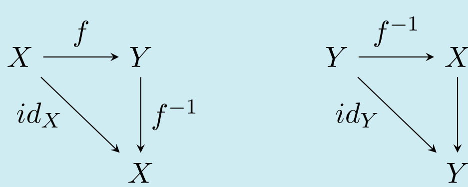
(c) Each $y \in Y$ is mapped to $x = f^{-1}(y)$ which is the only element of X such that $f(x) = y$.

(d) It follows from **(b)** and **(c)** that

$$\text{if } x \in X, y \in Y \text{ then } f(x) = y \ \Leftrightarrow \ x = f^{-1}(y). \tag{5.25}$$

(e) It also follows from **(b)** and **(c)** that $f^{-1}(f(x)) = x$ for all $x \in X$ and $f(f^{-1}(y)) = y$ for all $y \in Y$. In other words, $f^{-1} \circ f = id_X$ and $f \circ f^{-1} = id_Y$. Here is the picture:

(5.26) Inverse function:



$\square$

Theorem 5.1 (Characterization of inverse functions)**.** *Let X and Y be nonempty sets and $f : X \to Y$. The following are equivalent:*

(a) *f is bijective.*

(b) *There exists $g : Y \to X$ such that both $g \circ f = id_X$ and $f \circ g = id_Y$.*

Proposition 5.4. *Let $(R, \oplus, \odot, P)$ be an ordered integral domain*

(A) *Let $b \in R$. Then the function*

$$T : R \to R; \quad x \mapsto x \oplus b,$$

is a bijection.

(B) *Let $a \in R, a \neq 0$. Then the function*

$$D : R \to a \odot R; \quad x \mapsto a \odot x,$$

is a bijection. (As usual, $a \odot R = aR = \{a \odot r : r \in R\}$.)

Proposition 5.5. *Let $X, Y, Z \neq \emptyset$. Let $f : X \to Y$ and $g : Y \to Z$.*

- ***(a)*** *If both f, g are injective then $g \circ f$ is injective.*
- ***(b)*** *If both f, g are surjective then $g \circ f$ is surjective.*
- ***(c)*** *If both f, g are bijective then $g \circ f$ is bijective.*

Corollary 5.1. *Let $X, Y, Z \neq \emptyset$. Let $f : X \to Y$ and $g : Y \to Z$.*

- ***(a)*** *If f is bijective and g is injective then both $g \circ f$ and $f \circ g$ are injective.*
- ***(b)*** *If f is bijective and g is surjective then both $g \circ f$ and $f \circ g$ are surjective.*
- ***(c)*** *If f is bijective and g is bijective then both $g \circ f$ and $f \circ g$ are bijective.*

Proposition 5.6. ⋆ *Let X be an arbitrary set and let A be a nonempty proper subset of X. so that $X = A \uplus A^{\complement}$ is a partitioning of X into two nonempty subsets A and $A^{\complement}$. Let $a \in A$, $a_0 \in A^{\complement}$ and $A' := (A \setminus \{a\}) \uplus \{a_0\}$. Then the function*

$$\varphi : A' \xrightarrow{\sim} A; \qquad x \mapsto \begin{cases} x & \text{if } x \neq a_0, \\ a & \text{if } x = a_0 \end{cases}$$

is a bijection.

Proposition 5.7. *Let $X, Y \neq \emptyset$. Let $f : X \to Y$ and $g : Y \to X$ such that $g \circ f = id_X$. Then*

- ***(a)*** *f is injective,*
- ***(b)*** *g is surjective.*

Proposition 5.8. *Let $X, Y \neq \emptyset$.*

- ***(a)*** *Let $f : X \to Y$. If f is injective then there exists $g : Y \to X$ such that $g \circ f = id_X$ and any such function g is necessarily surjective.*
- ***(b)*** *Let $g : Y \to X$. If g is surjective then there exists $f : X \to Y$ such that $g \circ f = id_X$ and any such function f is necessarily injective.*

Definition 5.13 (Left inverses and right inverses)**.** Let $X, Y \neq \emptyset$. Let $f : X \to Y$ and $g : Y \to X$ such that $g \circ f = id_X$. We say that

(a) f possesses a **left inverse**,
(b) g is a **left inverse** of f,
(c) g possesses a **right inverse**,
(d) f is a **right inverse** of g. $\square$

Theorem 5.2. *Let $X, Y \neq \emptyset$.*

(a) *Let $f : X \to Y$. Then f is injective $\Leftrightarrow$ f has a left inverse (which is necessarily surjective).*
(b) *Let $g : Y \to X$. Then g is surjective $\Leftrightarrow$ g has a right inverse (which is necessarily injective).*
(c) *An injection $X \to Y$ exists $\Leftrightarrow$ a surjection $Y \to X$ exists.*

5.2.6 Binary Operations and Restrictions and Extensions of Functions

Definition 5.14 (Binary and unary operations)**.** ⋆ Let X be a nonempty set.

A **binary operation** on X is a function

(5.27) $$\diamond : X \times X \longrightarrow X; \quad (x, y) \mapsto x \diamond y := \diamond(x, y).$$

A **unary operation**, on X is a function

(5.28) $$\bullet : X \longrightarrow X; \quad x \mapsto \bullet(x). \ \square$$

One often writes $x^\bullet$ or $\bullet x$ instead of $\bullet(x)$. For example, $-x$ instead of $-(x)$ and x^{-1} rather than ${}^{-1}(x)$.

Definition 5.15 (Restriction/Extension of a function)**.** Given are three nonempty sets A, X and Y such that $A \subseteq X$ and a function $f : X \to Y$ with domain X.

(a) We define the **restriction of f to A** as the function

(5.29) $$f\big|_A : A \to Y \quad \text{defined as} \quad f\big|_A(x) := f(x) \text{ for all } x \in A.$$

(b) Conversely, let $f : A \to Y$ and $\varphi : X \to Y$ be functions such that $f = \varphi\big|_A$. We then call φ an **extension** of f to X. $\square$

Proposition 5.9. *Let X, Y be nonempty sets. Let $f : X \xrightarrow{\sim} Y$ be bijective*

(a) *Let $\emptyset \neq A \subseteq X$, $B := f\big|_A(A) = \{f(a) : a \in A\}$.* [4] *Let $f' : A \to B$; $x \mapsto f(x)$, i.e., $f' = f\big|_A$, except that we have shrunken the codomain Y to B. Then f' is bijective.*

(b) *Let $\emptyset \neq V \subseteq Y$. Let $U := \{x \in X : f(x) \in V\}$.* [5] *Let $f'' : U \to V$; $x \mapsto f(x)$, i.e., $f'' = f\big|_U$, except that we have shrunken the domain X to U. Then f'' is bijective.*

5.2.7 Real–Valued Functions and Polynomials

Definition 5.16 (Real–Valued Function)**.**
Let X be an arbitrary, nonempty set. If the codomain Y of a mapping

$$f : X \to Y; \qquad x \mapsto f(x)$$

is a subset of $\mathbb{R}$, then we call $f(\cdot)$ a **real function** or **real–valued function**. ☐

Definition 5.17 (Operations on real–valued functions)**.** ⋆

Let X be an arbitrary nonempty set.
Given are two real–valued functions $f(\cdot), g(\cdot) : X \to \mathbb{R}$ and a real number α. The **sum** $f + g$, **difference** $f - g$, **product** fg or $f \cdot g$, **quotient** f/g, and **scalar product** αf are defined by doing the operation in question with the numbers $f(x)$ and $g(x)$ for each $x \in X$. In other words these items are defined by the following equations:

$$(5.30)\qquad \begin{aligned} (f+g)(x) &:= f(x) + g(x), \\ (f-g)(x) &:= f(x) - g(x), \\ (fg)(x) &:= f(x)g(x), \\ (f/g)(x) &:= f(x)/g(x) \qquad \text{for all } x \in X \text{ where } g(x) \neq 0, \\ (\alpha f)(x) &:= \alpha \cdot g(x). \ \square \end{aligned}$$

Definition 5.18 (Negative function)**.** ⋆

Let X be an arbitrary, nonempty set and let $f : X \to \mathbb{R}$. The function

$$-f(\cdot) : X \to \mathbb{R}; \qquad x \mapsto -f(x).$$

is called **negative** f or **minus** f. We usually write $-f$ for $-f(\cdot)$. ☐

Definition 5.19 (Polynomials)**.** Let A be subset of the real numbers and let $p(\cdot) : A \to \mathbb{R}$ be a real–valued function on A. $p(\cdot)$ is called a **polynomial**. if there is an integer $n \geq 0$ and real numbers $a_1, a_2, \dots, a_n$ which are constant (they do not depend on x) so that $p(\cdot)$ can be written as a sum

$$p(x) \; = \; a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \; = \; \sum_{j=0}^{n} a_j x^j. \tag{5.31}$$

In other words, polynomials are linear combinations of the **monomials** $x \to x^k \quad (k \in (Z)_{\geq 0}$. If $a_n \neq 0$ then we call n the **degree** of p. The zero function $x \mapsto 0 = 0 \cdot x^0$ is a polynomial which we call the **zero polynomial**. Note that it has no degree because we cannot represent it in the form (5.31) with a non–zero coefficient a_n. We call $z \in A$ a **root** of the polynomial p if $p(z) = 0$.

If we talk about polynomials without explicitly specifying the domain then it is implied that the domain is $\mathbb{R}$. ☐

Proposition 5.10. *If p_1 and p_2 are polynomials and if $\lambda \in \mathbb{R}$ then*

- ***(a)*** *The sum $x \mapsto p_1(x) + p_2(x)$ is a polynomial.*
- ***(b)*** *The "scalar product" $x \mapsto \lambda p_1(x)$ is a polynomial.*

5.2.8 Families, Sequences, and Functions as Families

Definition 5.20 (Indexed families)**.** Let J and X be nonempty sets and assume that

for each $\imath \in J$ there exists exactly one indexed item $x_\imath \in X$.

- **(a)** We write $(x_\imath)_{\imath \in J}$ for this collection of indexed elements and call it an **indexed family** or simply a **family** in X.
- **(b)** J is called the **index set** of the family.
- **(c)** For each $\jmath \in J$, $x_\jmath$ is called a **member of the family** $(x_\imath)_{\imath \in J}$. ☐

Remark 5.4.

- Every family $(x_\imath)_{\imath \in J}$ in X can be interpreted as a function

$$x(\cdot) : J \longrightarrow X; \quad \imath \mapsto x_\imath \,. \; ☐$$

Definition 5.21 (Equality of families)**.** Two families $(x_i)_{i \in I}$ and $(y_j)_{j \in J}$ are equal if

- **(a)** $I = J$,
- **(b)** $x_i = y_i$ for all $i \in I$. ☐

Note 5.1 (Simplified notation for families)**.**
If there is no confusion about the index set then it can be dropped from the specification of a family and we simply write $(x_\imath)_\imath$ instead of $(x_\imath)_{\imath\in J}$. We even may shorten this to $(x_\imath)$ if doing so does not lead do confusion.

Definition 5.22 (Sequences and subsequences)**.** Let $n_\star \in \mathbb{Z}$, let

$$J := [n_\star, \infty[_\mathbb{Z} = \{k \in \mathbb{Z} : k \geq n_\star\}.$$

Let X be an arbitrary nonempty set. An indexed family $(x_n)_{n\in J}$ in X with index set J is called a **sequence** in X with **start index** $n_\star$. We will also write

$$(x_n)_{n\geq n_\star} \quad \text{or} \quad (x_n)_{n=n_\star}^\infty \quad \text{or} \quad x_{n_\star}, x_{n_\star+1}, x_{n_\star+2}, \dots$$

for this sequence. As for families, the name of the index variable of a sequence is unimportant as long as it is applied consistently. It does not matter whether one writes, e.g.,

$$(x_n)_{n\geq n_\star} \quad \text{or} \quad (x_j)_{j\geq n_\star} \quad \text{or} \quad (x_\beta)_{\beta\geq n_\star} \quad \text{or} \quad (x_A)_{A=n_\star}^\infty.$$ [6]

Let $(n_j)_{j=1}^\infty$ be a sequence of integers n_j such that

1) $n_j \in J$ (i.e., a sequence of indices for the above sequence $(x_j)_{j=n_\star}^\infty$)
2) $i < j \Rightarrow n_i < n_j$ for all $i, j \in \mathbb{N}$.

Note that $n_j \in J$ for all $j \in \mathbb{N}$ implies $n_\star \leq n_1 < n_2 < \dots$. If we write $I := \{n_j : j \in \mathbb{N}\}$ then we see that $(x_n)_{n\in I} = (x_{n_j})_{j\in\mathbb{N}}$, thus this object is an indexed family whose index set I is a subset of the original index set J. We call $(x_{n_j})_{j\in\mathbb{N}} = (x_{n_j})_{j=1}^\infty$ a **subsequence** of the sequence $(x_j)_{j=n_\star}^\infty$. This is an appropriate name since we obtain $(x_{n_j})_{j=1}^\infty$ from $(x_j)_{j\in J}$ by removing all members x_n such that none of the n_j equals n. Be sure to understand that, according to this definition, the sequence $(n_j)_{j\in\mathbb{N}}$ is a subsequence of the full sequence of indices $(n)_{n=n_\star}^\infty$. We will also write

$$(x_{n_j})_{j\in\mathbb{N}} \quad \text{or} \quad (x_{n_j})_{j\geq 1} \quad \text{or} \quad (x_{n_j})_{j=1}^\infty \quad \text{or} \quad x_{n_1}, x_{n_2}, x_{n_3}, \dots$$

for this subsequence. □

Note 5.2 (Simplified notation for sequences)**.**

(a) It is customary to choose either of i, j, k, l, m, n as the symbol of the index variable of a sequence and to stay away from other symbols whenever possible.
(b) By default the index set for a sequence is $\mathbb{N} = \{1, 2, 3, 4, \dots\}$.
(c) We are allowed to write $(x_n)_n$ or just (x_n) if there is no confusion about the value of $n_\star$ or if this value is irrelevant for the statement at hand.
(d) Customary simplified notation for subsequences is either of $(x_{n_j})_{j\in\mathbb{N}}$, $(x_{n_j})_{j\geq 1}$, $(x_{n_j})_j$ or simply (x_{n_j}).

Compare this to note 5.1 about simplified notation for families. □

Assumption 5.1 (indices of sequences)**.** Unless explicitly stated otherwise, sequences are always indexed $1, 2, 3, \dots$, i.e., the first index is 1, there is no largest index and, given any index, you obtain the next one by adding 1 to it. □

In contrast to sets, families and sequences allow us to incorporate duplicates.

Proposition 5.11 (Functions are families and families are functions)**.** *The following two ways of specifying a function $f : X \to Y$, $x \mapsto f(x)$ are equivalent:*

- ***(a)*** *f is defined by its graph $\{(x, f(x)) : x \in X\}$.*
- ***(b)*** *f is defined by the following family in Y: $(f(x))_{x \in X}$*

5.3 Right Inverses and the Axiom of Choice ⋆

Definition 5.23 (Choice function)**.** Let $\mathscr{A}$ be a collection of nonempty sets and let Ω be a set such that $\bigcup[A : A \in \mathscr{A}] \subseteq \Omega$. Let the function

$$c : \mathscr{A} \longrightarrow \Omega \qquad \text{satisfy} \qquad c(A) \in A \ \text{ for all } A \in \mathscr{A}$$

Then we call c a **choice function** [7] on $\mathscr{A}$. □

Proposition 5.12. *Let $X, Y \neq \emptyset$.*
Let $g : Y \to X$. If g is surjective then there exists $f : X \to Y$ such that $g \circ f = id_X$.

Proposition 5.13. *Assume that each surjective function possesses a right inverse. Assume further that $\mathscr{A}$ is a collection of nonempty sets. Then there exists a choice function on $\mathscr{A}$.*

Theorem 5.3. *The following are equivalent.*

- ***(a)*** *For any sets $X, Y \neq \emptyset$ and surjective $g : Y \to X$ there exists a right inverse for g, i.e., a function $f : X \to Y$ such that $g \circ f = id_X$.*
- ***(b)*** *The Axiom of Choice holds: For any collection $\mathscr{A}$ of nonempty sets there exists a choice function on $\mathscr{A}$, i.e., a function $c : \mathscr{A} \to \bigcup[A : A \in \mathscr{A}]$ such that $c(A) \in A$ for all $A \in \mathscr{A}$.*

6 The Integers

Note to Math 330 students: You should read this chapter in parallel with chapters 2, 4, 6 and 7 of [1] Beck/Geoghegan Art of Proof

6.1 The Integers, the Induction Axiom, and the Induction Principles

Since addition and multiplication are associative in integrals domains $(R, \oplus, \odot)$ we will heneforth write $a \oplus b \oplus c$ for either of $(a \oplus b) \oplus c$, $a \oplus (b \oplus c)$, and $a \odot b \odot c$ for either of $(a \odot b) \odot c$, $a \odot (b \odot c)$. Here we assumed that $a, b, c \in R$.

The case of more than three operands will be taken care of later by Theorem 6.7 (Generalized Law of Associativity) on p.55.

Axiom 6.1 (Integers and Natural Numbers)**.**
We postulate the existence of two sets, $\mathbb{Z}$ and $\mathbb{N}$, which satisfy the following:

(a) $\mathbb{Z}$ is endowed with two binary operations "$+$" (called addition) and "$\cdot$" (called multiplication) and with a positive cone $\mathbb{N}$ such that $(\mathbb{Z}, +, \cdot, \mathbb{N})$ is an ordered integral domain. We denote the additive unit of this integral domain by 0 and its multiplicative unit by 1.

(b) **Induction Axiom**: Let $A \subseteq \mathbb{Z}$ such that
  (1) $1 \in A$,
  (2) $k \in A$ implies $k + 1 \in A$.
  Then $A \supseteq \mathbb{N}$.

We call $\mathbb{Z}$ the set of **integers**, and we call $\mathbb{N}$ the set of **natural numbers**. $\square$

Definition 6.1 (Decimal Digits)**.** We use 1 (the neutral element for "$\cdot$") and addition $a + b$ to define the following integers:

$$2 := 1 + 1,\ 3 := 2 + 1,\ 4 := 3 + 1,\ 5 := 4 + 1,\ 6 := 5 + 1,\ 7 := 6 + 1,\ 8 := 7 + 1,\ 9 := 8 + 1\,.$$

We call the elements of the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ **digits** aka **decimal digits**. $\square$

Proposition 6.1. *Let $i, j, n \in \mathbb{Z}$. Then*

$$n + i \in [i, \infty[_{\mathbb{Z}} \Leftrightarrow\ n + j \in [j, \infty[_{\mathbb{Z}}\,.$$

Corollary 6.1. *Let $k_0, n \in \mathbb{Z}$. Then*

$$n \in [k_0, \infty[_{\mathbb{Z}} \Leftrightarrow \ n - k_0 + 1 \in \mathbb{N}\,.$$

Theorem 6.1 (Generalization of the Induction Axiom). *Let $k_0 \in \mathbb{Z}$ and let*

$$A_{k_0} \ := \ \{k \in \mathbb{Z} : j \geq k_0\} \ = \ [k_0, \infty[_{\mathbb{Z}}$$

be the set of all integers at least as big as k_0. Let $A \subseteq \mathbb{Z}$ such that

(a) $k_0 \in A$,

(b) $k \in A$ *implies* $k+1 \in A$.

Then $A \supseteq A_{k_0}$.

Theorem 6.2 (Principle of Mathematical Induction). *Assume that for each integer $k \geq k_0$ there is an associated statement $P(k)$ such that*

A. Base case. *The statement $P(k_0)$ is true.*

B. Induction Step. *For each $k \geq k_0$ we have the following: Assuming that $P(k)$ is true ("**Induction Assumption**"), it can be shown that $P(k+1)$ also is true.*

*It then follows that $P(k)$ is true for **each** $k \geq k_0$.*

Theorem 6.3 (Principle of Strong Mathematical Induction). *Let $k_0 \in \mathbb{Z}$ and assume that for each integer $k \geq k_0$ there is an associated statement $P(k)$ such that the following is valid:*

A. Base case. *The statement $P(k_0)$ is true.*

B. Induction Step. *For each $k \geq k_0$ we have the following: Assuming that $P(j)$ is true for all $j \in \mathbb{Z}$ such that $k_0 \leq j \leq k$ ("**Induction Assumption**"), it can be shown that $P(k+1)$ also is true.*

*It then follows that $P(k)$ is true for **each** $k \geq k_0$.*

6.2 The Discrete Structure of the Integers

Theorem 6.4 (B/G Prop.2.20). *If $k \in \mathbb{N}$, then*

$$k \ \geq \ 1. \tag{6.1}$$

Proposition 6.2 (B/G Prop.2.21)**.** *There exists no $x \in \mathbb{Z}$ such that $0 < x < 1$.*

Corollary 6.2 (B/G Cor.2.22)**.** *Let $n \in \mathbb{Z}$. There exists no $x \in \mathbb{Z}$ such that $n < x < n + 1$.*

Proposition 6.3 (sharpening of B/G Prop.2.13)**.** $\mathbb{N} = \{k \in \mathbb{Z} : k \geq 1\}$.

6.3 Divisibility

Definition 6.2 (Divisibility)**.**

(a) Let $m, n \in \mathbb{Z}$. We say that n **divides** m or, equivalently, that m **is divisible by** n if there exists $j \in \mathbb{Z}$ such that $m = jn$. We then write $n \mid m$, and we write $n \nmid m$ if n does not divide m, i.e., there is no $k \in \mathbb{Z}$ that satisfies $m = kn$.

(b) Let $m \in \mathbb{Z}$. We say that m is an **even** integer if $2 \mid m$. We say that m is an **odd** integer if m is not even, i.e., $2 \nmid m$. □

Proposition 6.4. *Let $m, n \in \mathbb{Z}$ such that $m \neq 0$ and $m \mid n$, i.e., there exists $j \in \mathbb{Z}$ be such that $n = j \cdot m$. Then j is uniquely determined.*

Definition 6.3 (Quotients)**.** Let $d, n \in \mathbb{Z}$ such that $d \mid n$ and $d, n \neq 0$.
Let $q \in \mathbb{Z}$ be the unique integer for which $n = q \cdot d$. We write either of

$$\frac{n}{d}, \quad n/d, \quad n \div d \quad \text{instead of } q\,,$$

and we call n the **dividend** or **numerator**, d the **divisor** or **denominator**, and q the **quotient** of the expression n/d. We also define $\frac{0}{d} := 0$ if $d \neq 0$, but we leave $\frac{n}{0}$ undefined for all $n \in \mathbb{Z}$. □

Proposition 6.5 (B/G prop.1.16)**.** *If m and n are even integers, then so are $m + n$ and mn.*

Proposition 6.6 (B/G prop.1.17)**.**

(a) *If m is an integer then $m \mid 0$. In particular, $0 \mid 0$.*

(b) *If m is a nonzero integer then $0 \nmid m$.*

Proposition 6.7 (B/G prop.2.18). *Let $n \in \mathbb{N}$. Then*

- ***(a)*** $n^3 + 2n$ *is divisible by* 3,
- ***(b)*** $n^4 - 6n^3 + 11n^2 - 6n$ *is divisible by* 4,
- ***(c)*** $n^2 + n$ *is even, i.e., divisible by* 2,
- ***(d)*** $n^3 + 5n$ *is divisible by* 6.

Proposition 6.8 (B/G Prop.2.24). *Let $n \in \mathbb{N}$. Then $n^2 + 1 > n$.*

Proposition 6.9 (B/G prop.2.23). *Let $m, n \in \mathbb{N}$. If $m \mid n$ then $m \leq n$*

6.4 Embedding the Integers Into an Ordered Integral Domain

Definition 6.4 (Natural Embedding of the Integers Into $(R, \oplus, \odot, P)$). ⋆
We define a function $e : \mathbb{Z} \to R$, partially by recursion, as follows.

$$e(0_{\mathbb{Z}}) := 0_R\,, \tag{6.2}$$

$$e(n + 1_{\mathbb{Z}}) := e(n) \oplus 1_R \quad \text{for } n \in [0, \infty[_{\mathbb{Z}}\,, \tag{6.3}$$

$$e(n) := \ominus\, e(-n) \quad \text{for } n \in\,]-\infty, -1\,]_{\mathbb{Z}}\,. \tag{6.4}$$

We call e the **natural embedding of $\mathbb{Z}$ into** $(R, \oplus, \odot, P)$. □

Theorem 6.5. *Let $R = (R, \oplus, \odot, P)$ be an ordered integral domain.*
The natural embedding $e : (\mathbb{Z}, +, \cdot, \mathbb{N}) \longrightarrow (R, \oplus, \odot, P)$ which is defined as follows:

$$e(0_{\mathbb{Z}}) = 0_R, \quad e(n + 1_{\mathbb{Z}}) = e(n) \oplus 1_R \ \text{ if } n \in \mathbb{N}, \quad e(n) = \ominus e(-n) \ \text{ if } n < 0$$

is an injective function with the following structure preserving properties ($m, n \in \mathbb{Z}$):

- ***(a)*** *e maps neutral element to neutral element: $e(0_{\mathbb{Z}}) = 0_R$ and $e(1_{\mathbb{Z}}) = 1_R$.*
- ***(b)*** *Image of the sum = sum of the images: $e(m + n) = e(m) \oplus e(n)$.*
- ***(c)*** *Image of the product = product of the images: $\Rightarrow$ $e(m \cdot n) = e(m) \odot e(n)$.*
- ***(d)*** *Image of the additive inverse = additive inverse of the image: $e(-m) = \ominus e(m)$.*
- ***(e)*** *e preserves the order: $m < n \Leftrightarrow e(m) \prec e(n)$ and $m \leq n \Leftrightarrow e(m) \preceq e(n)$.*

Definition 6.5 (Ring homomorphism). ⋆

A function $h : (R, \oplus, \odot) \longrightarrow (R', \oplus', \odot')$ between two commutative rings with unit and in particular between two ordered integral domains [8] which satisfies Theorem6.5.**a–d** is called a **ring homomorphism**.

Note that ring homomorphisms play for commutative rings with unit the role which group homomorphisms play for groups. □

Theorem 6.6. *Let $R = (R, \oplus, \odot, P)$ be an ordered integral domain which satisfies the induction axiom. See Axiom 6.1 (Integers and Natural Numbers) on p.50.*

Then the natural embedding $e : (\mathbb{Z}, +, \cdot, \mathbb{N}) \longrightarrow (R, \oplus, \odot, P)$ is an isomorphism of ordered integral domains, i.e., e is bijective and it's inverse, e^{-1}, also satisfies ***(a)–(e)*** *of Theorem 6.5.*

6.5 Recursive Definitions of Sums, Products and Powers in Integral Domains

Assume in this entire chapter that $R = (R, \oplus, \odot, P)$ is an ordered integral domain

Definition 6.6. Let $k \in \mathbb{Z}$ and let $(x_j)_{j=k}^{\infty} \in R$. For each $n \in \mathbb{Z}$ such that $k \leq n$, we define an element of R, denoted $\displaystyle\sum_{j=k}^{n} x_j$ or $x_k \oplus x_{k+1} \oplus \cdots \oplus x_n$, as follows.

$$\text{(6.5)} \qquad \text{(i)} \quad \sum_{j=k}^{k} x_j = x_k, \qquad \text{(ii)} \quad \sum_{j=k}^{n+1} x_j = \sum_{j=k}^{n} x_j \oplus x_{n+1}.$$

We call $\displaystyle\sum_{j=k}^{n} x_j$ the **sum** of $x_k, x_{k+1}, \dots, x_{n-1}, x_n$. □

Definition 6.7 (Definition of $\prod_{j=k}^{n} x_j$). Let $k \in \mathbb{Z}$ and let $(x_j)_{j=k}^{\infty} \in R$. For each $n \in \mathbb{Z}$ such that $k \leq n$, we define an element of R, denoted $\displaystyle\prod_{j=k}^{n} x_j$ or $x_k \odot x_{k+1} \cdots \odot x_n$, as follows.

$$\text{(6.6)} \qquad \text{(i)} \quad \prod_{j=k}^{k} x_j = x_k, \qquad \text{(ii)} \quad \prod_{j=k}^{n+1} x_j = \prod_{j=k}^{n} x_j \odot x_{n+1}.$$

We call $\displaystyle\prod_{j=k}^{n} x_j$ the **product** of $x_k, x_{k+1}, \dots, x_{n-1}, x_n$. □

Proposition 6.10 (B/G prop.4.15). *Let $m, n, k \in \mathbb{Z}$, $c \in R$, and let $(x_j)_{j=k}^{\infty}$ be a sequence in R. Then*

(a) $c \odot \left(\sum_{j=k}^{n} x_j \right) = \sum_{j=k}^{n} (c \odot x_j)$.

(b) *If $x_j = 1$ for all $j \in [k,n]_{\mathbb{Z}}$ then* $\sum_{j=k}^{n} x_j = n \ominus k \oplus 1$.

(c) *If $x_j = c$ for all $j \in [k,n]_{\mathbb{Z}}$ then* $\sum_{j=k}^{n} x_j = (n \ominus k \oplus 1) \odot c$.

Proposition 6.11 (B/G prop.4.16).
Let $m, n, p \in \mathbb{Z}$ such that $m \leq n < p$, and let $(x_j)_{j=m}^{p}$ and $(y_j)_{j=m}^{p}$ be sequences in R. Then

(a) $\sum_{j=m}^{p} x_j = \sum_{j=m}^{n} x_j \oplus \sum_{j=n+1}^{p} x_j$,

(b) $\sum_{j=m}^{p} (x_j \oplus y_j) = \sum_{j=m}^{p} x_j \oplus \sum_{j=m}^{p} y_j$.

Proposition 6.12 (B/G prop.4.17). *Let $m, n, p \in \mathbb{Z}$ such that $m \leq n$, and let $(x_j)_{j=m}^{n}$ be a sequence in R. Then* $\sum_{j=m}^{n} x_j = \sum_{j=m+p}^{n+p} x_{j-p}$.

Proposition 6.13 (B/G prop.4.18). *Let $m, n \in \mathbb{Z}$ such that $m \leq n$, and let $(x_j)_{j=m}^{n}$ and $(y_j)_{j=m}^{n}$ be sequences in R such that $x_j \leq y_j$ for all $m \leq j \leq n$. Then* $\sum_{j=m}^{n} x_j \leq \sum_{j=m}^{n} y_j$.

Theorem 6.7 (Generalized Law of Associativity).
Let $x_1, x_2, \dots, x_n \in R$. Then the formulas for associativity stated for $n = 3$,

- $x_1 \oplus (x_2 \oplus x_3) = (x_1 \oplus x_2) \oplus x_3$ *for sums*
- $x_1 \odot (x_2 \odot x_3) = (x_1 \odot x_2) \odot x_3$ *for products*

extend to $x_1, x_2, \dots, x_n$ in the following sense: It does not matter how parentheses are used to control the order how the sum and the product of those n items is evaluated.

- *Moreover, the value of any such grouping is* $\sum_{j=1}^{n} x_n$ *for summation and* $\prod_{j=1}^{n} x_n$ *for products.*

Definition 6.8. Let $\beta \in R$. For any $n \in \mathbb{Z}_{\geq 0}$ we define $\beta^n \in R$ recursively as follows:

(6.7) (i) $\beta^0 := 1,$ (ii) $\beta^{n+1} = \beta^n \odot \beta.$

In an expression of the form β^n we call β the **basis**, we call n the **exponent**, and we call β^n the n–th **power** of β. □

Proposition 6.14 (B/G prop.4.6: Arithmetic Rules for Exponentiation)**.** *Let $\beta \in R$ and $k, m \in \mathbb{Z}_{\geq 0}$. We have the following:*

- *(a)* *If $\beta > 0$ then $\beta^k > 0$,*
- *(b)* $\beta^m \odot \beta^k = \beta^{m+k}$,
- *(c)* $(\beta^m)^k = \beta^{mk}$.

Proposition 6.15 (B/G prop.10.9)**.** *Let $a \in R$ such that $0 \leq a \leq 1$, and let $m, n \in \mathbb{N}$ such that $m \geq n$. Then $a^m \leq a^n$.*

Proposition 6.16 (B/G prop.8.41)**.** *Let $a \in R$. Then $a^2 < a^3$ if and only if $a > 1$.*

Definition 6.9 (Finite Geometric Series)**.** Let $q \in R$ and $n \in \mathbb{Z}_{\geq 0}$.

We call a sum of the form $\sum_{j=0}^{n} q^j$ a **finite geometric series**. □

Proposition 6.17 (Finite Geometric Series Formula (B/G prop.4.13))**.** *Let $q \in R$. If $n \in \mathbb{Z}_{\geq 0}$ then*

$$(1 \ominus q) \odot \sum_{j=0}^{n} q^j = 1 \ominus q^{n+1}.$$

6.6 Binomial Coefficients

Definition 6.10 (Definition of Factorials)**.**
For any $n \in \mathbb{Z}_{\geq 0}$ we define a natural number $n!$ recursively as follows:

(6.8) (i) $0! := 1,$ (ii) $(n+1)! = n! \cdot (n+1).$

We pronounce $n!$ as n **factorial**. □

Definition 6.11 (Binomial coefficients)**.** Let $n, k \in \mathbb{Z}$ such that $0 \le k \le n$. We define the **binomial coefficient** $\binom{n}{k}$ (pronounced "n choose k") as follows:

$$\binom{n}{k} := \begin{cases} 1 & \text{if k = 0 or k = n,} \\ \binom{n-1}{k-1} + \binom{n-1}{k} & \text{otherwise, i.e., } n \ge 2 \text{ and } 0 < k < n. \end{cases} \tag{6.9}$$
□

Proposition 6.18. *Let $n, k \in \mathbb{Z}$ such that $0 \le k \le n$. Then*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{6.10}$$

Lemma 6.1 (Symmetry and reduction lemma)**.**

$$\binom{n}{k} = \binom{n}{n-k} \quad (0 \le k \le n) \qquad \textit{symmetry} \tag{6.11a}$$

$$\binom{n}{k} = \frac{n}{k} \cdot \binom{n-1}{k-1} \quad (1 \le k \le n) \qquad \textit{reduction} \tag{6.11b}$$

Theorem 6.8 (Binomial theorem)**.** *Let $R = (R, \oplus, \odot)$ be an integral domain. If $n \in \mathbb{Z}_{\ge 0}$ and $a, b \in R$, then*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \tag{6.12}$$

Corollary 6.3. *Let $n \in \mathbb{Z}_{\ge 0}$. Then* $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

6.7 Bernstein Polynomials ⋆

Definition 6.12 (Bernstein Polynomials). ⋆

Let $f : [0,1] \to \mathbb{R}$ be a real–valued function on the unit interval which need not necessarily be continuous. If $n \in \mathbb{N}$ then

$$B_n^f : \mathbb{R} \to \mathbb{R}; \qquad x \mapsto B_n^f(x) := \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}. \tag{6.13}$$

defines a function of the form (**??**) (see p.**??**), thus B_n^f is a polynomial which we call the n–th **Bernstein polynomial** associated with $f(\cdot)$. □

Proposition 6.19 (The Bernstein polynomials for $1, id(\cdot), id^2(\cdot)$). *Let*

$$1 : x \mapsto 1; \qquad id : x \mapsto x; \qquad id^2 : x \mapsto x^2; \qquad (0 \le x \le 1) \tag{6.14}$$

be the constant function 1*, the identity function, and the square function on the unit interval* $[0,1]$*. Then*

$$B_n^1 = 1, \tag{6.15a}$$

$$B_n^{id} = id, \tag{6.15b}$$

$$B_n^{id^2} = \frac{1}{n} id + \frac{n-1}{n} id^2. \tag{6.15c}$$

In other words, for any real number x we have

$$\begin{aligned} B_n^1(x) &= 1 \\ B_n^{id}(x) &= id(x) = x \\ B_n^{id^2}(x) &= \frac{1}{n} id(x) + \frac{n-1}{n} id^2(x) = \frac{1}{n} x + \frac{n-1}{n} x^2. \end{aligned}$$

6.8 The Well–Ordering Principle

Theorem 6.9 (Well–Ordering Principle).
Every nonempty subset of $\mathbb{N}$ possesses a minimum, i.e., a smallest element.

Theorem 6.10 (Extended Well–Ordering Principle).

(a) *Let A be a nonempty subset of $\mathbb{Z}$ which is bounded below. Then A possesses a minimum in $\mathbb{Z}$.*
(b) *Let B be a nonempty subset of $\mathbb{Z}$ which is bounded above. Then B possesses a maximum in $\mathbb{Z}$.*
(c) *Let C be a nonempty bounded subset of $\mathbb{Z}$. Then C possesses both minimum and maximum in $\mathbb{Z}$.*

Proposition 6.20. *Let $\emptyset \neq A \subseteq B \subseteq \mathbb{Z}$.*

- **(a)** *If B is bounded below (resp., above), then $\min(A) \geq \min(B)$ (resp., $\max(A) \leq \max(B)$).*
- **(b)** *If also $\min(B) \notin A$ (resp., $\max(B) \notin A$), then $\min(A) > \min(B)$ (resp., $\max(A) < \max(B)$).*

Proposition 6.21 ($\mathbb{N}$ is unbounded in $\mathbb{Z}$).
For any $k \in \mathbb{Z}$ there exists $n \in \mathbb{N}$ such that $n > k$, i.e., there are no upper bounds for $\mathbb{N}$ in $\mathbb{Z}$.

6.9 The Division Algorithm

Theorem 6.11 (Division Algorithm for Integers (B/G thm.6.13)).
Let $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ Then there exists a unique pair of integers q and r such that

$$m = qn + r \quad \text{and} \quad 0 \leq r < n. \tag{6.16}$$

*We call q the **quotient** and r the **remainder** when dividing n into m.*

Proposition 6.22 (B/G prop.6.15). *Let $m \in \mathbb{Z}$.*
Then m is odd if and only if there exists $q \in \mathbb{Z}$ such that $m = 2q + 1$.

Proposition 6.23. *Any product of odd numbers is odd.*

Proposition 6.24 (B/G prop.6.16). *Let $n \in \mathbb{Z}$. Then n is even or $n + 1$ is even.*

Proposition 6.25 (B/G prop.6.17). *Let $n \in \mathbb{Z}$. Then n is even if and only if n^2 is even.*

Proposition 6.26 (Division Algorithm for Polynomials (B/G prop. 6.18)). *Let $\alpha, \beta \in \mathbb{Z}_{\geq 0}$ and let*

$$n(x) := \sum_{j=0}^{\alpha} a_j x^j, \qquad m(x) := \sum_{j=0}^{\beta} b_j x^j, \tag{6.17}$$

be two polynomials with real coefficients a_j, b_j such that $n(x)$ is not the null polynomial $p(x) = 0$. Then there exist polynomials $q(x)$ and $r(x)$ such that $r(x)$ has degree less than α or $r(x) = 0$ (and hence has no degree), such that

$$m(x) := q(x)n(x) + r(x). \tag{6.18}$$

Proposition 6.27 (B/G prop.6.19)**.** *Let $p(x)$ be a polynomial and $z \in \mathbb{R}$. Then z is a root of p if and only if there exists a polynomial $q(x)$ such that*

$$p(x) = (x - z)q(x) \text{ for all } x \in \mathbb{R}. \tag{6.19}$$

6.10 The Integers Modulo n

Proposition 6.28 (B/G prop.6.24)**.** *For two integers a and b we define*

$$a \sim b \quad \text{if and only if} \quad n \mid (a - b). \tag{6.20}$$

Then

- ***(a)*** (6.20) *defines an equivalence relation on $\mathbb{Z}$,*
- ***(b)*** *The equivalence class for $m \in \mathbb{Z}$ is $[m] = [r]$, where r is the remainder of m modulo n. See thm.6.11 (division algorithm for integers) on p.59.*
- ***(c)*** *If $r \in [0, n-1]_{\mathbb{Z}}$ then $[r] = \{qn + r : q \in \mathbb{Z}\}$.*
- ***(d)*** *This equivalence relation has exactly n distinct equivalence classes $[0], [1], \dots, [n-1]$.*

Definition 6.13 (Equivalence Modulo n)**.**

- **(a)** We write $a \equiv b \mod n$ for $a \sim b$. We call n the **modulus**, and we say that a **equals** b **modulo** n.
- **(b)** We write

$$\mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z} := \{\, [0], [1], \dots, [n-1] \,\} \tag{6.21}$$

for the set of equivalence classes resulting from the equivalence relation $a \sim b$. (See prop.6.28**(b)** above.) We call $\mathbb{Z}_n$ the set of **integers modulo** n. $\square$

Proposition 6.29 (B/G prop.6.25)**.**
Let $a, a', b, b' \in \mathbb{Z}$ such that $a \sim a'$ and $a \sim a'$, i.e., $n \mid (a - a')$ and $n \mid (b - b')$.
Then $a + b \sim a' + b'$ and $ab \sim a'b'$.

Definition 6.14. Let $a, b \in \mathbb{Z}$.
We define addition $[a] \oplus [b]$ and multiplication $[a] \odot [b]$ for the corresponding equivalence classes $[a], [b] \in \mathbb{Z}_n$ in terms of ordinary addition and multiplication in $\mathbb{Z}$ as follows.

$$[a] \oplus [b] := [a + b]; \qquad [a] \odot [b] := [ab]. \tag{6.22}$$

We further define $[a]^0 := [1]$. $\square$

Theorem 6.12 (B/G prop.6.26 and B/G project 6.27).
- ***(a)*** *The operations* $\oplus$ *and* $\odot$ *on* $\mathbb{Z}_n$ *of Definition 6.14 above turn* $(\mathbb{Z}_n, \oplus, \odot)$ *into a commutative ring with unit.*
- ***(b)*** $(\mathbb{Z}_n, \oplus, \odot)$ *is an integral domain, i.e., there are no zero divisors, if and only if* n *is prime.*

Proposition 6.30 (Arithmetic mod n). *Let* $m_1, m_2, \dots m_k,\ a_1, a_2, \dots a_k \in \mathbb{Z}$. *Then*

$$[m_1 + m_2 + \dots + m_k] = [m_1] \oplus [m_2] \oplus \dots \oplus [m_k], \tag{6.23}$$

$$[m_1 \cdot m_2 \cdots m_k] = [m_1] \odot [m_2] \odot \dots \odot [m_k], \tag{6.24}$$

$$[\sum_{j=1}^{k} a_j x^j\ ] = \sum_{j=1}^{k} [a_j] \odot [x]^j. \tag{6.25}$$

6.11 The Greatest Common Divisor

Lemma 6.2 (B/G prop.2.34). *For* $m, n \in \mathbb{Z}$ *let*

$$S := S(m,n) := \{k \in \mathbb{N} : k = mx + ny \text{ for some } x, y, \in \mathbb{Z}\}. \tag{6.26}$$

Then S *is empty if and only if* $m = n = 0$.

Lemma 6.3. *For* $m, n \in \mathbb{Z}$ *let* $S(m,n)$ *be defined as in (6.26). Then*
- ***(a)*** $S(m,n) = S(n,m)$,
- ***(b)*** $S(m,n) = S(-m,n) = S(m,-n) = S(-m,-n)$,
- ***(c)*** $S(m,n) = S(|m|,|n|)$.

Definition 6.15 (Greatest Common Divisor).
For $m, n \in \mathbb{Z}$ let $S = S(m,n)$ be the set defined in (6.26). Let

$$\gcd(m,n) := \begin{cases} 0 & \text{if m = n = 0}, \\ \min(S) & \text{otherwise}. \end{cases} \tag{6.27}$$

We call $\gcd(m,n)$ the **greatest common divisor** of m and n. □

Proposition 6.31 (B/G prop.6.29). *Let $m, n \in \mathbb{Z}$. Then*

- ***(a)*** $gcd(m, n) \mid m$ *and* $gcd(m, n) \mid n$,
- ***(b)*** *If* $m \neq 0$ *or* $n \neq 0$ *then* $gcd(m, n) > 0$,
- ***(c)*** *Let* $k \in \mathbb{Z}$ *such that* $k \mid m$ *and* $k \mid n$. *Then* $k \mid gcd(m, n)$.

Proposition 6.32 (B/G prop.6.30). *Let $k, m, n \in \mathbb{Z}$. Then* $gcd(km, kn) = |k| \cdot gcd(m, n)$.

6.12 Prime Numbers

Definition 6.16 (Prime numbers and prime factorizations).

- **(a)** Let $p \in \mathbb{N}, p \geq 2$. p is a **prime number** or p is **prime** if $q \in \mathbb{Z}$ and $q \mid p$ implies that $q = \pm 1$ or $q = \pm p$. We note that 1 is **not** prime.
- **(b)** Let $p \in \mathbb{N}, p \geq 2$. p is called a **composite number** or just a **composite** if p is not prime.
- **(c)** Let $m \in \mathbb{N}, m \geq 2$. If there are primes $p_1, \dots, p_k$ such that $m = p_1 \cdot p_2 \cdots p_k$ then $p_1, \dots, p_k$ are called **factors** or **prime factors** of m and $p_1 \cdot p_2 \cdots p_k$ is called a **prime factorization** or just a **factorization** of m.
- **(d)** If the prime factorizations of $m, n \in \mathbb{N}$ both contain the prime number p then we call p a **common factor** of m and n.
- **(e)** If $m \in \mathbb{Z}$ satisfies $m \leq -2$ and if $p_1 \cdot p_2 \cdots p_k$ is a prime factorization of the positive(!) integer $-m$ then we call $-(p_1 \cdot p_2 \cdots p_k)$ a prime factorization of m. $\square$

Proposition 6.33 (B/G prop.6.28). *Let $n \in \mathbb{N}$ such that $n > 1$. Then n has a prime factorization.*

Lemma 6.4. *Let p be prime and let $n \in \mathbb{N}$. We have the following:*

- ***(a)*** *Either* $gcd(p, n) = 1$ *or* $gcd(p, n) = p$.
- ***(b)*** *If* $p \nmid n$ *(p does not divide n) then* $gcd(p, n) = 1$.

Definition 6.17 (relatively prime). Let $m, n \in \mathbb{Z}$. We say that m and n are **relatively prime** if their greatest common divisor satisfies

$$\text{gcd}(m, n) = 1. \quad \square$$

Proposition 6.34.

Two natural numbers are relatively prime $\Leftrightarrow$ *they possess no common factors.*

We next look at Euclids Lemma and the unqueness of prime number factorizations. Thm.6.13 below states the following: Every integer ≥ 2 can be factored uniquely (i.e. up to permutation) into primes. The proof of that theorem requires Euclid's lemma which in turn uses lemma 6.4 above.

Proposition 6.35 (B/G prop.6.31: Euclid's Lemma for Two Factors)**.**
Let p be prime and $m, n \in \mathbb{N}$. If $p \mid (mn)$ then $p \mid m$ or $p \mid n$.

The generalization of Euclid's lemma to more than two factors is a straightforward proof by induction.

Proposition 6.36 (Euclid's Lemma for more than two factors)**.**
Let p be prime and $m_1, m_2, \dots, m_k \in \mathbb{N}$. If $p \mid (m_1 m_2 \cdots m_k)$ then $p \mid m_j$ for some $1 \leq j \leq k$.

Theorem 6.13 (B/G thm 6.32: Uniqueness of prime factorizations)**.**
Every integer ≥ 2 can be factored uniquely (i.e., up to permutation) into primes.

Notation 6.1 ("The" prime factorization of an integer greater than 1)**.**
When we talk about prime factorizations of some $n \in [2, \infty[$ it usually does not matter in which order the prime factors of n occur. We will in such instances talk about **the** prime factorization of n. For example, We might say, "The prime factorization of n does not contain the number 2." $\square$

(a) $p_1 \cdots p_i \cdot q_1 \cdots q_j$ is the prime factorization of $m \cdot n$.
(b) If p is a prime factor of m then $p = p_k$ for some suitable $1 \leq k \leq i$.
(c) It follows from **(b)** that if $p > p_k$ for each $1 \leq k \leq i$ then p is not a prime factor of m.
(d) If p is prime and $p \mid mn$ then p is a prime factor of mn. If p is not a prime factor of m then it follows from **(a)** that p is a prime factor of n. That is of course just a reformulation of Euclid's lemma, but note that we used the uniqueness of prime factorizations to deduce this. $\square$

Proposition 6.37 (B/G Prop.6.33). *Let $a, b \in \mathbb{N}$, and assume that $a \mid b$. Further, assume that p is a prime factor of b that is not a prime factor of a. Then $a \mid \dfrac{b}{p}$.*

Proposition 6.38 (B/G Prop.6.34). *Let p be a prime and $k \in \mathbb{N}$ such that $0 < k < p$. Then $p \mid \binom{p}{k}$.*

Theorem 6.14 (Fermat's Little Theorem (B/G thm 6.35)).
If $m \in \mathbb{Z}$ and p is prime, then $m^p \equiv m \mod p$.

Proposition 6.39 (Corollary to Fermat's Little Theorem (B/G cor.6.36)).
Let p be prime and let $m \in \mathbb{N}$ such that $p \nmid m$. Then

$$m^{p-1} \equiv 1 \mod p.$$

6.13 The Base–β Representation of the Integers

Definition 6.18. ⋆ If $\beta \in \mathbb{Z}_{\geq 2}$ then we mean by a set of **base β digits** a set of $\beta - 1$ distinct symbols $\{d_i : i \in \mathbb{Z}, 0 \leq i < \beta$ such that each d_i represents the integer i. □

Proposition 6.40 (B/G thm.7.7: Existence of base–β representations).
Let $n \in \mathbb{N}$ and $\beta \in \mathbb{N}$ such that $\beta \geq 2$. Then there exists a nonnegative integer $\mu = \mu(n)$, and there exist integers d_j $(0 \leq j \leq \mu)$ such that $0 \leq d_j < \beta$ for each j and $d_\mu > 0$, and also

$$n = \sum_{j=0}^{\mu} d_j \beta^j. \tag{6.28}$$

Proposition 6.41 (B/G prop.7.9: Uniqueness of base–β representations).
Let $n \in \mathbb{N}$ and $\beta \in \mathbb{N}$ such that $\beta \geq 2$. Assume that

$$n = \sum_{j=0}^{\mu} d_j \beta^j = \sum_{j=0}^{\mu'} d'_j \beta^j \tag{6.29}$$

*where $\mu, \mu' \in \mathbb{Z}_{\geq 0}$, each d_i and each d'_i is a base β digit, $d_\mu \neq 0$ and $d'_{\mu'} \neq 0$.
Then $\mu = \mu'$ and $d_i = d'_i$ for all i.*

Proposition 6.42 (B/G Prop.7.11). *Let* $n := \sum_{j=0}^{\mu} x_j 10^j$*, where each* x_j *is a digit and* $x_\mu \neq 0$*. Then*

$$n = x_0 + x_1 + \cdots + x_\mu \mod 3. \tag{6.30}$$

6.14 The Addition Algorithm for Two Nonnegative Numbers (Base 10)

7 Cardinality I: Finite and Countable Sets

Notation: In this entire chapter, if $n \in \mathbb{N}$, the symbol $[n]$ does not denote an equivalence class of any kind but the set $[1,n]_{\mathbb{Z}} = \{1, 2, \dots, n\}$ of the first n natural numbers. we further define $[0] := \emptyset$.

7.1 The Size of a Set

Proposition 7.1. *Let $n \in \mathbb{N}$. Let $\emptyset \neq A \subsetneq [n]$ be a proper, nonempty subset of $[n]$. Then there is no surjection from A onto $[n]$.*

Corollary 7.1. *The following contains B/G thm.13.4 and B/G cor.13.5. Let $m, n \in \mathbb{N}$.*

- ***(a)*** *If $m < n$ then there exists no surjective function $f : [m] \to [n]$.*
- ***(b)*** *If $m > n$ then there exists no injective function $g : [m] \to [n]$. This is commonly referred to as the **pigeonhole principle**.*
- ***(c)*** *If $m \neq n$ then there exists no bijective function $f : [m] \to [n]$.*
- ***(d)*** *There exists no surjective function $h : [m] \to \mathbb{N}$.*

Definition 7.1 (Finite and infinite sets)**.**

- **(a)** Let $X \neq \emptyset$ and $n \in \mathbb{N}$ such that there is a bijective mapping $F : [n] \to X$. By Corollary 7.1**(c)**, n is uniquely defined by the property that $[n]$ can be bijected to X. This allows us to define n as the **size** of the set X. We write $|X| = n$.
- **(a)** If we write x_j for $F(j)$, we see that X is of the form

$$X = F([n]) = \{F(j) : j \in [n]\} = \{x_j : j \in \mathbb{Z} \text{ and } 1 \leq j \leq n\}.$$

In other words, its elements can be enumerated as $x_1, x_2, \dots, x_n$. This is the mathematician's way of stating th

- **(b)** We say that the empty set $\emptyset$ has size $|\emptyset| = 0$.
- **(c)** We call a set X **finite**, if there exists $n \in [0, \infty[_{\mathbb{Z}}$ such that X has size n. Note that this implies that the empty set is finite. We say that X is **infinite** and we write $|X| = \infty$, if X is not finite.

- **(d)** Let X be a set such that there is a bijection $f : \mathbb{N} \xrightarrow{\sim} X$. In other words, all of the elements of X can be arranged in a sequence $(x_n)_{n \in \mathbb{N}}$ such that

$$X = \{ x_n : n \in \mathbb{N}, x_n = f(n) \}.$$

Then we call X a **countably infinite** set.
- **(e)** We call a set that is either finite or countably infinite a **countable** set.

(f) A set that is not countable is called **uncountable**

(g) We use the phrase **"finitely many"** items, **"countably many"** items, **"infinitely many"** items, etc., if they would constitute a finite set, a countable set, an infinite set, etc. □

Proposition 7.2. *A countably infinite set is infinite (and not finite).*

Proposition 7.3.
Let X and Y be two nonempty sets with a bijection $f : X \xrightarrow{\sim} Y$. Then

(a) *Y is finite if and only if X is finite,*
(b) *Y is countably infinite if and only if X is countably infinite,*
(c) *Y is countable if and only if X is countable,*
(d) *Y is uncountable if and only if X is uncountable.*
(e) *$|Y| = |X|$.*

Proposition 7.4.
let A and B two mutually disjoint, finite sets. Then $A \uplus B$ is finite and

$$|A \uplus B| = |A| + |B| .$$

Proposition 7.5. *Let $n \in \mathbb{Z}_{\geq 0}$. Let Ω be a set such that $|\Omega| = n$. Then its power set has size $|2^\Omega| = 2^n$.*

7.2 The Subsets of $\mathbb{N}$ and Their Size

Proposition 7.6. *Let $\emptyset \neq A \subseteq \mathbb{N}$ and let $A_j \subseteq A$ and $a_j \in A$ $(j \in \mathbb{N})$ be recursively defined as follows.*

$$A_1 := A, \qquad a_1 := \min(A_1); \tag{7.1}$$

$$A_{n+1} := A \setminus \{a_j : j \in \mathbb{N}, j \leq n\}; \qquad a_{n+1} := \begin{cases} \min(A_{n+1}) & \text{if } A_{n+1} \neq \emptyset, \\ a_n & \text{else.} \end{cases} \tag{7.2}$$

The following is true for all $i, j, n \in \mathbb{N}$.

(a) *The sequence of sets $A_1, A_2, A_3 \dots$ is nonincreasing: if $i < j$ then $A_i \supseteq A_j$.*
(b) *If $j < n$ and $A_n \neq \emptyset$ then $a_j < a_n$.*
(c) *If $A_n \neq \emptyset$ then $a_n \geq n$.*
(d) *Let $n \geq 2$. If $a \in A$ and $a < a_n$ then $a = a_j$ for some $j < n$.*
(e) *Let $n \in \mathbb{N}$. There is no $a \in A$ such that $a_n < a < a_{n+1}$.*

(f) *If $A_n = \emptyset$ for some $n \in \mathbb{N}$ then A is bounded. Let $K := \max\{j \in \mathbb{N} : A_j \neq \emptyset\}$. Then $\max(A) = a_K$. Figure* **??** *illustrates this for the case $K = 4$. Moreover,*

$$A = \{a_j : j \in \mathbb{N}, j \leq K\} = \{\min(A_j) : j \in \mathbb{N}, j \leq K\}, \tag{7.3}$$

$$\text{If } n \geq K \text{ then } a_n = a_K. \tag{7.4}$$

(g) *The sequence $a_j : j \in \mathbb{N}$ is nondecreasing: if $i < j$ then $a_i \leq a_j$.*

(h) *If $A_n \neq \emptyset$ for all $n \in \mathbb{N}$ then A is unbounded and*

$$A = \{a_j : j \in \mathbb{N}\} = \{\min(A_j) : j \in \mathbb{N}\}.$$

Proposition 7.7. *Let A be a nonempty subset of $\mathbb{N}$. Let $A_j \subseteq A$ and $a_j \in A$ ($j \in \mathbb{N}$) be defined as in prop. 7.6 on p.67. Then*

- *either $A_n \neq \emptyset$ for all $n \in \mathbb{N}$. In this case A is not bounded and there exists a bijection $\mathbb{N} \xrightarrow{\sim} A$. Further $A = \{a_n : n \in \mathbb{N}\}$*
- *or A_n is empty for some $n \in \mathbb{N}$. In this case A is bounded and there exists a bijection $[1, K]_{\mathbb{Z}} \xrightarrow{\sim} A$ for some suitable $K \in \mathbb{N}$. Further $A = \{a_n : n \in \mathbb{N}$ such that $1 \leq n \leq K\}$*

In both cases the integers a_n and a_{n+1} are adjacent for each index n in the sense that there is no $a \in A$ such that $a_n < a < a_{n+1}$.

Proposition 7.8. *Let J be a nonempty set of integers which is bounded below.Then*

(a) *If J is bounded above then there exists $K \in \mathbb{N}$ and integers n_j ($1 \leq j \leq K$) such that $J = \{n_j : 1 \leq j \leq K\}$.*

(b) *If J is not bounded above then there exist integers n_j ($j \in \mathbb{N}$) such that $J = \{n_j : j \in \mathbb{N}\}$.*

(c) *In both cases **(a)** and **(b)** the integers n_j satisfy $i < j \Rightarrow n_i < n_j$, and n_j and n_{j+1} are adjacent for each index j: There is no $n \in J$ such that $n_j < n < n_{j+1}$.*

Notation 7.1 (Notation Alert for bounded below subsets of the integers)**.**
If J is a nonempty subset of the integers which is bounded below then the last proposition makes it natural to introduce the following notation:

(a) If J is finite, i.e., bounded above and hence of the form $J = \{n_j : 1 \leq j \leq K\}$ then we also say that J consists of the numbers $n_1 < n_2 < ... < n_K$.

(b) If J is infinite, i.e., not bounded above and hence of the form $J = \{n_j : j \in \mathbb{N}\}$ then we also say that J consists of the numbers $n_1 < n_2 < ....$ ☐

Proposition 7.9. *Let A be a nonempty, finite subset of $\mathbb{N}$. Then A is bounded.*

Proposition 7.10. *Let $B \subseteq A \subseteq \mathbb{N}$ and assume that A is finite. Then B is finite.*

Theorem 7.1. *Let A be a nonempty subset of the natural numbers. Then*

(a) *A is finite if and only if A is bounded,*
(b) *A is countably infinite if and only if A is not bounded.*
(c) *All subsets of $\mathbb{N}$ are countable.*

Theorem 7.2.

(a) *Let X be a finite set and $A \subseteq X$. Then A is finite.*
(b) *Let $X_1, X_2, \dots, X_n$ be finite sets. Then $\bigcup_{j=1}^{n} X_j$ is finite.*

Theorem 7.3. *Let A be a nonempty set of integers. Then*

(a) *A is finite if and only if A is bounded,*
(b) *A is countably infinite if and only if A is not bounded.*

7.3 Finite Sequences and Subsequences and Eventually True Properties

Definition 7.2 (Finite sequences)**.** Let $n_\star, n^\star \in \mathbb{Z}$ such that $n_\star \leq n^\star$, let $J := [n_\star, n^\star]_{\mathbb{Z}}$. Then J is a finite set of integers since it is bounded below by $n_\star$ and above by $n^\star$. Let X be a nonempty set. We call an indexed family $(x_n)_{n \in J}$ in X with index set J a **finite sequence**. We write

$$(x_n)_{n_\star \leq n \leq n^\star} \quad \text{or} \quad (x_n)_{n=n_\star}^{n^\star} \quad \text{or} \quad x_{n_\star}, x_{n_\star+1}, \dots, x_{n^\star-1}, x_{n^\star} \quad \text{or} \quad (x_{n_\star}, x_{n_\star+1}, \dots, x_{n^\star-1}, x_{n^\star})$$

for such a finite sequence. We will sometimes call a sequence $(y_n)_{n=n_\star}^{\infty}$ an **infinite sequence** if we want to stress that its set of indices $[n_\star, \infty[$ is infinite.

If all members x_j of the finite sequence are (real) numbers then we also talk about a **vector** [9] of dimension $\left|[n_\star, n^\star]_{\mathbb{Z}}\right| = n^\star - n_\star + 1$. In this case we always must surround the members of that finite sequence with parentheses, and we will often use a symbol with "arrow notation"

$$\vec{x} = (x_1, x_2, x_3, \dots, x_{n-1}, x_n) \tag{7.5}$$

when working with such vectors. $\square$

Definition 7.3 (Finite subsequences)**.** Assume that either $J := [n_\star, \infty[_{\mathbb{Z}}$ or $J := [n_\star, n^\star]_{\mathbb{Z}}$ ($n_\star, n^\star \in \mathbb{Z}$ and $n_\star \leq n^\star$). Let $(n_j)_{j=1}^K (K \in \mathbb{N})$ be a finite sequence of integers $n_j \in J$ such that $i < j \Rightarrow n_i < n_j$ for all $i, j \in \mathbb{N}$. Note that if $J = [n_\star, \infty[_{\mathbb{Z}}$ then $n_j \in J$ for all j implies $n_\star \leq n_1 < n_2 < \cdots < n_K$, and if $J = [n_\star, n^\star]_{\mathbb{Z}}$ then this implies $n_\star \leq n_1 < n_2 < \cdots < n_K \leq n^\star$. Let $(x_n)_{n \in J}$ be a sequence in a nonempty set X. We call $(x_{n_j})_{j=1}^K$ a **finite subsequence** of the original sequence since its index set $\{n_j : 1 \leq j \leq K\}$ is finite and we obtain $(x_{n_j})_{j=1}^K$ from $(x_n)_{n \in J}$ by omitting all members x_n for which there is no n_j which equals n. □

Definition 7.4. Let X be a nonempty set, $n_\star \in \mathbb{Z}$, $J := \{k \in \mathbb{Z} : k \geq n_\star\}$, and let $(x_n)_{n=n_\star}^\infty$ be a sequence in X. If the set of indices $n \in J$ for which a certain property does not hold is empty or bounded then we say that the sequence $(x_n)_n$ satisfies this property **eventually** or that it satisfies this property for **eventually all indices** n. □

Proposition 7.11. *We have the following equivalent ways to state that a sequence (x_n) satisfies a property P eventually:*

- ***(a)*** *There is $K \in J$ such that if P is false for some x_j then $j \leq K$.*
- ***(b)*** *There is $K \in J$ such that P is true for all x_j such that $j > K$.*
- ***(c)*** *The set of all indices j such that P is false for x_j is finite.*

7.4 Countable Sets

Proposition 7.12 (Countability Criterion)**.** *Let $X \neq \emptyset$.*

The following are equivalent:

- ***(a)*** *X is countable.*
- ***(b)*** *There exists an injective function $f : X \to \mathbb{N}$.*
- ***(c)*** *There exists a surjective function $g : \mathbb{N} \to X$.*

Theorem 7.4. *Let X be a countable set and $A \subseteq X$. Then A is countable.*

Corollary 7.2.

- ***(a)*** *subsets of countable sets are either finite or countably infinite.*
- ***(b)*** *supersets of uncountable sets are uncountable.*
- ***(c)*** *Supersets of infinite sets are infinite,*

[9] Vectors can be of a more general nature than just being a finite sequence of numbers. See ch.11.2 (General Vector Spaces) on p.105 (General Vector Spaces).

Proposition 7.13 (B/G prop.13.11). *Every infinite set contains a proper subset that is countably infinite.*

Proposition 7.14 (B/G prop.13.12).
A set is infinite if and only if it contains a proper subset that is countably infinite.

Proposition 7.15 (B/G Cor.13.16, p.122). $\mathbb{N}^2$ *is countable.*

Proposition 7.16. *Let* $n \in \mathbb{N}$. *Then*

- ***(a)*** *There exist unique* $k \in \mathbb{Z}_{\geq 0}$ *and* $m \in \mathbb{N}$ *such that* m *is odd and* $n = 2^k m$.
- ***(b)*** *If* $n \neq 1$ *then* k *is the number of times the factor 2 occurs in its prime factorization. Further, either* m *is the product of all other prime factors, or* $m = 1$ *if there are no prime factors different from* 2.

Proposition 7.17.

- ***(a)*** *The function* $G : \left([0, \infty[_{\mathbb{Z}}\right)^2 \to \mathbb{N}$; $(i, j) \mapsto 2^i\,(2j+1)$ *is a bijection.*
- ***(b)*** *The function* $F : \mathbb{N}^2 \to \mathbb{N}$; $(i, j) \mapsto 2^{i-1}\,(2j-1)$ *is a bijection.*

Theorem 7.5 (B/G prop.13.19: Countable unions of countable sets).

The union of countably many countable sets is countable.

Corollary 7.3. *Let the set* X *be uncountable and let* $A \subseteq X$ *be countable. Then the complement* $A^{\complement}$ *of* A *is uncountable.*

Corollary 7.4. *The set* $\mathbb{Z}$ *of all integers is countable.*

Corollary 7.5. *The rational numbers are countable.*

Theorem 7.6 (Finite Cartesians of countable sets are countable).
The Cartesian product of finitely many countable sets is countable.

Corollary 7.6. *Let $n \in \mathbb{N}$. The sets $\mathbb{Q}^n$ and $\mathbb{Z}^n$ are countable.*

Theorem 7.7. *Let X be a set which contains at least two elements. Then $X^{\mathbb{N}} = \{(x_n)_{n\in\mathbb{N}} : x_j \in X \ \forall j \in \mathbb{N}\}$ (the set of all sequences with values in X) is uncountable.*

8 More on Sets, Relations, Functions and Families

8.1 More on Set Operations

Definition 8.1. We define

(8.1) $$\bigcup_{i\in\emptyset} A_i \;:=\; \emptyset, \qquad \text{If there is a universal set } \Omega\text{:} \bigcap_{i\in\emptyset} A_i \;:=\; \Omega. \;\square$$

Lemma 8.1 (Inclusion lemma)**.**
Let J be an arbitrary, nonempty index set. Let U, X_j, Y, Z_j, W $(j \in J)$ be sets such that

$$U \subseteq X_j \subseteq Y \subseteq Z_j \subseteq W$$

for all $j \in J$. Then

(8.2) $$U \subseteq \bigcap_{j\in J} X_j \subseteq Y \subseteq \bigcup_{j\in J} Z_j \subseteq W.$$

Definition 8.2 (Disjoint families)**.** Let J be a nonempty set. We call a family of sets $(A_i)_{i\in J}$ a **mutually disjoint family** if for any two different indices $i, j \in J$ it is true that $A_i \cap A_j = \emptyset$, i.e., if any two sets in that family with different indices are mutually disjoint. $\square$

Definition 8.3 (Partition)**.** Let J be an arbitrary nonempty set, let $(A_j)_{j\in J}$ be a family of subsets of Ω. We call $(A_j)_{j\in J}$ a **partition** or a **partitioning** of Ω if it is a mutually disjoint family which satisfies $\Omega = \biguplus \big[A_j : j \in J\big]$.

In other words,

- $(A_j)_{j\in J}$ is a partition of Ω if and only if $\mathfrak{A} := \{A_j : j \in J\}$ is a partition of Ω. $\square$

Theorem 8.1 (De Morgan's Law)**.** *Let there be a universal set Ω (see (2.8) on p.9). Then the following "duality principle" holds for any indexed family $(A_\alpha)_{\alpha\in I}$ of sets:*

(8.3) $$\textbf{(a)}\; \Big(\bigcup_\alpha A_\alpha\Big)^{\complement} = \bigcap_\alpha A_\alpha^{\complement} \qquad \textbf{(b)}\; \Big(\bigcap_\alpha A_\alpha\Big)^{\complement} = \bigcup_\alpha A_\alpha^{\complement}$$

Proposition 8.1 (Distributivity of unions and intersections). *Let $(A_i)_{i\in I}$ be an arbitrary family of sets and let B be a set. Then*

$$\bigcup_{i\in I}(B\cap A_i) = B\cap\bigcup_{i\in I}A_i, \tag{8.4}$$

$$\bigcap_{i\in I}(B\cup A_i) = B\cup\bigcap_{i\in I}A_i. \tag{8.5}$$

Proposition 8.2 (Rewrite unions as disjoint unions). *Let $(A_j)_{j\in\mathbb{N}}$ be a sequence of sets which all are contained within the universal set Ω. Let*

(a) $B_n := \bigcup_{j=1}^{n} A_j = A_1\cup A_2\cup\dots\cup A_n \; (n\in\mathbb{N})$,

(b) $C_1 := A_1 = B_1, \quad C_{n+1} := A_{n+1}\setminus B_n \; (n\in\mathbb{N})$.

Then,

(c) *The sequence $(B_j)_j$ is increasing: $m<n \Rightarrow B_m\subseteq B_n$.*

(d) *For each $n\in\mathbb{N}$, $\bigcup_{j=1}^{n} A_j = \bigcup_{j=1}^{n} B_j$.*

(e) *The sets C_j are mutually disjoint and $\bigcup_{j=1}^{n} A_j = \biguplus_{j=1}^{n} C_j$.*

(f) *The sets C_j $(j\in\mathbb{N})$ form a partition of the set $\bigcup_{j=1}^{\infty} A_j$.*

8.2 Rings and Algebras of Sets ⋆

Definition 8.4 (Rings, algebras, and σ–Algebras of Sets). A subset $\mathscr{R}$ of 2^Ω (a set of sets!) is called a **ring of sets** if it is closed with respect to the operations "$\cup$" and "$\setminus$", i.e.,

$$R_1\cup R_2\in\mathscr{R} \text{ and } R_1\setminus R_2\in\mathscr{R} \quad \text{whenever } R_1,R_2\in\mathscr{R}. \tag{8.6}$$

A subset $\mathscr{A}$ of 2^Ω is called an **algebra of sets** if $\Omega\in\mathscr{A}$ and $\mathscr{A}$ is a ring of sets.
A subset $\mathscr{F}$ of 2^Ω is called a σ**–algebra** if $\mathscr{F}$ is an algebra of sets which satisfies

$$(A_n)_{n\in\mathbb{N}}\in\mathfrak{F} \quad\Rightarrow\quad \bigcup_{n\in\mathbb{N}}A_n\in\mathfrak{F} \;\square$$

Proposition 8.3.
(1) *Let $\mathscr{R}$ be a ring of sets and $A,B\in\mathscr{R}$. Then $\emptyset\in\mathscr{R}$, $A\triangle B\in\mathscr{R}$, and $A\cap B\in\mathscr{R}$.*

(2) *Let* A, B, C, Ω *be sets such that* $A, B, C \subseteq \Omega$*. Then*

(a) $(A \triangle B) \triangle C = A \triangle (B \triangle C)$ *(associativity of* $\triangle$*)*

(b) $A \triangle \emptyset = \emptyset \triangle A = A$ *(neutral element* $\emptyset$ *for* $\triangle$*)*

(c) $A \triangle A = \emptyset$ *(inverse element* $A^{-1} = A$ *for* $\triangle$*)*

(d) $A \triangle B = B \triangle A$ *(commutativity of* $\triangle$*)*

Further, we have the following for the intersection operation:

(e) $(A \cap B) \cap C = A \cap (B \cap C)$ *(associativity of* $\cap$*)*

(f) $A \cap \Omega = \Omega \cap A = A$ *(neutral element* Ω *for* $\cap$*)*

(g) $A \cap B = B \cap A$ *(commutativity of* $\cap$*)*

Also, we have the following interrelationship between $\triangle$ *and* $\cap$*:*

(h) $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$ *(distributivity)*

Remark 8.1 (Algebras of Sets as Rings)**.**

(1) Prop.8.3**(1)** states that the assignments $(A, B) \mapsto A \triangle B$ and $(A, B) \mapsto A \cap B$ are binary operations on $\mathscr{R}$.

(2) Items **(a)** – **(d)** of prop.8.3**(2)** assert that $(\mathscr{R}, \triangle)$ is an abelian group with neutral element $\emptyset$ and inverse $A^{-1} = A$.

(3) If $\Omega \in \mathscr{R}$, i.e., $\mathscr{R}$ is an algebra of sets, Items **(e)** – **(g)** of prop.8.3**(2)** assert that $(\mathscr{R}, \cap)$ is a commutative monoid with unit Ω.

(4) Assume that Ω is not empty. Then the "additive" neutral elenent $\emptyset$ is different from Ω, the "multiplicative" neutral elenent.

(5) **(1)** – **(4)** plus Proposition 8.3**(2).(h)** imply that, if $\Omega \neq \emptyset$, then $(\mathscr{R}, \triangle, \cap)$ satisfies Definition 3.7 on p.24, i.e., $(\mathscr{R}, \triangle, \cap)$ is a commutative ring with unit.

8.3 Cartesian Products of More Than Two Sets

Definition 8.5 (Cartesian Product of a family of sets)**.** ⋆

Let I be an arbitrary, nonempty set (the index set). Let $(X_i)_{i \in I}$ be a family of nonempty sets X_i.

The **cartesian product** of the family $(X_i)_{i \in I}$ is the set

(8.7) $$\prod_{i \in I} X_i := \Big(\prod X_i \Big)_{i \in I} := \{ (x_i)_{i \in I} : x_k \in X_k \ \forall k \in I \}$$

of all familes $(x_i)_{i \in I}$ each of whose members x_j belongs to the corresponding set X_j.

$(x_i)_{i \in I}$, $(y_k)_{k \in I} \in \prod_{i \in I} X_i$ are called **equal** (we write $(x_i)_{i \in I} = (y_k)_{k \in I}$), if $x_j = y_j$ for all $j \in I$.

If all sets X_i are equal to one and the same set X, we also write

(8.8) $$X^I := \prod_{i \in I} X := \prod_{i \in I} X_i \,. \ \square$$

$$Y^X = \{f : f \text{ is a function with domain } X \text{ and codomain } Y\}. \quad \square \tag{8.9}$$

8.4 Set Operations involving Direct Images and Preimages

Unless stated otherwise, X, Y and f are as defined above for the remainder of this chapter: $f : X \to Y$ is a function with domain X and codomain Y.

Proposition 8.4 (f^{-1} is compatible with all basic set ops)**.** *Let J be an arbitrary index set. Let $B \subseteq Y$, $B_j \subseteq Y$ for all j. Then*

$$f^{-1}\Big(\bigcap_{j\in J} B_j\Big) = \bigcap_{j\in J} f^{-1}(B_j) \tag{8.10}$$

$$f^{-1}\Big(\bigcup_{j\in J} B_j\Big) = \bigcup_{j\in J} f^{-1}(B_j) \tag{8.11}$$

$$f^{-1}(B^{\complement}) = \big(f^{-1}(B)\big)^{\complement} \tag{8.12}$$

$$f^{-1}(B_1 \setminus B_2) = f^{-1}(B_1) \setminus f^{-1}(B_2) \tag{8.13}$$

$$f^{-1}(B_1 \Delta B_2) = f^{-1}(B_1) \Delta f^{-1}(B_2) \tag{8.14}$$

Proposition 8.5 (Properties of the direct image)**.** *Let J be an arbitrary index set. Let $A \subseteq X$, $A_j \subseteq X$ for all j. Then*

$$f\Big(\bigcap_{j\in J} A_j\Big) \subseteq \bigcap_{j\in J} f(A_j) \tag{8.15}$$

$$f\Big(\bigcup_{j\in J} A_j\Big) = \bigcup_{j\in J} f(A_j) \tag{8.16}$$

Proposition 8.6 (Direct images and preimages of function composition)**.** *Let X, Y, Z be arbitrary, nonempty sets.*
Let $f : X \to Y$ and $g : Y \to Z$, and let $U \subseteq X$ and $W \subseteq Z$. Then

$$(g \circ f)(U) = g\big(f(U)\big) \text{ for all } U \subseteq X. \tag{8.17}$$

$$(g \circ f)^{-1}(W) = f^{-1}\big(g^{-1}(W)\big) \text{ for all } W \subseteq Z, \text{ i.e., } (g \circ f)^{-1} = f^{-1} \circ g^{-1}. \tag{8.18}$$

Proposition 8.7 (Indirect image and fibers of f)**.** *Let X, Y be nonempty sets and let $f : X \to Y$ be a function. We define on the domain X a relation "$\sim$" as follows:*

$$x_1 \sim x_2 \quad \Leftrightarrow \quad f(x_1) = f(x_2). \tag{8.19}$$

(a) *"$\sim$" is an equivalence relation. Its equivalence classes, which we denote by $[x]_f$,* [10] *are*

$$[x]_f = \{a \in X : f(a) = f(x)\} = f^{-1}\{f(x)\}. \quad (x \in X) \tag{8.20}$$

(b) *If $A \subseteq X$ then*

$$f^{-1}(f(A)) = \bigcup_{a \in A} [a]_f. \tag{8.21}$$

Corollary 8.1.

$$\text{If } A \subseteq X \quad \text{then} \quad f^{-1}(f(A)) \supseteq A. \tag{8.22}$$

Proposition 8.8.

$$\text{If } B \subseteq Y \quad \text{then} \quad f(f^{-1}(B)) = B \cap f(X). \tag{8.23}$$

Corollary 8.2.

$$\text{If } B \subseteq Y \quad \text{then} \quad f(f^{-1}(B)) \subseteq B. \tag{8.24}$$

Proposition 8.9. ***(a)*** *Let $A \subseteq X$. If $f : X \to Y$ is injective then $f^{-1}(f(A)) = A$.*
(b) *Let $B \subseteq Y$. If $f : X \to Y$ is surjective then $f\big(f^{-1}(B)\big) = B$.*
(c) *Let $A \subseteq X$ and $B \subseteq Y$. If $f : X \to Y$ is injective and if $B = f(A)$ then $f^{-1}(B) = A$.*
(d) *Let $A \subseteq X$ and $B \subseteq Y$. If $f : X \to Y$ is surjective and if $f^{-1}(B) = A$ then $B = f(A)$.*
(e) *Let $A \subseteq X$ and $B \subseteq Y$. If $f : X \to Y$ is bijective then $B = f(A) \;\Leftrightarrow\; f^{-1}(B) = A$.*

Proposition 8.10. *Let J be an arbitrary nonempty index set and let $A \subseteq X$, $A_j \subseteq X$ for all j.*

Let $f : X \to Y$ be bijective. Then the following all are true:

$$f\Big(\bigcap_{j\in J} A_j\Big) = \bigcap_{j\in J} f(A_j) \tag{8.25}$$

$$f\Big(\bigcup_{j\in J} A_j\Big) = \bigcup_{j\in J} f(A_j) \tag{8.26}$$

$$f(A^{\complement}) = f(A)^{\complement} \tag{8.27}$$

$$f(A_1 \setminus A_2) = f(A_1) \setminus f(A_2) \tag{8.28}$$

$$f(A_1 \Delta A_2) = f(A_1) \Delta f(A_2) \tag{8.29}$$

8.5 Indicator Functions ⋆

Definition 8.6 (indicator function for a set)**.** Let Ω be "the" universal set, i.e., we restrict our scope of interest to subsets of Ω. Let $A \subseteq \Omega$. Let $\mathbf{1}_A : \Omega \to \{0,1\}$ be the function defined as

$$\mathbf{1}_A(\omega) := \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A. \end{cases} \tag{8.30}$$

$\mathbf{1}_A$ is called the **indicator function** of the set A. □

Proposition 8.11. *Let $\mathscr{F}(\Omega, \{0,1\}) := \{0,1\}^\Omega$ denote the set of all functions $f : \Omega \to \{0,1\}$, i.e., all functions f with domain Ω for which the only possible function values $f(\omega)$ are zero or one.* [11]

(a) *The mapping*

$$F : 2^\Omega \to \mathscr{F}(\Omega, \{0,1\}), \quad \text{defined as} \quad F(A) := \mathbf{1}_A \tag{8.31}$$

which assigns to each subset of Ω its indicator function is injective.

(b) *Let $f \in \mathscr{F}(\Omega, \{0,1\})$. Further, let $A := \{f = 1\} = f^{-1}(\{1\}) = \{a \in A : f(a) = 1\}$. Then $f = \mathbf{1}_A$.*

(c) *The function F above is bijective.*

Its inverse function is

$$G : \mathscr{F}(\Omega, \{0,1\}) \to 2^\Omega, \quad \text{defined as} \quad G(f) := \{f = 1\}. \tag{8.32}$$

Proposition 8.12. *Let $m, n, p \in \mathbb{Z}$. Then addition mod 2 is associative, i.e.,*

$$\big(m + n \mod 2\big) + p \mod 2 = m + \big(n + p \mod 2\big) \mod 2. \tag{8.33}$$

[11] See remark **??** on p.**??**, ch.8.3 (Cartesian Products of More Than Two Sets).

Proposition 8.13. *Let A, B, C be subsets of Ω. Then*

$$\mathbb{1}_{A\cup B} = \max(\mathbb{1}_A, \mathbb{1}_B), \tag{8.34}$$

$$\mathbb{1}_{A\cap B} = \min(\mathbb{1}_A, \mathbb{1}_B), \tag{8.35}$$

$$\mathbb{1}_{A^\complement} = 1 - \mathbb{1}_A, \tag{8.36}$$

$$\mathbb{1}_{A\triangle B} = \mathbb{1}_A + \mathbb{1}_B \mod 2. \tag{8.37}$$

Proposition 8.14 (Symmetric set differences $A\triangle B$ are associative)**.** *Let $A, B, C \subseteq \Omega$. Then*

$$(A\triangle B)\triangle C = A\triangle(B\triangle C). \tag{8.38}$$

9 The Real Numbers

9.1 The Ordered Fields of the Real and Rational Numbers

Definition 9.1 (Fields). ⋆

Let $(F, \oplus, \odot)$ be a commutative ring with unit (see Definition 3.7 on p.24) such that each nonzero element possesses an inverse element with respect to multiplication, i.e., the set $(F \setminus \{0\}, \odot)$ with neutral element 1 is an abelian group. Then we call $(F, \oplus, \odot)$ a **field**. ☐

Proposition 9.1 (B/G prop.8.6). *Let $(F, \oplus, \odot)$ be a field and $a, b \in F \setminus \{0\}$. Then*

$$(ab)^{-1} = b^{-1}a^{-1}.$$

Proposition 9.2. *Fields are integral domains.*

Corollary 9.1 (B/G prop.8.7). *Let $a, b, c \in F$ and $a \neq 0$. If $ab = ac$ then $b = c$.*

Theorem 9.1. *For $n \in \mathbb{N}$ the following holds true:*
The commutative ring with unit $(\mathbb{Z}_n, \oplus, \odot)$ is a field if and only if n is prime.

Definition 9.2 (Division and Quotients). Let a, b be elements of a field $(F, \oplus, \odot)$, and let $b \neq 0$. Since b possesses a unique multiplicative inverse b^{-1} (see rem.**??** on p.**??**) we can define the function

$$\text{div} : F \times (F \setminus \{0\}) \longrightarrow F; \qquad (a, b) \mapsto a \odot b^{-1}.$$

We call this function the **division** operation on F. It is customary to also write $\frac{a}{b}$ or a/b instead of $a \odot b^{-1}$, and we follow that convention. In particular we may also write $\frac{1}{b}$ instead of b^{-1}. As in the case of the integers we call a the **dividend** or **numerator**, b the **divisor** or **denominator**, and $\frac{a}{b}$ the **quotient** of the expression $\frac{a}{b}$. ☐

Proposition 9.3. *Let $(F, \oplus, \odot, P)$ be a field and let $a \in R$. If $a \neq 0$ then the function*

$$D : F \to F; \quad x \mapsto a \odot x,$$

is a bijection.

Proposition 9.4 (B/G prop.11.2)**.** *Let $a, b, c, d \in F$ such that $b, d \neq 0$.*

$$\text{If} \quad \frac{a}{b} = \frac{c}{d} \qquad \text{then} \quad ad = bc\,.$$

Proposition 9.5 (B/G prop.11.3)**.** *Let $a, b, c \in F$ such that $b, c \neq 0$. Then*

$$\frac{ac}{bc} = \frac{a}{b}\,.$$

Proposition 9.6 (B/G prop.11.6)**.** *Let $a, b, c, d \in F$ such that $b, d \neq 0$. Then*

$$\frac{a}{b} \oplus \frac{c}{d} = \frac{ad \oplus bc}{bd}\,. \qquad \text{In particular, } \frac{a}{b} \oplus \frac{(\ominus a)}{b} = 0\,.$$

Proposition 9.7. *Let $a, b, c, d \in F$ such that $b, d \neq 0$.*

$$\text{Then} \quad \frac{a}{b} \odot \frac{c}{d} = \frac{ac}{bd}. \qquad \text{In particular, } \left(\frac{b}{d}\right)^{-1} = \frac{d}{b}\,.$$

Definition 9.3 (Ordered fields)**.** ⋆ Let $(F, \oplus, \odot)$ be a field which is ordered by a positive cone P. Then we call $(F, \oplus, \odot, P)$ an **ordered field**. □

Proposition 9.8 (B/G prop.8.40)**.**

(a) *Let $a \in F$. Then $a > 0$ if and only if $a^{-1} > 0$, and $a < 0$ if and only if $a^{-1} < 0$.*

(b) *Let $a, b \in F$. If $0 < a < b$ then $0 < \frac{1}{b} < \frac{1}{a}$.*

Corollary 9.2 (B/G prop.11.7)**.** *Let $a, b \in F_{\neq 0}$. Then*

(a) $\frac{a}{b} > 0 \Leftrightarrow \frac{b}{a} > 0$ *and* $\frac{a}{b} < 0 \Leftrightarrow \frac{b}{a} < 0$,

(b) $\frac{a}{b} > 0 \Leftrightarrow$ *either both $a, b > 0$ or both $a, b < 0$.*

Theorem 9.2 (B/G thm.8.43)**.** *Let $a, b \in F$ such that $a < b$. Then*

$$a \ < \ \frac{a+b}{2} \ < \ b\,.$$

Theorem 9.3 (B/G thm.8.42)**.** *The positive cone P does not have a minimum.*

Axiom 9.1 (Real Numbers)**.** We postulate the existence of a set $\mathbb{R}$ which satisfies the following:

- **(a)** $\mathbb{R}$ is endowed with two binary operations "$+$" (called addition) and "$\cdot$" (called multiplication) and with a positive cone $\mathbb{R}_{>0}$ such that $(\mathbb{R}, +, \cdot, \mathbb{R}_{>0})$ is an ordered integral domain. As usual we denote the additive unit of this integral domain by 0 and its multiplicative unit by 1.
- **(b)** The set $\mathbb{R}_{\neq 0} = \{x \in \mathbb{R} : x \neq 0\}$ is a group with respect to multiplication; thus for each $x \in \mathbb{R}_{\neq 0}$ there exists a unique $x^{-1} \in \mathbb{R}_{\neq 0}$ such that $xx^{-1} = 1$.
- **(c)** $\mathbb{R}$ satisfies the **completeness axiom**: Any nonempty subset A of $\mathbb{R}$ which is bounded above possesses a supremum in $\mathbb{R}$ (i.e., $\sup(A) \neq \pm\infty$).

We call this set $\mathbb{R}$ the set of **real numbers**. ☐

Definition 9.4 (Rational numbers)**.** We call the set

$$\mathbb{Q} \ := \ \{\,n/d : n \in \mathbb{Z}, d \in \mathbb{N}\,\}$$

(this is a subset of $\mathbb{R}$!) the set of **rational numbers**.

In other words rational numbers are fractions of integers. ☐

Theorem 9.4 (The Rational Numbers are an Ordered Field)**.**

- ***(a)*** *The assignments $(a, b) \mapsto a + b$ and $(a, b) \mapsto a \cdot b$ are binary operations on $\mathbb{Q}$, i.e., sums and products of rational numbers are rational numbers.*
- ***(b)*** *The triplet $(\mathbb{Q}, +, \cdot)$ is an integral domain.*
- ***(c)*** *Let $\mathbb{Q}_{>0} := \mathbb{R}_{>0} \cap \mathbb{Q}$. Then $(\mathbb{Q}, +, \cdot, \mathbb{Q}_{>0})$ is an ordered integral domain which satisfies the following: if $a, b \in \mathbb{Q}$ then $a < b$ with respect to the ordering induced by $\mathbb{Q}_{>0}$ if and only if $a < b$ with respect to the ordering induced by $\mathbb{R}_{>0}$*
- ***(d)*** *$(\mathbb{Q}_{\neq 0}, \cdot)$ is a (commutative) group.*

Theorem 9.5 (B/G thm.10.1: $\mathbb{N}$ is unbounded in $\mathbb{R}$)**.** *For any $x \in \mathbb{R}$ there exists $n \in \mathbb{N}$ such that $n > x$, i.e., there are no upper bounds for $\mathbb{N}$ in $\mathbb{R}$.*

Corollary 9.3. *There are no upper bounds for $\mathbb{N}$ in $\mathbb{Q}$.*

Remark 9.1 (Contrasting $\mathbb{Z}$ and $\mathbb{R}$)**.**

The Integers:

(a) $\mathbb{Z} = (\mathbb{Z}, +, \cdot)$ is a commutative ring with unit

(b) Cancellation rule (no zero divisors: $\mathbb{Z}$ is an integral domain)

(c) Ordered by the positive cone $P := \mathbb{N}$

(d) Induction axiom: If $A \subseteq \mathbb{N}$ satisfies **(1)** $1 \in A$, **(2)** $\left[n \in A \Rightarrow n+1 \in A\right]$, then $A \supseteq \mathbb{N}$

The Real Numbers:

(a) $\mathbb{R} = (\mathbb{R}, +, \cdot)$ is a commutative ring with unit

(b) $(\mathbb{R}_{\neq 0}, \cdot)$ is an abelian group: each $x \neq 0$ has a multiplicative inverse $\frac{1}{x}$ (implies the cancellation rule, hence $\mathbb{R}$ is an integral domain)

(c) Ordered by the positive cone $P := \mathbb{R}_{>0}$

(d) Completeness axiom: If nonempty $A \subseteq \mathbb{R}$ has upper bounds then $\sup(A)$ exists (as an element of $\mathbb{R}$, i.e. $\sup(A) < \infty$) □

9.2 Minima, Maxima, Infima and Suprema in $\mathbb{R}$ and $\mathbb{Q}$

Remark 9.2. Let $A \subseteq \mathbb{R}$ be nonempty.

(a) If A is bounded above then it follows from the completeness axiom that its least upper bound $\sup(A) = \min(A_{\mathscr{uppb}})$ exists (see axiom 9.1 (Real Numbers) on p.82).

(b) If A is bounded below then it follows from the completeness axiom and cor.3.4 on p.35 that its greatest lower bound $\inf(A) = \max(A_{\mathscr{lowb}})$ exists.

The above is the core distinction between real numbers and rational numbers. There are bounded sets of rational numbers which do not possess a supremum in $\mathbb{Q}$. □

Proposition 9.9. *Let $A \subseteq B \subseteq \mathbb{R}$. Then $\inf(A) \geq \inf(B)$ and $\sup(A) \leq \sup(B)$.*

Proposition 9.10 (Supremum and infimum are positively homogeneous)**.** *Let A be a nonempty subset of $\mathbb{R}$ and let $\lambda \in \mathbb{R}_{\geq 0}$. If $\lambda > 0$ or if $\lambda = 0$ and $\sup(A) < \infty$ then*

$$\text{If } \lambda > 0 \text{ or if } \lambda = 0 \text{ and } \sup(A) < \infty \quad \text{then} \quad \sup(\lambda A) = \lambda \sup(A), \tag{9.1}$$

$$\text{If } \lambda > 0 \text{ or if } \lambda = 0 \text{ and } \inf(A) > -\infty \quad \text{then} \quad \inf(\lambda A) = \lambda \inf(A). \tag{9.2}$$

Definition 9.5 (bounded functions). ⋆

Given are a nonempty set X and a real-valued function f with domain X.

We call f **bounded above** if the image $f(X) = \{f(x) : x \in X\}$ is bounded above, i.e., if there exists a (possibly very large) number $\gamma_1 > 0$ such that

(9.3) $$f(x) < \gamma_1 \quad \text{for all arguments } x.$$

We call f **bounded below** if the image $f(X) = \{f(x) : x \in X\}$ is bounded below, i.e., if there exists $\gamma_2 > 0$ such that

(9.4) $$f(x) > -\gamma_2 \quad \text{for all arguments } x.$$

We call f a **bounded function** if it is both bounded above and below, i.e., if there exists $\gamma > 0$ such that

(9.5) $$|f(x)| < \gamma \quad \text{for all arguments } x. \ \square$$

Definition 9.6 (supremum and infimum of functions).

Let X be an arbitrary set, $A \subseteq X$ a subset of X, $f : X \to \mathbb{R}$ a real–valued function on X. Consider the set $f(A) = \{f(x) : x \in A\}$, the image of A under f.

The **supremum of** $f(\cdot)$ **on** A is defined as

(9.6) $$\sup_A f := \sup_{x \in A} f(x) := \sup f(A)$$

The **infimum of** $f(\cdot)$ **on** A is defined as

(9.7) $$\inf_A f := \inf_{x \in A} f(x) := \inf f(A). \ \square$$

Definition 9.7 (supremum and infimum of families).

The **supremum** and **infimum** of a family of real numbers $(x_i)_{i \in I}$ $(x_i)_{i \in I}$ are defined as

(9.8) $$\sup(x_i) := \sup_i (x_i) := \sup(x_i)_i := \sup(x_i)_{i \in I} := \sup_{i \in I} x_i := \sup\{x_i : i \in I\}.$$

(9.9) $$\inf(x_i) := \inf_i (x_i) := \inf(x_i)_i := \inf(x_i)_{i \in I} := \inf_{i \in I} x_i := \inf\{x_i : i \in I\}. \ \square$$

Definition 9.8 (supremum and infimum of sequences).

Let $I = [k_0, \infty[_{\mathbb{Z}}$ and $x_n \in \mathbb{R}$ for $n \in I$. **Supremum** and **infimum** of $(x_n)_{n \in I}$ are defined as

(9.10) $$\sup(x_n) := \sup(x_n)_{n \in I} := \sup_{n \in I} x_n = \sup\{x_n : n \in I\}$$

(9.11) $$\inf(x_n) := \inf(x_n)_{n \in I} := \inf_{n \in I} x_n = \inf\{x_n : n \in I\}. \ \square$$

Proposition 9.11. *Let X be a nonempty set and $\varphi, \psi : X \to \mathbb{R}$. Let $\emptyset \neq A \subseteq X$. Then*

$$\sup\{\varphi(x)+\psi(x) : x \in A\} \;\leq\; \sup\{\varphi(y) : y \in A\} \;+\; \sup\{\psi(z) : z \in A\}, \tag{9.12}$$

$$\inf\{\varphi(x)+\psi(x) : x \in A\} \;\geq\; \inf\{\varphi(y) : y \in A\} \;+\; \inf\{\psi(z) : z \in A\}. \tag{9.13}$$

9.3 Convergence and Continuity in $\mathbb{R}$

Definition 9.9 (convergence of sequences of real numbers [12])**.** Let $a \in \mathbb{R}$. We say that a sequence (x_n) of real numbers **converges** to a for $n \to \infty$ if the following is true:
For any $\delta \in]0, \infty[$ (no matter how small), there exists $n_0 \in \mathbb{N}$ such that

$$|a - x_j| \;<\; \delta \quad \text{for all } \; j \geq n_0. \tag{9.14}$$

We write either of

$$a \;=\; \lim_{n\to\infty} x_n \qquad \text{or} \qquad x_n \to a \;\text{ as } n \to \infty \tag{9.15}$$

and we call a the **limit** of the sequence (x_n). $\square$

(b) Definition 9.9 can be worded as follows:

- For any $\delta > 0$, $|a - x_j| \;<\; \delta|$, **eventually**.

Definition 9.10 (Open ε-Neighborhood in $\mathbb{R}$)**.** For $x_0 \in \mathbb{R}$ and $\varepsilon > 0$, let

$$N_\varepsilon(x_0) \;:=\;]x_0 - \varepsilon, x_0 + \varepsilon[\;=\; \{x \in \mathbb{R} : |x - x_0| < \varepsilon\}$$

be the set of all elements of $\mathbb{R}$ with a distance to x_0 of strictly less than the number ε (the open interval with center x_0 and radius ε from which the points on the boundary (those with distance equal to ε) are excluded).

(a) We call $N_\varepsilon(x_0)$ the ε**–neighborhood** of x_0. [13] $N_\varepsilon(x_0)$ is often called the **open** ε**–neighborhood** of x_0, to differentiate it from the closed interval $[\, x_0 - \varepsilon, x_0 + \varepsilon \,]$, which is also called the **closed** ε–neighborhood of x_0.

(b) Let $x, y \in \mathbb{R}$ and $\varepsilon > 0$. We say that x and y are ε**–close** if $|x - y| < \varepsilon$. $\square$

[12] We will define convergence of a sequence of items more general than real numbers in ch.12.4 (see Definition 12.10 (convergence of sequences in metric spaces) on p.119).

[13] This will generalized to metric spaces in Definition 12.6 on p.118.

There are two equivalent ways of expressing convergence to $a \in \mathbb{R}$:

(a) No matter how small a δ–neighborhood of a you choose: at most finitely many of the x_n will be located outside that neighborhood.

(b) No matter how small a δ–neighborhood of a you choose: eventually all of the x_n will be found inside that neighborhood.

Definition 9.11 (Limit infinity). ⋆

Given a real number $K > 0$, we define

(9.16a) $$N_K(\infty) := \{x \in \mathbb{R} : x > K\}$$

(9.16b) $$N_K(-\infty) := \{x \in \mathbb{R} : x < -K\}$$

We call $N_K(\infty)$ the K**–neighborhood of** ∞ and $N_K(-\infty)$ the K**–neighborhood of** $-\infty$. We say that a sequence (x_n) has limit ∞ and we write either of

(9.17) $$x_n \to \infty \qquad \text{or} \qquad \lim_{n\to\infty} x_n = \infty$$

if the following is true for any $K \in \mathbb{R}$ (no matter how big): There is an integer n_0 such that all x_j belong to $N_K(\infty)$ for all $j \geq n_0$, i.e., if

for all $K \in \mathbb{N}$ there exists $n_0 \in \mathbb{N}$ such that if $j \geq n_0$ then $x_j > K$.

We say that the sequence (x_n) has limit $-\infty$ and we write either of

(9.18) $$x_n \to -\infty \qquad \text{or} \qquad \lim_{n\to\infty} x_n = -\infty$$

if the following is true for any $K \in \mathbb{R}$ (no matter how big): There is an integer n_0 such that all x_j belong to $N_K(-\infty)$ for all $j \geq n_0$. □

(a) There is an equivalent way of stating that the sequence (x_n) has limit ∞: No matter how big a threshold $K > 0$ you choose: eventually all of the x_n will be located above that threshold.

(b) $x_n \to -\infty$ can also be expressed as follows: No matter how big a threshold $K > 0$ you choose: eventually all of the x_n will be located below $-K$.

Remark 9.3. The majority of mathematicians agrees that there is no "convergence to ∞" or "divergence to ∞". Rather, they say that a sequence has the limit ∞. We will follow that convention in this document. □

Theorem 9.6 (Limits are uniquely determined). *Let $(x_n)_n$ be a convergent sequence of real numbers. Then its limit is uniquely determined.*

Proposition 9.12 (B/G prop.10.11). *Let $a, b \in \mathbb{R}$. Then $a = b \Leftrightarrow |a - b| < \varepsilon$ for all $\varepsilon > 0$.*

Proposition 9.13 (Subsequences of real number sequences with limits). *Let $(x_n)_n$ be a sequence of real numbers with limit $L := \lim\limits_{n\to\infty} x_n$. Let (x_{n_j}) be a subsequence. Then $\lim\limits_{j\to\infty} x_{n_j} = L$.*

Note 9.1 (Notation for limits of monotone sequences).
Let (x_n) be a nondecreasing and y_n a nonincreasing sequence of real numbers.

(a) If $\xi = \lim\limits_{k\to\infty} x_k$ (that limit might be $+\infty$), then we write • $x_n \uparrow \xi \quad (n \to \infty)$

(b) If $\eta = \lim\limits_{j\to\infty} y_j$ (that limit might be $-\infty$), then we write • $y_j \downarrow \eta \quad (j \to \infty)$. □

Proposition 9.14. *[See B/G prop.10.16]*

Let $(x_n)_n$ be a sequence of real numbers such that $\lim\limits_{n\to\infty} x_n$ exists. Let $K \in \mathbb{N}$. For $n \in \mathbb{N}$ let $y_n := x_{n+K}$. Then $(y_n)_n$ has the same limit as $(x_n)_n$.

Proposition 9.15 (convergent ⇒ bounded). *Let $(x_n)_n$ be a sequence in $\mathbb{R}$.*

- *If the sequence converges, then it is bounded.*

Proposition 9.16 (bounded times zero–convergent is zero–convergent). *Let $(x_n)_n$ and $(\alpha_n)_n$ be two sequences in $\mathbb{R}$ and let $\alpha \in \mathbb{R}$.*

- *If $\lim\limits_{n\to\infty} x_n = 0$ and if $|\alpha_j| \leq \alpha$ for all $j \in \mathbb{N}$, then*

$$\lim_{j\to\infty} (\alpha_j x_j) = 0. \tag{9.19}$$

Proposition 9.17 (Rules of arithmetic for limits). *Let $(x_n)_n$ and $(y_n)_n$ be sequences in $\mathbb{R}$ and $x, y, \alpha \in \mathbb{R}$. Let $\lim\limits_{j\to\infty} x_j = x$ and $\lim\limits_{j\to\infty} y_j = y$. Then*

(a) $\lim_{j\to\infty} \alpha = \alpha$,
(b) $\lim_{j\to\infty} (\alpha \cdot x_j) = \alpha \cdot x$, *(constant sequence)*
(c) $\lim_{j\to\infty} (x_j + y_j) = x + y$,
(d) $\lim_{j\to\infty} (x_j \cdot y_j) = x \cdot y$,
(e) *if* $x \neq 0$ *then* $\lim_{j\to\infty} \frac{1}{x_j} = \frac{1}{x}$.

Proposition 9.18.
(a) *Let x_n be a sequence of real numbers that is nondecreasing, i.e., $x_n \leq x_{n+1}$ for all n (see def. 18.1 on p.161), and which is bounded above. Then $\lim_{n\to\infty} x_n$ exists and coincides with* $\sup\{x_n : n \in \mathbb{N}\}$
(b) *If y_n is a sequence of real numbers that is nonincreasing, i.e., $y_n \geq y_{n+1}$ for all n, and which is bounded below. Then $\lim_{n\to\infty} y_n$ exists and coincides with* $\inf\{y_n : n \in \mathbb{N}\}$.

Proposition 9.19 (Domination Theorem for Limits)**.**
Let $x_n, y_n \in \mathbb{R}$ be two sequences of real numbers both of which have limits. Assume there is $K \in \mathbb{N}$ such that $x_n \leq y_n$ for all $n \geq K$. Then
$$\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n.$$

Corollary 9.4. *Let $x_n, y_n \in \mathbb{R}$ be two sequences of real numbers and $L \in \mathbb{R}$. Assume there is $K \in \mathbb{N}$ such that $x_n = y_n$ for all $n \geq K$. Then*
$$\lim_{n\to\infty} x_n = L \Leftrightarrow \lim_{n\to\infty} y_n = L, \qquad \lim_{n\to\infty} x_n = \pm\infty \Leftrightarrow \lim_{n\to\infty} y_n = \pm\infty.$$

Proposition 9.20. *Let $a, b \in \mathbb{R}$. Then*

$$[a,b] = \bigcap_{n\in\mathbb{N}} \left]a - \frac{1}{n}, b + \frac{1}{n}\right[. \tag{9.20}$$

$$]a,b[= \bigcup_{n\in\mathbb{N}} \left[a + \frac{1}{n}, b - \frac{1}{n}\right], \tag{9.21}$$

Definition 9.12 (Continuity in $\mathbb{R}$)**.** Let $A \subseteq \mathbb{R}$, $x_0 \in A$, and let $f : A \to \mathbb{R}$.
We say that f is **continuous at x_0** and we write

$$\lim_{x \to x_0} f(x) = f(x_0) \tag{9.22}$$

if **any** sequence (x_n) with values in A satisfies the following:

$$\text{if } x_n \to x_0 \text{ then } f(x_n) \to f(x_0).[14] \tag{9.23}$$

In other words, the following must be true for any sequence (x_n) in A:

$$\lim_{n\to\infty} x_n = x_0 \quad \Rightarrow \quad \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x_0). \tag{9.24}$$

We say that f is **continuous** if f is continuous at x_0 for all $x_0 \in A$. □

Proposition 9.21. *Let $A \subseteq \mathbb{R}$ and $\gamma \in \mathbb{R}$. The following functions $A \to \mathbb{R}$ are continuous.*

- ***(a)*** *The constant function $x \mapsto \gamma$,*
- ***(b)*** *The identity function $id\Big|_A : x \mapsto x$.*

Theorem 9.7 (Rules of arithmetic for continuous real–valued functions with domain in $\mathbb{R}$)**.** *Let $A \subseteq \mathbb{R}$ and $\alpha \in \mathbb{R}$. Assume that the functions*

$$f(\cdot), g(\cdot), f_1(\cdot), f_2(\cdot), f_3(\cdot), \dots, f_n(\cdot) : A \longrightarrow \mathbb{R}$$

all are continuous at $x_0 \in A$. Then

- ***(a)*** *Constant functions are continuous everywhere on A.*
- ***(b)*** *The product $fg(\cdot) : x \mapsto f(x)g(x)$ is continuous at x_0. Specifically, $\alpha f(\cdot) : x \mapsto \alpha \cdot f(x)$ is continuous at x_0. In particular $-f(\cdot) : x \mapsto -f(x) = (-1) \cdot f(x)$ is continuous at x_0.*
- ***(c)*** *The sum $f + g(\cdot) : x \mapsto f(x) + g(x)$ is continuous at x_0.*
- ***(d)*** *If $g(x_0) \neq 0$ then the quotient $f/g(\cdot) : x \mapsto f(x)/g(x)$ is continuous at x_0.*
- ***(e)*** *Any linear combination $\sum_{j=0}^{n} a_j f_j(\cdot) : x \mapsto \sum_{j=0}^{n} a_j f_j(x)$ is continuous in x_0.*

Proposition 9.22. *All polynomials are continuous*

[14] Since dontinuity is expressed here in terms of sequences, we speak of the **sequence continuity** of a function. See Definition 13.1 (Sequence continuity) on p.132 where continuity is generalized to metric spaces.

Proposition 9.23 (The composition of continuous functions is continuous)**.**
Let $A, B \subseteq \mathbb{R}$ be nonempty, $f : A \to \mathbb{R}$ continuous at $x_0 \in A$, and $g : B \to \mathbb{R}$ continuous at $f(x_0)$. Assume further that $f(A) \subseteq B$, i.e., $f(x) \in B$ for all $x \in A$.

Then the composition $g \circ f : X \longrightarrow Y$ is continuous at x_0.

Theorem 9.8. *Let $A \subseteq \mathbb{R}$, $x_0 \in A$, and let $f : A \to \mathbb{R}$ be a real–valued function with domain A. Then f is continuous at x_0 if and only if for any $\varepsilon > 0$, no matter how small, there exists $\delta > 0$ such that either one of the following equivalent statements is satisfied:*

$$f\left(N_\delta(x_0) \cap A\right) \subseteq N_\varepsilon\big(f(x_0)\big)\,, \tag{9.25}$$

$$f\big(\{x \in A : |x - x_0| < \delta\}\big) \subseteq \{y \in \mathbb{R} : |y - f(x_0)| < \varepsilon\}, \tag{9.26}$$

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon \text{ for all } x \in A. \tag{9.27}$$

Proposition 9.24. *Let $A \subseteq \mathbb{R}$, $x_0 \in A$, and let $f : A \to \mathbb{R}$ be a real–valued function with domain A. Then f is continuous at x_0 if and only if there exists $\varepsilon^\star > 0$ which satisfies the following: for any $\varepsilon \in]0, \varepsilon^\star[$ there exists $\delta > 0$ such that either one of the following equivalent statements is satisfied:*

(a) $f\big(\{x \in A : |x - x_0| < \delta\}\big) \subseteq \{y \in \mathbb{R} : |y - f(x_0)| < \varepsilon\}$,
(b) $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$ *for all* $x \in A$.

9.4 Rational and Irrational Numbers

Proposition 9.25 (B/G thm.10.25)**.**
Let $A := \{a \in \mathbb{R}_{>0} : a^2 < 2\}$. Then $r := \sup(A)$ exists and $r^2 = 2$.

Definition 9.13 (Lowest terms representation of rational numbers)**.**
We repeat the following from Definition 2.15 on page 12 of Chapter 2.
Let $q := \frac{d}{n}$ $(d, n \in \mathbb{Z}, d \neq 0)$ be a rational number. We say that d and n are a representation of q in **lowest terms** or that q is written in lowest terms if

a. d and n have no common factors,
b. $n \in \mathbb{N}$. $\square$

Proposition 9.26. *Let* $q = \dfrac{m}{n}$ $(m, n \in \mathbb{Z}, n \neq 0)$ *be a nonzero rational number. Then* q *is written in lowest terms if and only if* $n \in \mathbb{N}$ *and* m *and* n *are relatively prime.*

Proposition 9.27 (B/G prop.11.5)**.** *Let* $m, n, s, t \in \mathbb{Z}$ *be such that* m *and* n *do not have any common factors.*

$$\text{If } \frac{m}{n} = \frac{s}{t}, \qquad \text{then } m \text{ divides } s \text{ and } n \text{ divides } t.$$

Proposition 9.28 (B/G prop.11.10)**.** *The real number* $\sqrt{2}$ *is irrational.*

Definition 9.14 (Perfect Squares)**.** Let $n \in \mathbb{Z}$ We call n a **perfect square** if there exists $k \in \mathbb{Z}$ such that $n = k^2$. In other words, the set of all perfect squares is the set $0, 1, 4, 9, \ldots$. ☐

Theorem 9.9 (B/G thm.11.12)**.** *Let* $n \in \mathbb{Z}_{\geq 0}$*. If* n *is not a perfect square then* $\sqrt{n}$ *is irrational.*

If n is a nonnegative integer then its square root is either an integer or irrational.

Proposition 9.29 (B/G prop.11.13)**.** *Let* m *and* n *be nonzero integers. Then* $\frac{m}{n}\sqrt{2}$ *is irrational.*

Theorem 9.10 (B/G ch.11: n–th root)**.** *Let* n *be an integer* ≥ 2 *and* $x \in \mathbb{R}_{>0}$*. Then there exists* $r \in \mathbb{R}_{>0}$ *such that* $r^n = x$ *and* r *is uniquely determined.*

Definition 9.15 (n–th root)**.** Let n be an integer ≥ 2 and $x \in \mathbb{R}_{>0}$. We write $\sqrt[n]{x}$ for the uniquely defined $r \in \mathbb{R}_{>0}$ such that $r^n = x$, and we extend this definition to $n = 1$ by defining $\sqrt[1]{x} := x$. We call $\sqrt[n]{x}$ the n**–th root** of x. ☐

Proposition 9.30 (B/G prop.11.16)**.** *Let* $n \in \mathbb{Z}_{\geq 2}$*. Then* $\sqrt[n]{2}$ *is irrational.*

Proposition 9.31 (B/G prop.11.17). *Let $x, y \in \mathbb{R}$ such that $x < y$. Then there exists irrational z such that $x < z < y$.*

Proposition 9.32 (B/G cor.11.18). *There is no smallest positive irrational number.*

9.5 Geometric Series

Definition 9.16 (Real–valued Sequences and Series). ⋆

A sequence (a_j) is called a **real–valued sequence** if each a_j is a real number.
For any such sequence, we can build another sequence (s_n) as follows:

(9.28) $$s_1 := a_1; \quad s_2 := a_1 + a_2; \quad s_3 := a_1 + a_2 + a_3; \cdots \quad s_n := \sum_{k=1}^{n} a_k$$

We write this more compactly as

(9.29) $$a_1 + a_2 + a_3 + \cdots = \sum a_k,$$

and we call any such object which represents a sequence of partial sums a **series**. Loosely speaking, a series is a sum of infinitely many terms. We call (s_n) the sequence of **partial sums** associated with the series $\sum a_k$.
Let $s \in \mathbb{R}$. We say that the **series converges** to s and we write

(9.30) $$\sum_{k=1}^{\infty} a_k = s$$

if this is true for the associated sequence of partial sums (9.28), i.e., if $\lim\limits_{n\to\infty} s_n = s$. We then also say that the **series has limit** s.
We say that the **series has limit** $\pm\infty$ if $\lim\limits_{n\to\infty} s_n = \pm\infty$. In this case we write

(9.31) $$\sum_{k=1}^{\infty} a_k = \pm\infty.$$

We adopt for series the convention we did in rem.9.3 on p.86 for sequences: A series with limit $-\infty$ or ∞ never ever converges or diverges to $\pm\infty$. Instead we say that $\sum a_k$ diverges. □

Proposition 9.33 (Limits of Geometric Series).

(a) *Let $|q| < 1$. Then $\lim\limits_{j\to\infty} q^n = 0$.*

(b) (9.32) $$\sum_{j=0}^{n} q^j = \frac{1 - q^{n+1}}{1 - q},$$

(c) (9.33) $$\sum_{j=0}^{\infty} q^j = \frac{1}{1 - q}.$$

9.6 Decimal Expansions of Real and Rational Numbers

Notation 9.1 (Decimal digits). Note that $[0,9]_{\mathbb{Z}}$ is according to notations 2.1 on p.14 (and also according to Definition 3.12 on p.30 equal to the set $\{0,1,2,3,4,5,6,7,8,9\}$ of decimal digits. $\square$

Definition 9.17 (Decimal Expansion). **A:** Let $x \in \mathbb{R}_{\geq 0}$, $d_0 \in \mathbb{Z}_{\geq 0}$, and $(d_j)_{j\in\mathbb{N}}$ a sequence of decimal digits d_j such that

$$x = d_0 + \sum_{j=1}^{\infty} d_j 10^{-j} = \sum_{j=0}^{\infty} d_j 10^{-j}. \tag{9.34}$$

Then we call both the word $d_0.d_1d_2d_3\dots$ (of infinite length) and also the corresponding sequence $(d_0, d_1, d_2, \dots) = (d_j)_{j=0}^{\infty}$ a **decimal expansion** of the nonnegative real number x.

B: We do not distinguish between $\sum_{j=0}^{\infty} d_j 10^{-j}$, $d_0.d_1d_2d_3\dots$, and $(d_j)_{j=0}^{\infty}$ and think of those expression as different notations for the same real number.
We extend the above definition to $x \in \mathbb{R}_{<0}$ as follows. If $-x$ has a decimal expansion $-x = d_0 + \sum_{j=1}^{\infty} d_j 10^{-j}$ then we call the word $-d_0.d_1d_2d_3\dots$ and also the corresponding sequence $(-d_0, d_1, d_2, \dots) = -d_0, (d_j)_{j=1}^{\infty}$ a decimal expansion of x.
We may omit leading zeros of the integer d_0 and trailing zeros of the digits $d_1, d_2, \dots$. We further may omit the decimal point together with all digits d_j to the right of that decimal point if $d_j = 0$ for all $j \in \mathbb{N}$. In other words, if $x = \sum_{j=0}^{\infty} d_j 10^{-j}$ and if $d_j = 0$ for all $j \in \mathbb{N}$ then we may write either of d_0, $d_0.$, or $d_0.0$ for x. $\square$

Proposition 9.34 (Geometric series for decimals). *Let $n \in \mathbb{N}$ and $d_j \in [0,9]_{\mathbb{Z}}$ for $j \geq n$. Then,*

(a) $$0 \leq 9 \sum_{j=n}^{\infty} 10^{-j} = \frac{1}{10^{n-1}},$$

(b) $$\sum_{j=n}^{\infty} d_j 10^{-j} \leq \frac{1}{10^{n-1}},$$

(c) $$\sum_{j=n}^{\infty} d_j 10^{-j} = \frac{1}{10^{n-1}} \quad \Leftrightarrow \quad d_j = 9 \text{ for all } j \geq n.$$

Theorem 9.11 (Existence of Decimal Expansions (B/G thm.12.6)). *Every real number has a decimal expansion.*

Theorem 9.12 (Uniqueness of Decimal Expansions (B/G thm.12.8)). *Let $x \in \mathbb{R}_{\geq 0}$ have two different decimal representations*

$$x = d_0 + \sum_{j=1}^{\infty} \frac{d_j}{10^j} = e_0 + \sum_{j=1}^{\infty} \frac{e_j}{10^j}, \tag{9.35}$$

where $d_0, e_0 \in [0, \infty[_{\mathbb{Z}}$ and $d_j, e_j \in [0, 9]_{\mathbb{Z}}$ for all $j \in \mathbb{N}$. Further, let K be the smallest subscript such that $d_K \neq e_K$. Then we have the following:

$$\text{If } d_K < e_K, \text{ then } \bullet\ e_K = d_K + 1,\ \bullet\ e_j = 0 \text{ and } d_j = 9 \text{ for all } j > K.$$

Corollary 9.5. *If a real number has different decimal expansions then it is rational.*

Proposition 9.35 (B/G prop.11.8).
Let $x, y \in \mathbb{R}$ be such that $x < y$. Then there exists $q \in \mathbb{Q}$ be such that $x < q < y$.

Definition 9.18 (Repeating Decimals). A nonnegative decimal

$$x = m.d_1 d_2 \ldots = m + \sum_{j=1}^{\infty} d_j 10^{-j} \quad (d_j \in \{0, 1, 2 \ldots, 9\})$$

is **repeating** if there are natural numbers N and p such that

$$d_{N+n+kp} = d_{N+n} \ \forall\, 0 \leq n < p,\ k \in \mathbb{N}. \ \square$$

Proposition 9.36 (B/G Prop.12.11, p.119). *Every repeating decimal represents a rational number.*

Note 9.2 (Decimal expansions of real numbers)**.** Let $x \in \mathbb{R}$.

(a) x has at most two different decimal expansions.

(b) If x has two expansions then one is all zeros except for finitely many digits and the other is all nines except for finitely many digits.

(c) If x has more than one expansion then x is rational.

(d) x is a repeating decimal if and only if $x \in \mathbb{Q}$. ☐

9.7 Countable and Uncountable Subsets of the Real Numbers

Theorem 9.13. *The real numbers are uncountable.*

Definition 9.19 (algebraic numbers)**.** Let $x \in \mathbb{R}$ be the root (zero) of a polynomial with integer coefficients. We call such x an **algebraic number** and we call any real number that is not algebraic a **transcendental number**. ☐

Proposition 9.37 (B/G Prop.13.21)**.** *The set of all algebraic numbers is countable.*

Proposition 9.38. *Let $k, m, n \in \mathbb{N}$. Then $\sqrt[k]{\frac{m}{n}}$ is algebraic.*

Proposition 9.39. *Let $r \in \mathbb{Q}$. Then r is algebraic.*

Proposition 9.40. *The set of all transcendental numbers and that of all irrational numbers are uncountable.*

9.8 Limit Inferior and Limit Superior

Definition 9.20 (Tail sets of a sequence)**.** Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. Let

$$T_n \;:=\; \{x_j : j \in \mathbb{N} \text{ and } j \geq n\} \;=\; \{x_n, x_{n+1}, x_{n+2}, x_{n+3}, \dots\} \tag{9.36}$$

be what remains in the sequence after we discard the first $n-1$ elements. We call $(T_n)_{n\in\mathbb{N}}$ the n–th **tail set** of the sequence $(x_k)_k$. ☐

Definition 9.21. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ with tail sets $T_n = \{x_j : j \in \mathbb{N}, j \geq n\}$.
Assume that T_n is bounded above for some $n \in \mathbb{N}$ (and hence for all $n \in \mathbb{N}$). We call

$$\limsup_{n\to\infty} x_j \;:=\; \lim_{n\to\infty}\big(\sup_{j\geq n} x_j\big) \;=\; \inf_{n\in\mathbb{N}}\big(\sup_{j\geq n} x_j\big) \;=\; \inf_{n\in\mathbb{N}}\big(\sup(T_n)\big)$$

the **lim sup** or **limit superior** of the sequence (x_n).
If, for each n, T_n is not bounded above then we say $\limsup_{n\to\infty} x_j = \infty$.

Assume that T_n is bounded below for some n (and hence for all $n \in \mathbb{N}$). We call

$$\liminf_{n\to\infty} x_j \;:=\; \lim_{n\to\infty}\big(\inf_{j\geq n} x_j\big) \;=\; \sup_{n\in\mathbb{N}}\big(\inf_{j\geq n} x_j\big) \;=\; \sup_{n\in\mathbb{N}}\big(\inf(T_n)\big)$$

the **lim inf** or **limit inferior** of the sequence (x_n).
If, for each n, T_n is not bounded below then we say $\liminf_{n\to\infty} x_j = -\infty$. $\square$

Theorem 9.14 (Characterization of limsup and liminf)**.** *Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathbb{R}$. Then*

- ***a1.*** *$\limsup_{n\to\infty} x_n$ is the largest of all real numbers x for which $n_1 < n_2 < \cdots \in \mathbb{N}$ can be found such that $x = \lim_{j\to\infty} x_{n_j}$.*
- ***a2.*** *$\limsup_{n\to\infty} x_n$ is the only real number u such that, for all $\varepsilon > 0$, the following is true:*
  $x_n > u + \varepsilon$ for at most finitely many n and $x_n > u - \varepsilon$ for infinitely many n.
- ***b1.*** *$\liminf_{n\to\infty} x_n$ is the smallest of all real numbers x for which $n_1 < n_2 < \cdots \in \mathbb{N}$ can be found such that $x = \lim_{j\to\infty} x_{n_j}$.*
- ***b2.*** *$\liminf_{n\to\infty} x_n$ is the only real number l such that, for all $\varepsilon > 0$, the following is true:*
  $x_n < l - \varepsilon$ for at most finitely many n and $x_n < l + \varepsilon$ for infinitely many n.

Theorem 9.15 (Characterization of limits via limsup and liminf)**.** *Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathbb{R}$.*
The sequence (x_n) converges to a real number if and only if liminf and limsup for that sequence coincide. Moreover, if such is the case then

$$\lim_{n\to\infty} x_n \;=\; \liminf_{n\to\infty} x_n \;=\; \limsup_{n\to\infty} x_n. \tag{9.37}$$

Proposition 9.41. *Let $x_n, x'_n \in \mathbb{R}$ be two sequences of real numbers.*

Assume there is $K \in \mathbb{N}$ such that $x_n \le x'_n$ for all $n \ge K$. Then

$$\liminf_{n\to\infty} x_n \le \liminf_{n\to\infty} x'_n \quad \text{and} \quad \limsup_{n\to\infty} x_n \le \limsup_{n\to\infty} x'_n.$$

Corollary 9.6. *Let $x_n, y_n \in \mathbb{R}$ be two sequences of real numbers.*
Assume there is $K \in \mathbb{N}$ such that $x_n = y_n$ for all $n \ge K$. Then

$$\limsup_{n\to\infty} x_n = \limsup_{n\to\infty} y_n \quad \text{and} \quad \liminf_{n\to\infty} x_n = \liminf_{n\to\infty} y_n.$$

Corollary 9.7. *Let $x_n \ge 0$ such that $\limsup_{n\to\infty} x_n = 0$. Then $(x_n)_n$ converges to zero.*

Proposition 9.42. ⋆ *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ which is bounded above with tail sets T_n.*
***(A)** Let*

$$\begin{aligned} \mathscr{U} &:= \{y \in \mathbb{R} : T_n \cap [y, \infty[\neq \emptyset \text{ for all } n \in \mathbb{N}\}, \\ \mathscr{U}_1 &:= \{y \in \mathbb{R} : \text{ for all } n \in \mathbb{N} \text{ there exists } k \in \mathbb{Z}_{\ge 0} \text{ such that } x_{n+k} \ge y\}, \\ \mathscr{U}_2 &:= \{y \in \mathbb{R} : \exists \text{ subsequence } n_1 < n_2 < n_3 < \dots \in \mathbb{N} \text{ such that } x_{n_j} \ge y \text{ for all } j \in \mathbb{N}\}, \\ \mathscr{U}_3 &:= \{y \in \mathbb{R} : x_n \ge y \text{ for infinitely many } n \in \mathbb{N}\}. \end{aligned} \tag{9.38}$$

Then $\mathscr{U} = \mathscr{U}_1 = \mathscr{U}_2 = \mathscr{U}_3$.
***(B)** There exists $z = z(\mathscr{U}) \in \mathbb{R}$ such that $\mathscr{U}$ is either an interval $]-\infty, z]$ or an interval $]-\infty, z[$.*
***(C)** Let $u := \sup(\mathscr{U})$. Then $u = z = z(\mathscr{U})$ as defined in part B. Further, u is the only real number such that*

C1. $$u - \varepsilon \in \mathscr{U} \quad \text{and} \quad u + \varepsilon \notin \mathscr{U} \quad \text{for all } \varepsilon > 0. \tag{9.39}$$

C2. *There exists a subsequence $(n_j)_{j\in\mathbb{N}}$ of integers such that $u = \lim_{j\to\infty} x_{n_j}$ and u is the largest real number for which such a subsequence exists.*

Corollary 9.8. ⋆ *As in prop.9.42, let $u := \sup(\mathscr{U})$. Then $\mathscr{U} =]-\infty, u]$ or $\mathscr{U} =]-\infty, u[$. Further, u is determined by the following property: For any $\varepsilon > 0$, $x_n > u - \varepsilon$ for infinitely many n and $x_n > u + \varepsilon$ for at most finitely many n.*

Proposition 9.43. ⋆ *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ with tail sets T_n which is bounded below.*
(A) *Let*

$$\begin{aligned}\mathscr{L} &:= \{y \in \mathbb{R} : T_n \cap]-\infty, y] \neq \emptyset \text{ for all } n \in \mathbb{N}\},\\ \mathscr{L}_1 &:= \{y \in \mathbb{R} : \text{ for all } n \in \mathbb{N} \text{ there exists } k \in \mathbb{Z}_{\geq 0} \text{ such that } x_{n+k} \leq y\},\\ \mathscr{L}_2 &:= \{y \in \mathbb{R} : \exists \text{ subsequence } n_1 < n_2 < n_3 < \dots \in \mathbb{N} \text{ such that } x_{n_j} \leq y \text{ for all } j \in \mathbb{N}\},\\ \mathscr{L}_3 &:= \{y \in \mathbb{R} : x_n \leq y \text{ for infinitely many } n \in \mathbb{N}\}.\end{aligned} \tag{9.40}$$

Then $\mathscr{L} = \mathscr{L}_1 = \mathscr{L}_2 = \mathscr{L}_3$.
(B) *There exists $z = z(\mathscr{L}) \in \mathbb{R}$ such that $\mathscr{L}$ is either an interval $[z, \infty[$ or an interval $]z, \infty[$.*
(C) *Let $l := \inf(\mathscr{L})$. Then $l = z = z(\mathscr{L})$ as defined in part B. Further, l is the only real number such that*

C1.
$$l + \varepsilon \in \mathscr{L} \quad \text{and} \quad l - \varepsilon \notin \mathscr{L} \tag{9.41}$$

C2. *There exists a subsequence $(n_j)_{j\in\mathbb{N}}$ of integers such that $l = \lim\limits_{j\to\infty} x_{n_j}$ and l is the smallest real number for which such a subsequence exists.*

Proposition 9.44. ⋆

Let (x_n) be a bounded sequence of real numbers.
As in prop. 9.42 and prop 9.43, let

$$\begin{aligned} u &= \sup(\mathscr{U}) = \sup\{y \in \mathbb{R} : T_n \cap [y, \infty[\neq \emptyset \text{ for all } n \in \mathbb{N}\},\\ l &= \inf(\mathscr{L}) = \inf\{y \in \mathbb{R} : T_n \cap]-\infty, y] \neq \emptyset \text{ for all } n \in \mathbb{N}\},\end{aligned} \tag{9.42}$$

Then $u = \limsup\limits_{n\to\infty} x_j$ *and* $l = \liminf\limits_{n\to\infty} x_j$.

9.9 Sequences of Sets and Indicator functions and their liminf and limsup ⋆

Definition 9.22 (limsup and liminf of a sequence of real–valued functions). ⋆
Let Ω be a nonempty set and let $f_n : \Omega \to \mathbb{R}$ be a sequence of real-valued functions such that $f_n(\omega)$ is bounded for all $\omega \in \Omega$. We define

$$\liminf_{n\to\infty} f_n : \Omega \to \mathbb{R} \quad \text{as follows: } \omega \mapsto \liminf_{n\to\infty} f_n(\omega), \tag{9.43}$$

$$\limsup_{n\to\infty} f_n : \Omega \to \mathbb{R} \quad \text{as follows: } \omega \mapsto \limsup_{n\to\infty} f_n(\omega). \ \square \tag{9.44}$$

Proposition 9.45 (liminf and limsup of $\{0,1\}$–functions). *Let $\Omega \neq \emptyset$ and $f_n : \Omega \to \{0,1\}$. Let $\omega \in \Omega$. Then both $\liminf_n f_n(\omega)$ and $\limsup_n f_n(\omega)$ can only be equal to zero or one. Further,*

$$\liminf_{n\to\infty} f_n(\omega) = 1 \Leftrightarrow f_n(\omega) = 1 \text{ eventually}, \tag{9.45}$$

$$\limsup_{n\to\infty} f_n(\omega) = 1 \Leftrightarrow f_n(\omega) = 1 \text{ for infinitely many } n \in \mathbb{N}. \tag{9.46}$$

Definition 9.23. ⋆ Let $A_n \subseteq \Omega\,(n \in \mathbb{N})$. We define

$$A_\star := \bigcup_{n\in\mathbb{N}} \bigcap_{j\geq n} A_j, \qquad A^\star := \bigcap_{n\in\mathbb{N}} \bigcup_{j\geq n} A_j. \; \square \tag{9.47}$$

Proposition 9.46. *Let $\omega \in \Omega$. Then*

$$\omega \in A_\star \Leftrightarrow \omega \in A_n \text{ eventually, i.e., } \omega \in A_n \text{ for all except at most finitely many } n \in \mathbb{N}. \tag{9.48}$$

$$\omega \in A^\star \Leftrightarrow \omega \in A_n \text{ for infinitely many } n \in \mathbb{N}, \tag{9.49}$$

Proposition 9.47 (liminf and limsup of indicator functions). *Let $A_n \subseteq \Omega\,(n \in \mathbb{N})$ and let $A_\star, A^\star$ be the sets defined in (9.47). Then*

$$\mathbf{1}_{A_\star} = \liminf_{n\to\infty} \mathbf{1}_{A_n} \quad \text{and} \quad \mathbf{1}_{A^\star} = \limsup_{n\to\infty} \mathbf{1}_{A_n} \tag{9.50}$$

Definition 9.24 (limsup and liminf of a sequence of sets). ⋆
Let Ω be a nonempty set and let $A_n \subseteq \Omega$ $(n \in \mathbb{N})$. We define

$$\liminf_{n\to\infty} A_n := \bigcup_{n\in\mathbb{N}} \bigcap_{j\geq n} A_j, \tag{9.51}$$

$$\limsup_{n\to\infty} A_n := \bigcap_{n\in\mathbb{N}} \bigcup_{j\geq n} A_j. \tag{9.52}$$

We call $\liminf\limits_{n\to\infty} A_n$ the **limit inferior** and $\limsup\limits_{n\to\infty} A_n$ the **limit superior** of the sequence A_n.

We note that $\liminf\limits_{n\to\infty} A_n = \limsup\limits_{n\to\infty} A_n$ if and only if the functions $\liminf\limits_{n\to\infty} \mathbf{1}_{A_n}$ and $\limsup\limits_{n\to\infty} \mathbf{1}_{A_n}$ coincide (prop. 9.47) which is true if and only if the sequence $\mathbf{1}_{A_n}(\omega)$ has a limit for all $\omega \in \Omega$ (thm.9.15 on p.96). In this case we define

$$\lim_{n\to\infty} A_n := \liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n \tag{9.53}$$

and we call this set the **limit** of the sequence A_n. $\square$

Note 9.3 (Notation for limits of monotone sequences of sets)**.**
Let (A_n) be a nondecreasing sequence of sets, i.e., $A_1 \subseteq A_2 \subseteq \dots$ and let $A := \bigcup_n A_n$.
Further, let B_n be a nonincreasing sequence of sets, i.e., $B_1 \supseteq B_2 \supseteq \dots$ and let $B := \bigcap_n B_n$.
We write suggestively

$$A_n \uparrow A \quad (n \to \infty)\,, \qquad B_n \downarrow B \quad (n \to \infty)\,. \ \square$$

9.10 Sequences that Enumerate Parts of $\mathbb{Q}$ ⋆

Theorem 9.16 (Universal sequence of rational numbers with convergent subsequences to any real number)**.** ⋆ *There is a sequence $(q_n)_{n\in\mathbb{N}}$ of fractions which satisfies the following:*
For any $x \in \mathbb{R}$ there is a sequence $n_1, n_2, n_3, \dots,$ of natural numbers such that $x = \lim_{k\to\infty} q_{n_k}$ ■.

10 Cardinality II: Comparing Uncountable Sets

10.1 The Cardinality of a Set

Definition 10.1 (Cardinality Comparisons)**.** Given are two arbitrary sets X and Y. We say that

(a) X, Y **have same cardinality**, and we write $\mathbf{card}(X) = \mathbf{card}(Y)$, if either both $X, Y \neq \emptyset$ and there is a bijection $f : X \xrightarrow{\sim} Y$, or if both X and Y are empty. Otherwise we write $\mathbf{card}(X) \neq \mathbf{card}(Y)$

(b) the **cardinality of** X **is less than or equal to the cardinality of** Y, and we write $\mathbf{card}(X) \leq \mathbf{card}(Y)$, if there is an injective mapping $f : X \to Y$ or if X is empty.

(c) the **cardinality of** X **is less than the cardinality of** Y, and we write $\mathbf{card}(X) < \mathbf{card}(Y)$, if both $\text{card}(X) \leq \text{card}(Y)$ and $\text{card}(Y) \neq \text{card}(X)$, i.e., if either $X = \emptyset$ and $Y \neq \emptyset$, or there is an injective mapping but not a bijection $f : X \to Y$.

(d) the **cardinality of** X **is greater than or equal to the cardinality of** Y, and we write $\mathbf{card}(X) \geq \mathbf{card}(Y)$, if $\text{card}(Y) \leq \text{card}(X)$.

(e) the **cardinality of** X **is greater than the cardinality of** Y, and we write $\mathbf{card}(X) > \mathbf{card}(Y)$, if $\text{card}(Y) < \text{card}(X)$. ☐

Example 10.1. Let A, B be two sets such that $A \subseteq B$. Then $\text{card}(A) \leq \text{card}(B)$.

Theorem 10.1 (B/G thm.13.31)**.** *Let X be a set. Then $\mathbf{card}(X) < \mathbf{card}(2^X)$.*
In other words, X can be injected into 2^X, but it is not possible to find bijective $f : X \xrightarrow{\sim} 2^X$.

Proposition 10.1. *Let X, Y be two sets such that $card(X) = card(Y)$. Then $card(2^X) = card(2^Y)$.*

10.2 Cardinality as a Partial Ordering

Definition 10.2 (Cardinality as an Equivalence Class)**.** ⋆ Let $X, Y \subseteq \Omega$.
We call X and Y equivalent and we write $X \sim Y$, if and only if $\mathbf{card}(X) = \mathbf{card}(Y)$, i.e., either both X and Y are empty, or both are not empty and there is a bijection $f : X \xrightarrow{\sim} Y$.
The proposition following this definition shows that "$\sim$" is indeed an equivalence relation on 2^Ω. This justifies to define for a set $X \subseteq \Omega$ its **cardinality** as follows:

$$\mathbf{card}(X) := [X] \quad \text{(the equivalence class of } X \text{ w.r.t "}\sim\text{")}. \tag{10.1}$$

In other words,

$$\mathbf{card}(\emptyset) := \{\emptyset\}, \tag{10.2}$$
$$\mathbf{card}(X) := \{Y \subseteq \Omega : \exists \text{ bijection } X \to Y\} \text{ if } X \neq \emptyset. \tag{10.3}$$
☐

Proposition 10.2. *$X \sim Y$ as defined above is an equivalence relation on 2^Ω.*

Proposition 10.3. *Let X', X'', Y', Y'' be nonempty sets such that $X' \cap X'' = \emptyset$ and $Y' \cap Y'' = \emptyset$. Let $f' : X' \to Y'$ and $f'' : X'' \to Y''$. Then the function*

$$f : X' \biguplus X'' \longrightarrow Y' \biguplus Y'' ; \qquad x \mapsto \begin{cases} f'(x) & \text{if } x \in X', \\ f''(x) & \text{if } x \in X'', \end{cases}$$

satisfies the following:

(a) *If f' and f'' are injective then f is injective.*
(b) *If f' and f'' are surjective then f is surjective.*
(c) *If f' and f'' are bijective then f is bijective.*

Theorem 10.2 (Tarski's Fixed Point Theorem)**.**
Let Ω be a set and let $\varphi : 2^\Omega \longrightarrow 2^\Omega$ be nondecreasing with respect to "$\subseteq$", i.e.,

$$A, B \subseteq \Omega \text{ and } A \subseteq B \quad \Rightarrow \quad \varphi(A) \subseteq \varphi(B) .$$

*Then φ has a **fixed point**, i.e., there exists an argument $A_0 \in 2^\Omega$ such that $\varphi(A_0) = A_0$.*

Theorem 10.3 (Cantor–Schröder–Bernstein's Theorem)**.**
Let X and Y be nonempty sets. Let there be injective functions

$$f : X \rightarrowtail Y \qquad \text{and} \qquad g : Y \rightarrowtail X .$$

Then there exists a bijection $X \xrightarrow{\sim} Y$.

Corollary 10.1.

$$\text{The relation } \textbf{card}(X) \leq \textbf{card}(Y) \text{ partially orders the set } \mathscr{A} := \{\textbf{card}(X) : X \subseteq \Omega\} .$$

Theorem 10.4. *Let $X, Y \subseteq \Omega$. Then*

$$\textbf{card}(X) \leq \textbf{card}(Y) \qquad \text{or} \qquad \textbf{card}(Y) \leq \textbf{card}(X)$$

In other words, "$\leq$" is a total ordering [15] *on the set of all cardinalities for subsets of Ω.*

[15] See Definition 5.5 (Linear orderings) on p.38.

Theorem 10.5. *Let* $a, b \in \mathbb{R}$ *such that* $a < b$. *Let* A *be one of* $]a, b[,\ ]a, b\,],\ [\,a, b[,\ [\,a, b\,]$.

$$\textit{Then} \quad \textit{card}(A) \ = \ \textit{card}(\mathbb{R})\,.$$

Theorem 10.6.

$$\textit{card}(\mathbb{R}) \ = \ \textit{card}(2^{\mathbb{N}})\,. \tag{10.4}$$

11 Vectors and Vector spaces

11.1 $\mathbb{R}^n$: Euclidean Space

11.1.1 n–Dimensional Vectors

Definition 11.1 (n–dimensional vectors). ⋆

Let $n \in \mathbb{N}$. An n–**dimensional vector** is a finite, ordered collection $\vec{v} = (x_1, x_2, \dots, x_n)$ of real numbers $x_1, x_2, \dots, x_n$, n is called the **dimension** of the vector $\vec{v}$. □

Definition 11.2 (Transposed matrix). ⋆

Let A be a matrix with m rows and n columns. We will write $A = \big((a_{ij})\big)$ to express that a_{ij} denotes the "cell" at the intersection of row i and column j. ($i \in [1,m]_{\mathbb{Z}}$ and $j \in [1,n]_{\mathbb{Z}}$).

$$A = \begin{bmatrix} a_{11}, & a_{12}, & \dots, & a_{1n} \\ a_{21}, & a_{22}, & \dots, & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}, & a_{m2}, & \dots, & a_{mn}(t) \end{bmatrix}.$$

If A is a matrix with m rows and n columns, and if a_{ij} denotes the "cell" at the intersection of row i and column j, then we denote by $A^\top$ the "flipped" matrix which has row i of A as its i–th column, and column j of A as its j–th row.

In other words, if $A = \big((a_{ij})\big)$ and if $A^\top = \big((a^*_{k\ell})\big)$ then $a^*_{ij} = a_{ji}$ for all $i \in [1,m]_{\mathbb{Z}}$ and $j \in [1,n]_{\mathbb{Z}}$. We call $A^\top$ the **transpose** or **transposed matrix** of A. □

$$A^\top = \begin{bmatrix} a_{11}, & a_{21}, & \dots, & a_{m1} \\ a_{12}, & a_{22}, & \dots, & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n}, & a_{2n}, & \dots, & a_{mn}(t) \end{bmatrix}.$$

11.1.2 Addition and Scalar Multiplication for n–Dimensional Vectors

Definition 11.3 (Addition and scalar multiplication in $\mathbb{R}^n$). ⋆

Given are two n–dimensional vectors
$\vec{x} = (x_1, x_2, \dots, x_n)$ and $\vec{y} = (y_1, y_2, \dots, y_n)$ and a real number α.
We define the **sum** $\vec{x} + \vec{y}$ of $\vec{x}$ and $\vec{y}$ as the vector $\vec{z}$ with the components

$$z_1 = x_1 + y_1; \quad z_2 = x_2 + y_2; \quad \dots; \quad z_n = x_n + y_n; \tag{11.1}$$

We define the **scalar product** $\alpha\vec{x}$ of α and $\vec{x}$ as the vector $\vec{w}$ with the components

$$w_1 = \alpha x_1; \quad w_2 = \alpha x_2; \quad \dots; \quad w_n = \alpha x_n. \tag{11.2}$$ □

11.1.3 Length of n–Dimensional Vectors and the Euclidean Norm

It is customary to write $\|\vec{v}\|_2$ for the length, often also called the **Euclidean norm**, of the vector $\vec{v}$.

Definition 11.4 (Euclidean norm)**.** Let $n \in \mathbb{N}$ and $\vec{v} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ be an n–dimension vector. The **Euclidean norm** $\|\vec{v}\|_2$ of $\vec{v}$ is defined as follows:

$$\|\vec{v}\|_2 \quad = \quad \sqrt{{x_1}^2 + {x_2}^2 + \ldots + {x_n}^2} \quad = \quad \sqrt{\sum_{j=1}^{n} {x_j}^2}. \quad \square \tag{11.3}$$

Proposition 11.1 (Properties of the Euclidean norm)**.**
Let $n \in \mathbb{N}$. Then the Euclidean norm has the following properties, when viewed as a function

$$\|\cdot\|_2 : \mathbb{R}^n \to \mathbb{R}; \qquad \vec{v} = (x_1, x_2, \dots, x_n) \mapsto \|\vec{v}\|_2 = \sqrt{\sum_{j=1}^{n} {x_j}^2} :$$

(11.4a) $\|\vec{v}\|_2 \geq 0 \quad \forall \vec{v} \in \mathbb{R}^n$ *and* $\|\vec{v}\|_2 = 0 \;\Leftrightarrow\; \vec{v} = 0$ ***(positive definiteness)***
(11.4b) $\|\alpha\vec{v}\|_2 = |\alpha| \cdot \|\vec{v}\|_2 \quad \forall \vec{v} \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}$ ***(absolute homogeneity)***
(11.4c) $\|\vec{v} + \vec{w}\|_2 \leq \|\vec{v}\|_2 + \|\vec{w}\|_2 \quad \forall\, \vec{v}, \vec{w} \in \mathbb{R}^n$ ***(triangle inequality)***

11.2 General Vector Spaces

11.2.1 Vector spaces: Definition and Examples

Definition 11.5 (Vector spaces (linear spaces))**.** ⋆ A nonempty set V is called a **vector space** or **linear space** and we call its elements **vectors** if V satisfies the following:
(A) There exists a binary operation $+ : V \times V \to V;\ (x, y) \mapsto x + y$ on V such that $(V, +)$ is an abelian group (see def. 3.2 on p.20). We call $x + y$ the **sum** of x and y. Note that $(V, +)$ being an abelian group means that the following properties hold for "+":

1. $x + y \;=\; y + x$ for all $x, y \in V$ (**commutativity**);
2. $(x + y) + z \;=\; x + (y + z)$ for all $x, y, z \in V$ (**associativity**);
3. There exists an element $0 \in V$, called the **zero element**, or **zero vector**, or **null vector**, with the property that $x + 0 = x$ for each $x \in V$;
4. For every $x \in V$, there exists an element $-x \in V$, called the **negative** of x, with the property that $x + (-x) = 0$ for each $x \in V$. When adding negatives, then there is a convenient short form. We write $x - y$ as an abbreviation for $x + (-y)$;

(B) There exists a function $\cdot : \mathbb{R} \times V \to V;\ (\alpha, x) \mapsto \alpha \cdot x$, i.e., any real number α and vector x uniquely determine a vector $\alpha \cdot x$. It is customary to simply write αx for $\alpha \cdot x$. This vector is called the **scalar product** of α and x, and it has the following properties:

1. $\alpha(\beta x) = (\alpha\beta)x$;
2. $1x = x$;

(C) The operations of addition and scalar multiplication obey the two **distributive laws**

1. $(\alpha + \beta)x = \alpha x + \beta x$;
2. $\alpha(x + y) = \alpha x + \alpha y$; □

Remark 11.1. ★ A vector space V is an algebraic structure with the following properties:

(a) V is nonempty and comes with two assignments:
$+ : V \times V \to V;\ (x, y) \mapsto x + y$, the sum of x and y,
$\cdot : \mathbb{R} \times V \to V;\ (\alpha, x) \mapsto \alpha \cdot x$, (also written αx), the scalar product of α and x.

(c) $(V, +)$ is an abelian group. We write 0 (null vector) for its neutral element, $-x$ for the inverse of a vector x, and $x - y$ for $x + (-y)$.

(d) $\alpha(\beta x) = (\alpha\beta)x$ for all $\alpha, \beta \in \mathbb{R}$ and $x \in V$.

(e) $1 \cdot x = x$ for all $x \in V$. (1 is the real number 1).

(f) Two distributive laws:
$(\alpha + \beta)x = \alpha x + \beta x$,
$\alpha(x + y) = \alpha x + \alpha y$. □

Definition 11.6 (Subspaces of vector spaces)**.** Let V be a vector space and let $A \subseteq V$ be a nonempty subset of V such that

- For any $x, y \in A$ and $\alpha \in \mathbb{R}$ the sum $x + y$ and the scalar product αx also belong to A.

Then A is called a **subspace** of V.
The set $\{0\}$ which only contains the null vector 0 of V is called the **nullspace**. □

Proposition 11.2 (Subspaces are vector spaces)**.** *A subspace of a vector space is a vector space, i.e., it satisfies all requirements of definition (11.5).*

A subspace is a subset of a vector space which is closed with respect to vector addition and scalar multiplication.

The following example should be thought of as the **definition** of the very important function spaces $\mathscr{F}(X,\mathbb{R})$, $\mathscr{B}(X,\mathbb{R})$, $\mathscr{C}(X,\mathbb{R})$.

Example 11.1 (Vector spaces of real–valued functions)**.**

$$\begin{aligned}\mathscr{F}(X,\mathbb{R}) &= \{f(\cdot) : f(\cdot) \text{ is a real–valued function on } X\}\\ \mathscr{B}(X,\mathbb{R}) &= \{g(\cdot) : g(\cdot) \text{ is a bounded real–valued function on } X\}\\ \mathscr{C}([a,b],\mathbb{R}) &= \{h(\cdot) : h(\cdot) \text{ is a continuous real–valued function for } a \leq x \leq b\}\end{aligned}$$

- We have subspace relationships $\mathscr{B}(X,\mathbb{R}) \subseteq \mathscr{F}(X,\mathbb{R})$
- We have subspace relationships $\mathscr{C}([a,b],\mathbb{R}) \subseteq \mathscr{B}([a,b],\mathbb{R}) \subseteq \mathscr{F}([a,b],\mathbb{R})$ □

Definition 11.7 (linear combinations)**.** ⋆

Let V be a vector space and let $x_1, x_2, x_3, \dots, x_n \in V$ be a finite number of vectors in V.
Let $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n \in \mathbb{R}$. We call the finite sum

$$\sum_{j=0}^{n} \alpha_j x_j = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_n x_n \tag{11.5}$$

a **linear combination** of the vectors x_j . The multipliers $\alpha_1, \alpha_2, \dots$ are called **scalars**. □

Proposition 11.3 (Vector spaces are closed w.r.t. linear combinations)**.** *Let V be a vector space and let $x_1, x_2, x_3, \dots, x_n \in V$ be a finite number of vectors in V. Let $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n \in \mathbb{R}$. Then the linear combination $\sum\limits_{j=0}^{n} \alpha_j x_j$ also belongs to V. Note that this is also true for subspaces, because those are vector spaces, too.*

Proposition 11.4. *Let V be a vector space and let $(W_i)_{i \in I}$ be a family of subspaces of V. Let $W := \bigcap [W_i : i \in I]$. Then W is a subspace of V.*

Definition 11.8 (Linear span)**.** ⋆

Let V be a vector space and $A \subseteq V$. Then the set

$$\mathit{span}(A) := \{ \sum_{j=1}^{k} \alpha_j x_j : k \in \mathbb{N}, \alpha_j \in \mathbb{R}, x_j \in A \ (1 \leq j \leq k) \}. \tag{11.6}$$

of all linear combinations of vectors in A is called the **span** or **linear span** of A. □

Proposition 11.5. *Let V be a vector space and $A \subseteq V$. Then $span(A)$ is a subspace of V.*

Theorem 11.1. *Let V be a vector space and $A \subseteq V$.*

Let $\mathscr{V} := \{W \subseteq V : W \supseteq A$ and W is a subspace of $V\}$. Then $span(A) = \bigcap [W : W \in \mathscr{V}]$.

Remark 11.2 (Linear span(A) = subspace generated by A)**.** Let V be a vector space and $A \subseteq V$. Theorem 11.1 justifies to call $span(A)$ the **subspace generated by** A. □

Definition 11.9 (linear mappings)**.** ⋆

Let V_1, V_2 be two vector spaces. Let the function $f(\cdot) : V_1 \to V_2$ satisfy

(11.7a) $$f(x+y) \;=\; f(x) + f(y) \quad \forall x, y \in V_1$$ **additivity**

(11.7b) $$f(\alpha x) \;=\; \alpha f(x) \quad \forall x \in V_1,\ \forall \alpha \in \mathbb{R}$$ **homogeneity**

Then we call $f(\cdot)$ a **linear function** or **linear mapping**. □

Proposition 11.6 (Linear mappings preserve linear combinations)**.** *Let V_1, V_2 be two vector spaces. Let $f(\cdot) : V_1 \to V_2$ be a linear map and let $x_1, x_2, x_3, \dots, x_n \in V_1$ be a finite number of vectors in the domain V_1 of $f(\cdot)$. Let $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n \in \mathbb{R}$.*
Then $f(\cdot)$ preserves any such linear combination, i.e.,

(11.8) $$f(\sum_{j=0}^{n} \lambda_j x_j) \;=\; \sum_{j=0}^{n} \lambda_j f(x_j).$$

Lemma 11.1 ($F \circ span \;=\; span \circ F$)**.** *Let V, W be two vector spaces and $F : V \to W$ a linear mapping from V to W. Let $A \subseteq V$. Then*

(11.9) $$F(span(A)) \;=\; span(F(A)).$$

Definition 11.10 (Linear dependence and independence)**.** ⋆

Let V be a vector space and $A \subseteq V$

(a) A is called **linearly dependent** if the following is true: There exist distinct vectors $x_1, x_2, \dots x_k \in A$ and scalars $\alpha_1, \alpha_2, \dots \alpha_k \in \mathbb{R}$ ($k \in \mathbb{N}$) such that

- not all scalars α_j are zero $(1 \leq j \leq k)$
- $\sum_{j=1}^{k} \alpha_j x_j = 0.$

(b) A is called **linearly independent** if A is not linearly dependent, i.e., if the following is true: Let $x_1, x_2, \dots x_k \in A$ and $\alpha_1, \alpha_2, \dots \alpha_k \in \mathbb{R}$ ($k \in \mathbb{N}$).

- If $\sum_{j=1}^{k} \alpha_j x_j = 0$ then $\alpha_j = 0$, for all $1 \leq j \leq k$. □

Definition 11.11 (Basis of a vector space). ⋆

Let V be a vector space and $B \subseteq V$. B is called a **basis** of V if both

- B is linearly independent
- $span(B) = V$. □

Definition 11.12 (Standard basis of $\mathbb{R}^n$). ⋆

Let $n \in \mathbb{N}$. For $i \in [1, n]_{\mathbb{Z}}$, let $\vec{e}^{(i)} := \left(\delta_{i1}, \delta_{i2}, \dots, \delta_{in}\right)^\top$.
Here δ_{ij} denotes the Kronecker delta: $\delta_{ii} = 1$ for all i and $\delta_{ij} = 0$ for $i \neq j$. Thus,

$$\vec{e}^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}^{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \vec{e}^{(n)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Let $B := \{\vec{e}^{(i)} : i \in [1, n]_{\mathbb{Z}}\}$. Then B is a basis of $\mathbb{R}^n$ which we call the **standard basis**, also the **canonical basis**, of $\mathbb{R}^n$. □

Lemma 11.2. *Let V be a vector space and $A \subseteq V$.*
Assume that A is linearly independent but not a basis and that $y \in span(A)^{\complement}$.
Then $A \cup \{y\}$ is linearly independent.

Theorem 11.2. *Let V be a vector space with a finite basis $B = \{b_1, \dots, b_k\}$.*
Then any other basis of V has the same size k.

Definition 11.13 (Dimension of vector spaces). ⋆

- Let V be a vector space with a finite basis $B = \{b_1, \dots, b_k\}$. We call k the **dimension** of V and we write $\dim(V) = k$.
- If V does not possess a finite basis then we say that V has infinite dimension and we write $\dim(V) = \infty$. □

Proposition 11.7. *For $a \in \mathbb{R}$ define $f_a(\cdot) \in \mathscr{B}(\mathbb{R}, \mathbb{R})$ as follows.*

$$f_a(x) := \begin{cases} 0 & \text{if } x \neq a, \\ 1 & \text{if } x = a. \end{cases}$$

Then $\mathscr{A} := \{f_a : a \in \mathbb{R}\}$ is a linearly independent subset of $\mathscr{B}(\mathbb{R}, \mathbb{R})$.

Proposition 11.8. *Let V be a vector space and let U be a (linear) subspace of V. Let $x_0 \in V$. Let $\tilde{U} := \{u + \lambda x_0 : u \in U \text{ and } \lambda \in \mathbb{R}\}$. Then $\tilde{U} = \mathit{span}(U \cup \{x_0\})$.*

Proposition 11.9. *Let V and V' be two vector spaces and let U be a proper (linear) subspace of V, i.e., $U \subsetneq V$. Let $x_0 \in U^{\complement}$, $y_0 \in V'$. Let $f := U \to V'$ be a linear function from U into V'. Let $\alpha \in \mathbb{R}$. Then*

$$g : U \uplus \{x_0\} \to V'; \quad g(x) := \begin{cases} f(x) & \text{if } x \in U, \\ y_0 & \text{if } x = x_0, \end{cases} \tag{11.10}$$

uniquely extends to a linear function $\tilde{f} : \mathit{span}(U \uplus \{x_0\}) \to V'$ as follows:

$$\tilde{f}(x + \alpha\, x_0) := f(x) + \alpha\, y_0 \quad \text{for } x \in U, \alpha \in \mathbb{R}\,. \tag{11.11}$$

11.2.2 Normed Vector Spaces

Definition 11.14 (Inner product). Let V be a vector space with a function

$$\bullet(\cdot,\cdot) : V \times V \to \mathbb{R}; \qquad (x, y) \mapsto x \bullet y := \bullet(x, y)$$

which satisfies the following:

(11.12a) $x \bullet x \geq 0 \quad \forall x \in V$ and $x \bullet x = 0 \;\Leftrightarrow\; x = 0$ **positive definiteness**

(11.12b) $x \bullet y = y \bullet x \quad \forall x, y \in V$ **symmetry**

(11.12c) $(x + y) \bullet z = x \bullet z + y \bullet z \quad \forall\, x, y, z \in V$ **additivity**

(11.12d) $(\lambda x) \bullet y = \lambda(x \bullet y) \quad \forall\, x, y \in V \quad \forall\, \lambda \in \mathbb{R}$ **homogeneity**

We call such a function an **inner product**. □

Definition 11.15 (Bilinearity). ⋆
Let V be a vector space with a function

$$B : V \times V \to \mathbb{R}; \qquad (x, y) \mapsto B(x, y).$$

$B(\cdot, \cdot)$ is called **bilinear** if it is linear in each argument, i.e., the mappings

$$B_1 : V \to \mathbb{R}; \quad x \mapsto B(x, y)$$
$$B_2 : V \to \mathbb{R}; \quad y \mapsto B(x, y)$$

are both linear. □

Proposition 11.10 (Algebraic properties of the inner product).
Let V be a vector space with inner product $\bullet(\cdot, \cdot)$. Let $a, b, x, y \in V$. Then

(11.13a) $$(a + b) \bullet (x + y) = a \bullet x \; + \; b \bullet x \; + \; a \bullet y \; + \; b \bullet y$$

(11.13b) $$(x + y) \bullet (x + y) = x \bullet x \; + \; 2(x \bullet y) \; + \; y \bullet y$$

(11.13c) $$(x - y) \bullet (x - y) = x \bullet x \; - \; 2(x \bullet y) \; + \; y \bullet y$$

Proposition 11.11 (Inner product on $\mathbb{R}^n$)). *Let $n \in \mathbb{N}$. Then the real–valued function*

(11.14) $$(\vec{x}, \vec{y}) \; \mapsto \; x_1 y_1 + x_2 y_2 + \ldots + x_n y_n \; = \; \sum_{j=1}^{n} x_j y_j \, ,$$

where $\vec{x} = (x_1, \ldots x_n)$ and $\vec{y} = (y_1, \ldots y_n)$, is an inner product on $\mathbb{R}^n \times \mathbb{R}^n$.

Proposition 11.12 (Cauchy–Schwartz inequality for inner products).

Let V be a vector space with an inner product

$$\bullet(\cdot, \cdot) : V \times V \to \mathbb{R}; \qquad (x, y) \mapsto x \bullet y := \bullet(x, y)$$

Then,

$$(x \bullet y)^2 \; \leq \; (x \bullet x) \, (y \bullet y).$$

Definition 11.16 (sup–norm of bounded real–valued functions). Let X be an arbitrary, nonempty set. Let $f : X \to \mathbb{R}$ be a bounded real–valued function on X, i.e., there exists $K \geq 0$ such that $|f(x)| \leq K$ for all $x \in X$. Let

(11.15) $$\|f\|_\infty \; := \; \sup\{|f(x)| : x \in X\}$$

We call $\|f\|_\infty$ the **supremum norm** or **sup–norm** of the function f. □

Proposition 11.13 (Properties of the sup norm)**.** *Let X be an arbitrary, nonempty set. Let*

$$\mathscr{B}(X,\mathbb{R}) := \{h(\cdot) : h(\cdot) \text{ is a bounded real–valued function on } X\}$$

(see example 11.1 on p. 107). Then the sup–norm

$$\|\cdot\|_\infty : \mathscr{B}(X,\mathbb{R}) \to \mathbb{R}_+, \qquad h \mapsto \|h\|_\infty = \sup\{|h(x)| : x \in X\}$$

satisfies the following:

(11.16a) $\|f\|_\infty \geq 0 \ \forall\, f \in \mathscr{B}(X,\mathbb{R})$ *and* $\|f\|_\infty = 0 \ \Leftrightarrow \ f(\cdot) = 0$ ***positive definiteness***

(11.16b) $\|\alpha f(\cdot)\|_\infty = |\alpha| \cdot \|f(\cdot)\|_\infty \ \forall\, f \in \mathscr{B}(X,\mathbb{R}), \forall\, \alpha \in \mathbb{R}$ ***absolute homogeneity***

(11.16c) $\|f(\cdot) + g(\cdot)\|_\infty \leq \|f(\cdot)\|_\infty + \|g(\cdot)\|_\infty \ \forall\, f, g \in \mathscr{B}(X,\mathbb{R})$ ***triangle inequality***

Definition 11.17 (Normed vector spaces)**.** Let V be a vector space with a real–valued function

$$\|\cdot\| : \ V \to \mathbb{R} \qquad x \mapsto \|x\|$$

which satisfies

(11.17a) $\|x\| \geq 0 \quad \forall x \in V$ and $\|x\| = 0 \ \Leftrightarrow \ x = 0$ **positive definiteness**

(11.17b) $\|\alpha x\| = |\alpha| \cdot \|x\| \quad \forall x \in V, \forall \alpha \in \mathbb{R}$ **absolute homogeneity**

(11.17c) $\|x + y\| \leq \|x\| + \|y\| \quad \forall\, x, y \in V$ **triangle inequality**

We call $\|\cdot\|$ a **norm** on V and we call V a **normed vector space**.
We write $(V, \|\cdot\|)$ instead of V when we wish to emphasize what norm on V we are discussing. □

Definition 11.18 (p–norms for $\mathbb{R}^n$)**.** ⋆
Let $p \geq 1$. It will be proved in prop.11.16 on p.114 that the function

(11.18)
$$\vec{x} \mapsto \|\vec{x}\|_p := \Big(\sum_{j=1}^{n} |x_j|^p\Big)^{1/p}$$

is a norm on $\mathbb{R}^n$). This norm is called the **p–norm** on $\mathbb{R}^n$). The Euclidean norm is a p–norm; it is the 2-norm on $\mathbb{R}^n$). □

Theorem 11.3 (Inner products define norms)**.**
Let V be a vector space with an inner product

$$\bullet(\cdot,\cdot) : V \times V \to \mathbb{R}; \qquad (x,y) \mapsto x \bullet y$$

Then

$$\|\cdot\|_\bullet : x \ \mapsto \ \|x\| \ = \ \sqrt{(x \bullet x)}$$

defines a norm on V

Definition 11.19 (Norm for an inner product)**.** Let V be a vector space with an inner product

$$\bullet(\cdot,\cdot) : V \times V \to \mathbb{R}; \qquad (x,y) \mapsto x \bullet y$$

Then

(11.19) $$\|\cdot\|_\bullet : x \;\mapsto\; \|x\|_\bullet \;:=\; \sqrt{(x \bullet x)}$$

is called the **norm associated with the inner product** $\bullet(\cdot,\cdot)$. ☐

Corollary 11.1. *The Euclidean norm in* $\mathbb{R}^n$*:*

$$\|(x_1, x_2, \dots, x_n)\|_2 \;=\; \sqrt{\sum_{j=1}^{n} {x_j}^2}$$ *(see def.11.4 on p.105) is a norm.*

Definition 11.20. ★ Let $a, b \in \mathbb{R}$, $a < b$ and assume that $f, g : [a,b] \to \mathbb{R}$ are integrable functions. (See example **??** on p.**??**.)

- **(a)** We call the definite integral $\int_a^b f(x)dx$ the **net area** between the graph of f, the x–axis, and the vertical lines through $(a,0)$ $(y = a)$ and $(b,0)$ $(y = b)$. The above integral treats areas above the x–axis as positive and below the x–axis as negative, i.e., the net area is the difference between the areas above the x–axis and those below the x–axis.
- **(b)** We call $\int_a^b |f(x)|dx$ the **area** between the graph of f, the x–axis, and the vertical lines $y = a$ and $y = b$. Note that $f(x)$ has been replaced by its absolute value $|f(x)|$. In contrast to the net area, areas below the x–axis are also counted positive. ☐
- **(c)** We call $\int_a^b f(x) - g(x)dx$ the **net area** between the graphs of f and g and the vertical lines $y = a$ and $y = b$. We call $\int_a^b |f(x) - g(x)|dx$ the **area** between the graphs of f and g and the vertical lines $y = a$ and $y = b$. ☐

Proposition 11.14. *Let $a, b \in \mathbb{R}$ such that $a < b$. and let $f : [a,b] \to \mathbb{R}$ be continuous. Then*

$$\int_a^b f(x)dx \;=\; 0 \quad \text{if and only if} \quad f(x) \;=\; 0 \;\text{ for all } x \in \,]a,b[. \;\square$$

Proposition 11.15. *Let $a, b \in \mathbb{R}$ such that $a < b$. Then the mapping*

(11.20) $$(f,g) \;\mapsto\; f \bullet g \;:=\; \int_a^b f(x)g(x)dx$$

defines an inner product on $f \in \mathscr{C}([a,b],\mathbb{R})$. ☐

Definition 11.21 (L_2–Norm for continuous functions)**.** Let $a, b \in \mathbb{R}$ such that $a < b$. Let $f \bullet g$ be the the following inner product on the space $\mathscr{C}([a,b],\mathbb{R})$ of all continuous functions $[a,b] \to \mathbb{R}$:

$$f \bullet g \ := \ \int_a^b f(x)g(x)dx. \tag{11.21}$$

The L^2**–norm**. of f is the norm associated with that inner product:

$$\|\cdot\|_{L^2} : \ f \mapsto \|f\|_\bullet \ = \ \sqrt{\int_a^b f^2(x)dx}\,. \tag{11.22}$$

□

Definition 11.22 (L^p–norms for $\mathscr{C}([a,b],\mathbb{R})$)**.** ⋆ Let $a, b \in \mathbb{R}$ such that $a < b$ and $p \geq 1$. It will be shown in prop.11.17 (The L^p–norm is a norm) on p.114 that

$$f \mapsto \|f\|_{L^p} \ := \ \left(\int_a^b |f(x)|^p dx\right)^{1/p} \tag{11.23}$$

is a norm on $\mathscr{C}([a,b],\mathbb{R})$. This norm is called the L^p**–norm** of f. □

11.2.3 The Inequalities of Young, Hoelder, and Minkowski ⋆

Proposition 11.16 (The p–norm in $\mathbb{R}^n$ is a norm)**.** *Let $p \in [1, \infty[$.*
Then the p-norm $\vec{x} \mapsto \|\vec{x}\|_p \ = \ \left(\sum_{j=1}^n |x_j|^p\right)^{1/p}$ *is a norm in $\mathbb{R}^n$.*

Proposition 11.17 (The L^p–norm is a norm)**.** *Let $p \in [1, \infty[$ and let $a, b \in \mathbb{R}$ such that $a < b$.*
Then the L^p-norm $f \mapsto \|f\|_{L^p} \ = \ \left(\int_a^b |f(x)|^p\right)^{1/p}$ *is a norm in $\mathscr{C}([a,b],\mathbb{R})$.*

Proposition 11.18 (Young's Inequality)**.** *Let $a, b > 0$ and let $p, q > 1$ be* ***conjugate indices****, i.e.,*

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{11.24}$$

Then ***Young's inequality*** *holds:*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{11.25}$$

Theorem 11.4 (Hoelder's inequality for L^p–norms)**.**
Let $a, b \in \mathbb{R}$ such that $a < b$. Let $p, q > 1$ be conjugate indices, i.e.,

(11.26)
$$\frac{1}{p} + \frac{1}{q} = 1.$$

*Then **Hoelder's inequality** is true:*

(11.27)
$$\|fg\|_{L^1} \leq \|f\|_{L^p}\|g\|_{L^q}, \quad i.e., \int_a^b |f(x)g(x)|dx \leq \Big(\int_a^b |f(x)|^p dx\Big)^{1/p} \Big(\int_a^b |g(x)|^q dx\Big)^{1/q}.$$

Theorem 11.5 (Minkowski's inequality for L^p–norms)**.** *Let $a, b \in \mathbb{R}$ such that $a < b$ and let $p \in [1, \infty[$. Then **Minkowski's inequality** is true:*

(11.28)
$$\|f+g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}, \quad i.e.,$$

(11.29)
$$\Big(\int_a^b |f(x)+g(x)|^p \, dx\Big)^{1/p} \leq \Big(\int_a^b |f(x)|^p \, dx\Big)^{1/p} + \Big(\int_a^b |g(x)|^p \, dx\Big)^{1/p}.$$

Theorem 11.6 (Hoelder's inequality for the p–norms)**.** *Let $n \in \mathbb{N}$ and $\vec{x} = (x_1, \dots x_N), \vec{y} = (y_1, \dots y_N) \in \mathbb{R}^n$. Let $p, q > 1$ be conjugate indices, i.e.,*

(11.30)
$$\frac{1}{p} + \frac{1}{q} = 1.$$

*Then **Hoelder's inequality** in $\mathbb{R}^n$ is true:*

(11.31)
$$\sum_{j=1}^n |x_j y_j| \leq \|\vec{x}\|_p \|\vec{y}\|_q, \quad i.e., \sum_{j=1}^n |x_j y_j| \leq \Big(\sum_{j=1}^n |x_j|^p\Big)^{1/p} \Big(\sum_{j=1}^n |y_j|^q\Big)^{1/q}.$$

Theorem 11.7 (Minkowski's inequality for $(\mathbb{R}^n, \|\cdot\|_p)$)**.** *Let $n \in \mathbb{N}$ and $\vec{x} = (x_1, \dots x_N)$. Let $\vec{y} = (y_1, \dots y_N) \in \mathbb{R}^n$ and $p \in [1, \infty[$. Then **Minkowski's inequality for** $(\mathbb{R}^n, \|\cdot\|_p)$ is true:*

(11.32)
$$\|\vec{x}+\vec{y}\|_p \leq \|\vec{x}\|_p + \|\vec{y}\|_p, \quad i.e.,$$

(11.33)
$$\Big(\sum_j |x_j + y_j|^p\Big)^{1/p} \leq \Big(\sum_j |f(x)|^p\Big)^{1/p} + \Big(\sum_j |g(x)|^p\Big)^{1/p}.$$

12 Metric Spaces and Topological Spaces – Part I

12.1 Definition and Examples of Metric Spaces

Definition 12.1 (Metric spaces)**.** Let X be an arbitrary, nonempty set.
A **metric** on X is a real–valued function of two arguments

$$d(\cdot,\cdot) : X \times X \to \mathbb{R}, \qquad (x,y) \mapsto d(x,y)$$

with the following three properties:

(12.1a)	$d(x,y) \geq 0 \quad \forall x,y \in X \quad$ and $\quad d(x,y) = 0 \ \Leftrightarrow \ x = y$	**positive definiteness**
(12.1b)	$d(x,y) = d(y,x) \quad \forall x,y \in X$	**symmetry**
(12.1c)	$d(x,z) \leq d(x,y) + d(y,z) \quad \forall\, x,y,z \in X$	**triangle inequality**

Let $x,y \in X$ and $\varepsilon > 0$. We say that x and y are ε–**close** if $d(x,y) < \varepsilon$. The pair $(X, d(\cdot,\cdot))$, usually just written as (X,d), is called a **metric space**. We'll write X for short if it is clear which metric we are talking about. ☐

Remark 12.1 (Metric properties)**.** Let us examine what those properties mean.

"Positive definite":	The distance is never negative and two items x and y have distance zero if and only if they are equal.
"symmetry":	the distance from x to y is no different to that from y to x. That may come as a surprise to you if you have learned in Physics about the distance from point a to point b being the vector $\vec{v}$ that starts in a and ends in b and which is the opposite of the vector $\vec{w}$ that starts in b and ends in a, i.e., $\vec{v} = -\vec{w}$. We only care about size and not about direction.
"Triangle inequality":	If you directly drive from x to z then this will take less fuel than if you make a stopover at an intermediary y. ☐

Proposition 12.1. *Let (X,d) be a metric space. Let $n \in \mathbb{N}$ and $x_1, x_2, \dots, x_n \in X$. Then*

$$d(x_1,x_n) \ \leq \ \sum_{j=1}^{n-1} d(x_j,x_{j+1}) \ = \ d(x_1,x_2) \ + \ d(x_2,x_3) \ + \ d(x_{n-1},x_n). \tag{12.2}$$

Theorem 12.1 (Norms define metric spaces)**.** *Let $(V, \|\cdot\|)$ be a normed vector space. Then the function*

$$d_{\|\cdot\|}(\cdot,\cdot) : V \times V \to \mathbb{R}_{\geq 0}; \qquad (x,y) \mapsto d_{\|\cdot\|}(x,y) := \|y - x\| \tag{12.3}$$

defines a metric space $(V, d_{\|\cdot\|})$.

Definition 12.2 (Metric induced by a norm)**.** We say that the metric $d_{\|\cdot\|}(\cdot,\cdot)$ defined by (12.3) is **induced by the norm** $\|\cdot\|$. and that $d_{\|\cdot\|}(\cdot,\cdot)$ is **derived from the norm** $\|\cdot\|$, or that $d_{\|\cdot\|}(\cdot,\cdot)$ is **associated with the norm** $\|\cdot\|$. □

Definition 12.3 (Discrete metric)**.** Let X be nonempty. Then the function

$$d(x,y) = \begin{cases} 0 & \text{for } x = y \\ 1 & \text{for } x \neq y \end{cases}$$

on $X \times X$ is called the **discrete metric** on X. □

Proposition 12.2. *The discrete metric satisfies the properties of a metric.*

12.2 Measuring the Distance of Real–Valued Functions

Definition 12.4 (Maximal displacement distance between real–valued functions)**.** Let X be an arbitrary, nonempty set and let $f(\cdot), g(\cdot) : X \to \mathbb{R}$ be two real–valued functions on X. We define the **maximal displacement distance** , also called the **sup–norm distance** or $\|\cdot\|_\infty$ **distance** , between $f(\cdot)$ and $g(\cdot)$ as

$$d_\infty(f,g) \ := \ \|f(\cdot) - g(\cdot)\|_\infty \ = \ \sup\{|f(x) - g(x)| : x \in X\}, \tag{12.4}$$

i.e., as the metric induced by the sup–norm on the set $\mathscr{B}(X,\mathbb{R})$ of all bounded real–valued function on X. □

Definition 12.5 (Mean distances between real–valued functions)**.** Let $a, b \in \mathbb{R}$ such that $a < b$ and let $f(\cdot), g(\cdot) : X \to \mathbb{R}$ be two continuous real–valued functions on X. We define the **mean square distance** between $f(\cdot)$ and $g(\cdot)$ on $[a,b]$ as

$$d_{L^2}(f,g) \ := \ d_{\|\cdot\|_{L^2}(f,g)} \ = \ \|g - f\|_{L^2} \ = \ \left(\int_a^b (g(x) - f(x)^2 dx \right)^{1/2}, \tag{12.5}$$

i.e., as the metric induced by the L^2–norm on the set $\mathscr{C}_{\mathscr{B}}$([a,b], $\mathbb{R}$) of all continuous and bounded real–valued function on $[a,b]$.

We further define the **mean distance** between $f(\cdot)$ and $g(\cdot)$ on $[a,b]$ as

$$d_{L^1}(f,g) \ := \ d_{\|\cdot\|_{L^1}(f,g)} \ = \ \|g - f\|_{L^1} \ = \ \int_a^b |g(x) - f(x| dx, \tag{12.6}$$

i.e., as the metric induced by the L^1–norm on the set $\mathscr{C}_{\mathscr{B}}$([a,b], $\mathbb{R}$). □

12.3 Neighborhoods and Open Sets

Definition 12.6 (ε-Neighborhood)**.** Given a metric space (X, d), $x_0 \in X$ and $\varepsilon > 0$, let

(12.7) $$N_\varepsilon(x_0) \;=\; \{x \in X : d(x, x_0) < \varepsilon\}$$

be the set of all elements of X with a distance to x_0 of strictly less than the number ε (the open set around x_0 with "radius" ε from which the points on the boundary (those with distance equal to ε) are excluded). We call $N_\varepsilon(x_0)$ the ε**–neighborhood** of x_0. □

Definition 12.7 (Interior points in metric spaces)**.** Given is a metric space (X, d).
An element $a \in A \subseteq X$ is called an **inner point** or **interior point** of A if we can find some $\varepsilon > 0$ (no mattter how small), so that $N_\varepsilon(a) \subseteq A$. □

Definition 12.8 (Open sets in metric spaces)**.** Given is a metric space (X, d).
A set all of whose members are interior points is called an **open set**. □

Proposition 12.3. *Let (X, d) be a metric space. Let $x, y \in X$ and $\varepsilon > 0$ such that $y \in N_\varepsilon(x)$.*

If $\delta > 0$ Then $N_\delta(y) \subseteq N_{\delta+\varepsilon}(x)$.

Proposition 12.4. *$N_\varepsilon(x_0)$ is an open set*

Proposition 12.5 (Open intervals are open in $(\mathbb{R}, d_{|\cdot|})$)**.**
Let $a, b \in \mathbb{R}$ such that $a < b$. Then the open interval $]a, b[$ is an open set in $(\mathbb{R}, d_{|\cdot|})$.

Definition 12.9 (Neighborhoods in Metric Spaces)**.** Let (X, d) be a metric space, $x_0 \in X$. Any open set that contains x_0 is called an **open neighborhood** of x_0. Any superset of an open neighborhood of x_0 is called a **neighborhood** of x_0. □

Remark 12.2.

(a) You will see very often that **the important neighborhoods are the small ones**, not the big ones. The definition above says that, for any neighborhood A_x of a point $x \in X$, one can find an open neighborhood U_x of x such that $U_x \subseteq A_x$. Thus, very often **the open neighborhoods are the important ones**. Accordingly, there are many theorems where it is assumed that some given neighborhood is open.

(b) The empty set is not a neighborhood of any $x \in X$, since the condition $x \in \emptyset$ is never satisfied. □

Proposition 12.6 (Metric Spaces are Hausdorff Spaces)**.**
Let (X, d) be a metric space and let x, y be two different elements of X. Then there exist neighborhoods N_x of x and N_y of y such that $N_x \cap N_y = \emptyset$.

Theorem 12.2 (Metric spaces are topological spaces)**.**
The following is true about open sets of a metric space (X, d):

(12.8a) *An arbitrary union* $\bigcup_{i \in I} U_i$ *of open sets U_i is open.*

(12.8b) *A finite intersection* $U_1 \cap U_2 \cap \ldots \cap U_n$ $(n \in \mathbb{N})$ *of open sets is open.*

(12.8c) *The entire set X is open and the empty set $\emptyset$ is open.*

12.4 Convergence

Definition 12.10 (Convergence of Sequences in Metric Spaces)**.** Given is a metric space (X, d). We say that a sequence (x_n) of elements of X **converges** to $a \in X$ for $n \to \infty$ if the x_n will eventually come arbitrarily close to a in the following sense:

Let δ be a (arbitrarily small) positive real number. Then there is a (possibly extremely large) integer n_0 such that all x_j belong to $N_\delta(a)$ just as long as $j \geq n_0$.

This can also be expressed as follows:

(12.9) For all $\delta > 0$ there exists $n_0 \in \mathbb{N}$ such that $d(a, x_j) < \delta$ for all $j \geq n_0$.

Here is an yet another way of expressing convergence of $(x_n)_n$ to a:

- No matter how small a neighborhood of a is given, all members x_n will eventually be inside that neighborhood.

We write either of

(12.10) $$a = \lim_{n \to \infty} x_n \qquad \text{or} \qquad x_n \to a$$

and we call a the **limit** of the sequence (x_n) □

Theorem 12.3 (Limits in metric spaces are uniquely determined)**.**
Let (X, d) be a metric space and let $(x_n)_n$ be a convergent sequence in X. Then its limit is uniquely determined.

Proposition 12.7. *Let (X, d) be a metric space and $L, x_n \in X$ $(n \in \mathbb{N})$. Let $\delta_n \in \mathbb{R}_{>0}$ such that $\delta_n \to 0$ as $n \to \infty$. Assume further that $x_n \in N_{\delta_n}(L)$ for all $n \in \mathbb{N}$. Then $\lim\limits_{n\to\infty} x_n = L$.*

Corollary 12.1. *Let (X, d) be a metric space and $L, x_n \in X$ $(n \in \mathbb{N})$ such that $d(x_n, L) \leq \frac{1}{n}$ for all $n \in \mathbb{N}$.*
Then $\lim\limits_{n\to\infty} x_n = L$.

Proposition 12.8. *Let (X, d) be a metric space, $L \in X$ and $x_n = L$ for all $n \in \mathbb{N}$. Then $\lim\limits_{n\to\infty} x_n = L$.*

Proposition 12.9. *Let x_n, y_n be two sequences in a metric space (X, d). Assume there is $K \in \mathbb{N}$ such that $x_n = y_n$ for all $n \geq K$. Let $L \in X$ Then*

$$\lim_{n\to\infty} x_n = L \Leftrightarrow \lim_{n\to\infty} y_n = L.$$

Proposition 12.10 (Subsequences of sequences with limits)**.**
Let $(x_n)_n$ be a sequence in a metric space (X, d) with limit $L := \lim\limits_{n\to\infty} x_n$. Then it is true for any subsequence $(x_{n_j})_j$, that $\lim\limits_{j\to\infty} x_{n_j} = L$.

Proposition 12.11. *Let x_n be a convergent sequence in a metric space (X, d) with limit $L \in X$. Let $K \in \mathbb{N}$. For $n \in \mathbb{N}$ let $y_n := x_{n+K}$. Then $\lim\limits_{n\to\infty} (y_n)_n = L$.*

Remark 12.3. The following allows us to prove convergence of x_n to $L \in (X, d)$ by utilizing what we know about convergence in $(\mathbb{R}, d_{|\cdot|})$.

$$\lim_{n\to\infty} x_n = L \Leftrightarrow \lim_{n\to\infty} d(x_n, L) = 0. \square$$

Remark 12.4 (Opposite of convergence). $\big[\ \lim_{k\to\infty} x_k = L$ is **NOT true** $\big]$ $\Leftrightarrow$
$\big[$ there exists some $\varepsilon > 0$ such that for **all** $N \in \mathbb{N}$ there exists some natural number $j = j(N)$ such that $j \geq N$ and $d(x_j, L) \geq \varepsilon\,\big]$. $\square$

Proposition 12.12 (Opposite of convergence).
$\big[$ *A sequence* $(x_k)_k$ *with values in* (X,d) *does not have* $L \in X$ *as its limit* $\big]$ $\Leftrightarrow$
$\big[$ *there exists some* $\varepsilon > 0$ *and* $n_1 < n_2 < n_3 < \cdots \in \mathbb{N}$ *such that* $d(x_{n_j}, L) \geq \varepsilon$ *for* ***all*** j. $\big]$
In other words, there is a subsequence $(x_{n_j})_j$ *which completely stays out of some* ε*–neighborhood of* L.

12.5 Abstract Topological spaces

Definition 12.11 (Abstract topological spaces). Let X be an arbitrary nonempty set and let $\mathfrak{U}$ be a set of subsets of X whose members satisfy the properties a, b and c of (12.8) on p.119: [16]

(12.11a) An arbitrary union $\bigcup\limits_{i\in I} U_i$ of sets $U_i \in \mathfrak{U}$ belongs to $\mathfrak{U}$,

(12.11b) $U_1, U_2, \dots, U_n \in \mathfrak{U}\ (n \in \mathbb{N}) \quad\Rightarrow\quad U_1 \cap U_2 \cap \ldots \cap U_n \in \mathfrak{U}$,

(12.11c) $X \in \mathfrak{U}$ and $\emptyset \in \mathfrak{U}$.

Then $(X, \mathfrak{U})$ is called a **topological space** The members of $\mathfrak{U}$ are called **open sets** of $(X, \mathfrak{U})$. The collection $\mathfrak{U}$ of open sets is called the **topology** of X. $\square$

Every metric space (X,d) is a topological space in the following sense: If $\mathfrak{U}_d$ denotes the open sets of (X,d) then $(X, \mathfrak{U}_d)$ is a topological space.

Every normed vector space $(V, \|\cdot\|)$ is a topological space in the sense that If $\mathfrak{U}_d$ denotes the open subsets of a metric space (X,d) then $(V, \mathfrak{U}_d)$ is a topological space. $\square$

Definition 12.12 (Metric Topology and Norm Topology). ⋆

(a) Let (X,d) be a metric space and let $\mathfrak{U}_d$ be as defined in (**??**). We say that $\mathfrak{U}_d$ is **induced by the metric** $d(\cdot,\cdot)$ or that it is **generated by the metric** $d(\cdot,\cdot)$. or that it is the **metric topology** of X. If it is clear which metric d on X we mean then we also simply refer to "the" metric topology.

[16]Note that we encountered subsets of 2^X with special properties previously when looking at rings of sets in Definition 8.4 (Rings, algebras, and σ–algebras of Sets) on p.74.

(b) Let $(V, \|\cdot\|)$ be a normed vector space, and let $\mathfrak{U}_{\|\cdot\|}$ be as defined in (**??**), i.e., $\mathfrak{U}_{\|\cdot\|}$ is the topology defined by the metric $d_{\|\cdot\|}$. We say that this topology is **induced by the norm** $\|\cdot\|$ or that it is **generated by the norm** $\|\cdot\|$. If it is clear which norm on V we are studying then we call the topology associated with this norm the **norm topology** of V. □

Definition 12.13 (Discrete topology). ⋆ Let X be a nonempty set with the discrete metric

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

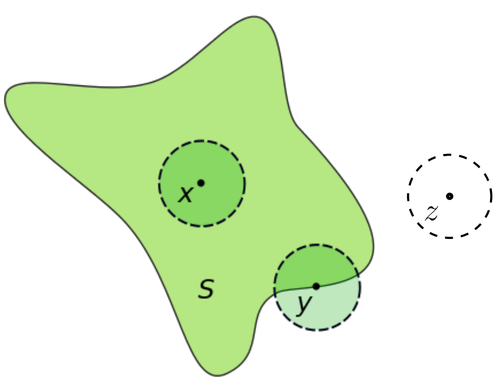
We call the topology associated with the discrete metric the **discrete topology** of X. □

Proposition 12.13. *Let (X,d) be a metric space with the discrete metric.*
Then its associated topology is $\mathfrak{U}_d = 2^X = \{A : A \subseteq X\}$.

Proposition 12.14. *Let X be an arbitrary nonempty set and let $\mathfrak{U} := \{\emptyset, X\}$.*
Then $(X, \mathfrak{U})$ is a topological space.

Definition 12.14 (Indiscrete topology)**.** Let X be a nonempty set.
The topology $\{\emptyset, X\}$ is called the **indiscrete topology** of X. □

Remark 12.5.



The picture to the right [17] demonstrates that there are exactly three mutually exclusive choices how a point in $(X, \mathfrak{U})$ is related to a subset S of X:

(a) either like the point x: There exists an open set U such that $x \in U \subseteq S$,

(b) or like the point z: There exists an open set U such that $z \in U \subseteq S^{\complement}$,

(c) or like the point y: There is no open set U such that $y \in U \subseteq S$ or $y \in U \subseteq S^{\complement}$, i.e., every open set that contains y intersects both S and $S^{\complement}$.

We can classify any element $x \in X$ accordingly: x satisfies either **(a)** or **(b)** or **(c)**.

[17] Source: Wikipedia, https://en.wikipedia.org/wiki/Interior_(topology). The author does not like to use the letter S for subsets of topological spaces, but it came with the picture.

Definition 12.15 (Neighborhoods and interior points in topological spaces)**.** Let $(X, \mathfrak{U})$ be a topological space, $x \in X$ and $S \subseteq X$. It is not assumed that S be open.

(a) S is called a **neighborhood** of x and x is called an **inner point** or **interior point** of S if there exists an open set U such that

$$x \in U \subseteq S.$$

We call the set $S^o := \{$ all interior points of S$\}$ the **interior** of S. An alternate but less commonly used notation for S^o is int(S).

(b) x is called an **exterior point** of S if x is an inner point of $S^{\complement}$, i.e., there exists an open set U' such that

$$x \in U' \subseteq S^{\complement},$$

We call the set ext(S) $:= \{$ all exterior points of S$\}$ the **open exterior** [18]of S.

(c) x is called a **boundary point** of S if any neighborhood of x intersects both S and $S^{\complement}$. We call this set the **boundary** of S and denote it ∂S. ☐

If S is an arbitrary subset of X, U is an open subset of X, and $x \in X$, then

(a) x is an interior point of $S \Leftrightarrow S$ is a neighborhood of x.
(b) x is an interior point of $U \Leftrightarrow x \in U$.
(c) If $U \subseteq S$ then all elements of U are interior points of S, i.e., $U \subseteq S^o$.

Proposition 12.15. *Let $(X, \mathfrak{U})$ be a topological space and let $A \subseteq X$. Then*

$$A^o = \bigcup \Big[U \in \mathfrak{U} : U \subseteq A \Big]. \tag{12.12}$$

In other words, the interior of A is the union of all open subsets of A.

The interior A^o of A is the largest of all open subsets of A.

Proposition 12.16. *Let $(X, \mathfrak{U})$ be a topological space.*

If $A \subseteq B \subseteq X$ then $A^o \subseteq B^o$. ☐

[18]The expression "open exterior" has been adopted from Wikipedia.
Source: https://en.wikipedia.org/wiki/Interior_(topology)

Proposition 12.17. *Let $(X, \mathfrak{U})$ be a topological space and let $A \subseteq X$. Then,*

$$X = A^o \biguplus ext(A) \biguplus \partial(A) . \tag{12.13}$$

Thus, X is partitioned into the interior, open exterior and boundary of anyone of its subsets.

Theorem 12.4 (Hierarchy of topological spaces)**.** *We have seen the following:*

- ***(a)*** *$\mathbb{R}^n$, in particular $\mathbb{R} = \mathbb{R}^1$, is an inner product space (see prop.11.11 on p.111).*
- ***(b)*** *All inner product spaces are normed spaces (see thm.11.3 on p.112).*
- ***(c)*** *All normed spaces are metric spaces (see thm.12.1 on p.116).*
- ***(d)*** *All metric spaces are topological spaces (see Definition 12.11 on p.121, Definition 12.12 on p.121).*

12.6 Bases and Neighborhood Bases ⋆

Definition 12.16 (Base of the topology)**.** Let $(X, \mathfrak{U})$ be a topological space. A subset $\mathfrak{B}$ of $\mathfrak{U}$ of open sets is called a **base of the topology** if any nonempty open set U can be written as a union of elements of $\mathfrak{B}$:

$$U = \bigcup_{i \in I} B_i \quad (B_i \in \mathfrak{B} \text{ for all } i \in I) \tag{12.14}$$

where I is a suitable index set, which of course will in general depend on U. □

Definition 12.17 (Neighborhood base of a point)**.** Let $(X, \mathfrak{U})$ be a topological space.

(a) The following set of subsets of X,

$$\mathfrak{N}(x) := \{A \subseteq X : A \text{ is a neighborhood of } x\} , \tag{12.15}$$

is called the **neighborhood system of** x

(b) Given a point $x \in X$, any subset $\mathfrak{B} := \mathfrak{B}(x) \subseteq \mathfrak{N}(x)$ of the neighborhood system of x is called a **neighborhood base of** x if it satisfies the following condition:

- For any $A \in \mathfrak{N}(x)$, there exists a set $B \in \mathfrak{B}(x)$ such that $B \subseteq A$. □

Definition 12.18 (First axiom of countability)**.** Let $(X, \mathfrak{U})$ be a topological space.
We say that X satisfies the **first axiom of countability** or X is **first countable** if we can find for each $x \in X$ a countable neighborhood base. □

Proposition 12.18 (ε-neighborhoods are a base of the topology)**.**
Let (X,d) be a metric space. Then both

$$\mathscr{B}_1 := \{N_\varepsilon(x) : x \in X, \varepsilon > 0\} \quad \text{and} \quad \mathscr{B}_2 := \{N_{1/n}(x) : x \in X, n \in \mathbb{N}\}$$

are bases for the topology of (X,d) (see 12.16 on p.124)

Theorem 12.5 (Metric spaces are first countable)**.** *Let (X,d) be a metric space. Then X is first countable.*

Proposition 12.19. *Let (X,d) be a metric space and let $\mathfrak{B} := \{N_{1/k}(x) : x \in X, k \in \mathbb{N}\}$. Then $\mathfrak{B}$ is a base of the topology for the associated topological space $(X, \mathfrak{U}_d)$.*

Definition 12.19 (Second axiom of countability)**.** Let $(X, \mathfrak{U})$ be a topological space.
We say that X satisfies the **second axiom of countability** or X is **second countable** if we can find a countable base for $\mathfrak{U}$. □

Theorem 12.6 (Euclidean space $\mathbb{R}^n$ is second countable)**.**
Let $\mathfrak{B}$ be the following collection of open subsets of $\mathbb{R}^n$:

(12.16) $$\mathfrak{B} := \{\, N_{1/j}(\vec{q}) : \vec{q} \in \mathbb{Q}^n,\ j \in \mathbb{N}\,\}.$$

Here,

$$\mathbb{Q}^n = \{\vec{q} = (q_1, \dots, q_n) : q_j \in \mathbb{Q},\ 1 \le j \le n\}$$

is the set of all points in $\mathbb{R}^n$ with rational coordinates. Then $\mathfrak{B}$ is a countable base of $\mathbb{R}^n$.

12.7 Metric and Topological Subspaces

Definition 12.20 (Metric subspaces)**.** Given is a metric space (X,d) and a nonempty $A \subseteq (X,d)$.
Let

$$d\big|_{A \times A} : A \times A \to \mathbb{R}_{\ge 0} \quad \text{be the restriction} \quad d\big|_{A \times A}(x,y) := d(x,y)(x,y \in A)$$

of the metric d to $A \times A$ (see Definition 5.15 on p.45).
It is trivial to verify that $(A, d\big|_{A \times A})$ is a metric space in the sense of Definition 12.1 on p.116.
We call $(A, d\big|_{A \times A})$ a **metric subspace** of (X,d) and we call $d\big|_{A \times A}$ the **metric induced by** d or the **metric inherited from** (X,d). □

Remark 12.6.

Metric subspaces come with their own collections of open and closed sets, neighborhoods, ε-neighborhoods, convergent sequences, ...
Watch out when looking at statements and their proofs whether those concepts refer to the entire space (X,d) or to the subspace $(A, d\big|_{AxA})$. □

Notation 12.1.

a) Because the only difference between d and $d_{A\times A}$ is the domain, it is customary to write d instead of $d\big|_{A\times A}$ to make formulas look simpler, if doing so does not give rise to confusion.

b) We often shorten "open in $(A, d\big|_{A\times A})$" to "open in A", "closed in $(A, d\big|_{A\times A})$" to "closed in A", "convergent in $(A, d\big|_{A\times A})$" to "convergent in A", ... □

Definition 12.21 (Traces of sets in a metric subspace). ★

Let (X, d) be a metric space and $A \subseteq X$ a nonempty subset of X, viewed as a metric subspace $(A, d\big|_{A\times A})$ of (X, d). Let $Q \subseteq X$.

We call $Q \cap A$ the **trace** of Q in A.

For $\varepsilon > 0$ and $a \in A$ let $N_\varepsilon(a)$ be the ε-neighborhood of a (in (X, d)). We define

$$N_\varepsilon^A(a) \;=\; N_\varepsilon(a) \cap A. \tag{12.17}$$

i.e., $N_\varepsilon^A(a)$ is defined as the trace of $N_\varepsilon(a)$ in A. □

Proposition 12.20 (Open sets in metric subspaces are traces of open sets in X).

Let (X, d) be a metric space and $A \subseteq X$ a nonempty subset of X.

(a) *Let $\varepsilon > 0$ and $a \in A$. Then*

$$N_\varepsilon^A(a) \;=\; N_\varepsilon(a) \cap A \;=\; \{x \in A : d\big|_{A\times A}(x, a) < \varepsilon\}\,, \tag{12.18}$$

i.e., $N_\varepsilon^A(a)$ is the "ordinary" ε–Neighborhood of a in the metric space $(A, d\big|_{A\times A})$ (as it was originally defined in Definition 12.6 on p.118). It thus follows from (12.17) that each ε-neighborhood in the subspace A is the trace of an ε-neighborhood in X.

(b) *Generalization: $U \subseteq A$ is open in $(A, d\big|_{A\times A})$ $\Leftrightarrow$ there is an open $V \subseteq (X, d)$ such that*

$$U \;=\; V \cap A\,. \tag{12.19}$$

In other words, U is the trace of a set V which is open in X.

Remark 12.7 (Convergence does not necessarily extend to metric subspaces).

Let (X, d) be a metric space, $A \subseteq (X, d)$ and $a_n \in A$ for all $n \in \mathbb{N}$. Be aware that convergence of the sequence (a_n) in the space (X, d) (i.e., there exists $x \in X$ such that $x = \lim\limits_{n\to\infty} a_n$) does **NOT** imply convergence of the sequence in the subspace $(A, d\big|_{A\times A})$! Rather, we have the following <u>dichotomy</u>:

(a) $x \in A$: Then a_n converges to x in the subspace $(A, d\big|_{A\times A})$ (and also in (X, d)).

(b) $x \in A^\complement$: Then a_n converges to x in (X, d) but not in $(A, d\big|_{A\times A})$. □

Definition 12.22 (Topological subspaces). ⋆

Let $(X, \mathfrak{U})$ be a topological space and $A \subseteq X$. We say that $V \subseteq A$ is **open in A** if V is the trace of an open set in X, i.e., if there is some $U \in \mathfrak{U}$ such that $V = U \cap A$. We denote the collection of all open sets in A as $\mathfrak{U}_A$. In other words,

$$\mathfrak{U}_A = \{V \cap A : V \in \mathfrak{U}\}.$$

We call $(A, \mathfrak{U}_A)$ a **topological subspace** or also just a **subspace** of $(X, \mathfrak{U})$ and we call $\mathfrak{U}_A$ the **subspace topology induced by** $(X, \mathfrak{U})$ or the **subspace topology inherited from** $(X, \mathfrak{U})$. □

Proposition 12.21 (Topological subspaces are topological spaces). *Let $(X, \mathfrak{U})$ be a topological space, $A \subseteq X$, and let $\mathfrak{U}_A$ be the collection of all open sets in A. Then $(A, \mathfrak{U}_A)$ is a topological space, i.e., it satisfies Definition 12.11 on p.121 of an abstract topological space.*

12.8 Contact Points and Closed Sets

Definition 12.23 (Contact points). Given is a topological space $(X, \mathfrak{U})$.
Let $A \subseteq X$ and $x \in X$ (x may or may not to belong to A). x is called a **contact point**, of A if

(12.20) $$A \cap N \neq \emptyset \text{ for any neighborhood } N \text{ of } x. \quad \square$$

Definition 12.24 (Closed sets). Let $(X, \mathfrak{U})$ be topological space and $A \subseteq X$. Let the set $\bar{A}$ be

(12.21) $$\bar{A} := \{x \in X : x \text{ is a contact point of } A\}.$$

We call $\bar{A}$ the **closure** of A. A set that contains all its contact points is called a **closed set**. □

Proposition 12.22. *If A is a subset of a topological space then*

(12.22) $$\bar{A} = A \cup \partial(A) = A^o \cup \partial(A).$$

Theorem 12.7 (Sequence criterion for contact points in metric spaces).
Given is a metric space (X, d). Let $A \subseteq X$ and $x \in X$. Then x is a contact point of A if and only if there exists a sequence $x_1, x_2, x_3, \dots$ of members of A which converges to x.

Theorem 12.8 (Open iff complement is closed).
Let (X, d) be a metric space and $A \subseteq X$. Then A is open if and only if $A^{\complement}$ is closed.

Proposition 12.23. *Let $(X, \mathfrak{U})$ be a topological space. The closed sets of X satisfy the following:*

(12.23)
- ***(a)*** *An arbitrary intersection of closed sets is closed.*
- ***(b)*** *A finite union of closed sets is closed.*
- ***(c)*** *The entire set X is closed and $\emptyset$ is closed.*

Proposition 12.24. *Let $(X, \mathfrak{U})$ be a topological space and $A \subseteq B \subseteq X$. Then $\bar{A} \subseteq \bar{B}$.*

Proposition 12.25. *Let $(X, \mathfrak{U})$ be a topological space and $A \subseteq X$. Then,*

$$\partial A = \bar{A} \cap \overline{A^{\complement}}. \tag{12.24}$$

In other words, $x \in X$ is a boundary point of A if and only if x is a contact point of both A and $A^{\complement}$.

Proposition 12.26 (Minimality of the closure of a set).
Let $(X, \mathfrak{U})$ be a topological space and $A \subseteq X$. Then

$$\bar{A} = \bigcap \Big[C \supseteq A : C \text{ is closed} \Big]. \tag{12.25}$$

In other words, the closure $\bar{A}$ of A is the smallest of all closed supersets of A.

Proposition 12.27 (Closure of a set as a hull operator).
Let $(X, \mathfrak{U})$ be a topological space. Consider the closure of sets as a function

$$^{-} : 2^X \longrightarrow 2^X; \qquad A \mapsto \bar{A}.$$

Then this function has the following properties for all $A, B \subseteq X$:

(a) $\bar{\emptyset} = \emptyset$, ***(b)*** $A \subseteq \bar{A}$, ***(c)*** $\bar{\bar{A}} = \bar{A}$, ***(d)*** $\overline{A \cup B} = \bar{A} \cup \bar{B}$.

Definition 12.25 (Contact points vs Limit points). ⋆

Given is a topological space $(X, \mathfrak{U})$. Let $A \subseteq X$ and $x_0 \in X$. x_0 is called a **limit point** or **cluster point** or **point of accumulation** of A if every neighborhood U of x_0 intersects A in at least one point <u>other than x_0</u>, i.e.,

$$U \cap (A \setminus x_0) \neq \emptyset. \ \square$$

12.9 Bounded Sets and Bounded Functions in Metric Spaces

Definition 12.26 (bounded sets). Given is a subset A of a metric space (X, d).
The **diameter** of A is defined as

(12.26) $$\text{diam}(\emptyset) := 0, \qquad \text{diam}(A) := \sup\{d(x,y) : x, y \in A\} \text{ if } A \neq \emptyset.$$

We call A a **bounded set** if $\text{diam}(A) < \infty$. □

Proposition 12.28. *Given is a metric space (X, d) and a nonempty subset A. The following are equivalent:*

(12.27)
(a) $\text{diam}(A) < \infty$, i.e., A is bounded.
(b) There exists $\gamma > 0$ and $x_0 \in X$ such that $A \subseteq N_\gamma(x_0)$.
(c) For all $x \in X$ there exists $\gamma > 0$ such that $A \subseteq N_\gamma(x)$.

Proposition 12.29. *Let (X, d) be a metric space. For $n \in \mathbb{N}$ let $A_n \subseteq X$ such that $\delta_n :=$ diam$(A_n) \to 0$ as $n \to \infty$. Let $A := \bigcap_n A_n$. Then,*

$$\text{either} \quad A = \emptyset, \quad \text{or} \quad \text{there is some } a \in X \text{ such that } A = \{a\}.$$

Proposition 12.30. *Let (X, d) be a metric space and $A \subseteq X$. Then,*

$$\text{diam}(A) = \text{diam}(\bar{A}).$$

Proposition 12.31. *Let (X, d) be a metric space. Let $A_1 \supseteq A_2 \supseteq \dots$ be subsets of X such that diam$(A_n) \to 0$ as $n \to \infty$ and let $A := \bigcap_j \bar{A}_j$. Let $x_n \in A_n$ for all n. Then*

- *$(x_n)_n$ converges if and only if A is not empty.*
- *If $A \neq \emptyset$, then A is the singleton set $A = \left\{\lim_{n\to\infty} x_n\right\}$.*

12.10 Completeness in Metric Spaces

Definition 12.27 (Cauchy sequences [19]). Given is a metric space (X, d). A sequence (x_n) in X is called a **Cauchy sequence** or, in short, it is Cauchy if for any $\varepsilon > 0$ (no matter how small), there exists some index $n_0 \in \mathbb{N}$ such that

(12.28) $$d(x_i, x_j) < \varepsilon \qquad \text{for all } i, j \geq n_0$$

This is called the **Cauchy criterion for convergence** of a sequence. □

Example 12.1 (Cauchy criterion for real numbers)**.** In $\mathbb{R}$ we have $d(x,y) = |x-y|$ and the Cauchy criterion requires for any given $\varepsilon > 0$ the existence of $n_0 \in \mathbb{N}$ such that

$$|x_i - x_j| \ < \ \varepsilon \qquad \text{for all } i,j \geq n_0. \quad \square \tag{12.29}$$

Proposition 12.32. *Let (X,d) be a metric space and $x_n \in X$ $(n \in \mathbb{N})$. Then the following are equivalent:*

- ***(a)*** *$(x_n)_n$ is Cauchy.*
- ***(b)*** *The diameters of the tail sets $T_n = \{x_j : j \geq n\}$ converge to zero.*
- ***(c)*** *There exists a nonincreasing sequence $A_1 \supseteq A_2 \supseteq \dots$ of subsets of X such that $x_n \in A_n$ and diam$(A_n) \to 0$ as $n \to \infty$.*

Proposition 12.33. *A Cauchy sequence in a metric space is bounded.*

Theorem 12.9 (Convergent sequences are Cauchy)**.**
Let $(x_n)_n$ be a convergent sequence in a metric space (X,d). Then $(x_n)_n$ is Cauchy.

Proposition 12.34. *Let $(x_n)_n$ be a Cauchy sequence in a metric space (X,d).*
If some subsequence x_{n_j} converges to a limit x_0. Then

- ***(a)*** *ANY subsequence of $(x_n)_n$ converges to L.*
- ***(b)*** *$(x_n)_n$ is a convergent sequence.*

Further, any subsequence y_{n_j} of a convergent sequence $(y_n)_n$ converges to the limit of $(y_n)_n$.

Definition 12.28 (Completeness in metric spaces)**.** A subset A of a metric space (X,d) is called **complete**, if any Cauchy sequence (a_n) with elements in A converges to some $a \in A$. $\square$

Theorem 12.10 (Completeness of the real numbers)**.**
Let (x_n) be a Cauchy sequence in $\mathbb{R}$. then there exists a real number L such that $L = \lim\limits_{n\to\infty} x_n$.

Theorem 12.11 (Completeness of $\mathbb{R}^n$)**.**
Let $(\vec{x}_j)$ be a Cauchy sequence in $\mathbb{R}^n$. Then there exists $\vec{a} \in \mathbb{R}^n$ such that $\vec{a} = \lim\limits_{j\to\infty} \vec{x}_j$.

[19] Cauchy sequence are named after the great french mathematician Augustin–Louis Cauchy (1789–1857) who contributed massively to the most fundamental ideas of Calculus.

Proposition 12.35. *Let* $\vec{x}_j = (x_{j,1}, x_{j,2}, \dots, x_{j,n})$ *and* $\vec{b} \in \mathbb{R}^n$. *Then,*

$$\lim_{j\to\infty} \vec{x}_j = \vec{b} \Leftrightarrow \lim_{j\to\infty} x_{j,k} = b_k \text{ for all } 1 \le k \le n\,. \tag{12.30}$$

Proposition 12.36. *The metric space* $(\mathbb{Q}, d_{|\cdot|})$ *(Euclidean metric) is not complete.*

Proposition 12.37. *Let* d *be the discrete metric on a nonempty set* X *and let* $(x_n)_n$ *a sequence in* X. *Then,*

$(x_n)_n$ *is Cauchy* $\Leftrightarrow$ $(x_n)_n$ *converges* $\Leftrightarrow$ $(x_n)_n$ *is constant eventually.*

Corollary 12.2. *Discrete metric spaces are complete.*

Theorem 12.12. *Any complete subset of a metric space is closed.*

Theorem 12.13 (Closed subsets of a complete space are complete)**.**
Let (X, d) *be a complete metric space and let* $A \subseteq X$ *be closed. Then* A *is complete, i.e., the metric subspace* $(A, d\big|_{A\times A})$ *is complete.*

13 Metric Spaces and Topological Spaces – Part II

13.1 Continuity

13.1.1 Definition and Characterizations of Continuous Functions

Definition 13.1 (Sequence continuity)**.** Given are two metric spaces (X, d_1) and (Y, d_2). Let $A \subseteq X$, $x_0 \in A$ and let $f : A \to Y$ be a mapping from A to Y. We say that f is **sequence continuous at x_0** and we write

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

if the following is true for any sequence (x_n) with values in A:

(13.2) $$\text{if } x_n \to x_0 \text{ then } f(x_n) \to f(x_0).$$

In other words, the following must be true for any sequence (x_n) in A and $x_0 \in A$:

(13.3) $$\lim_{n\to\infty} x_n = x_0 \;\Rightarrow\; \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x_0).$$

We say that f is **sequence continuous** if f is sequence continuous at x_0 for all $x_0 \in A$. □

Definition 13.2 (ε-δ continuity)**.** Given are two metric spaces (X, d_1) and (Y, d_2). Let $A \subseteq X$, $x_0 \in A$ and let $f(\cdot) : A \to Y$ be a mapping from A to Y. We say that $f(\cdot)$ is ε–δ **continuous at x_0** if the following is true: For any (whatever small) $\varepsilon > 0$ there exists $\delta > 0$ such that either one of the following equivalent statements is satisfied:

(13.4) $$f\big(N_\delta(x_0) \cap A\big) \;\subseteq\; N_\varepsilon(f(x_0)),$$

(13.5) $$d_1(x, x_0) < \delta \;\Rightarrow\; d_2(f(x), f(x_0)) < \varepsilon \;\text{ for all } x \in A.$$

We say that $f(\cdot)$ is ε-δ **continuous** if $f(\cdot)$ is ε-δ continuous at a for all $a \in A$. □

f is ε-δ continuous at x_0 $\Leftrightarrow$ for all $\varepsilon > 0$ there exists $\delta > 0$ s.t. $f\big(N_\delta^A(x_0)\big) \;\subseteq\; N_\varepsilon(f(x_0))$. □

Theorem 13.1 (Continuity criterion)**.**
Let (X, d_1) and (Y, d_2) be two metric spaces. Let $A \subseteq X$, $x_0 \in A$ and let $f(\cdot) : A \to Y$. Then,

- *f is sequence continuous at x_0 $\Leftrightarrow$ f is ε-δ continuous at x_0.*
- *In particular f is sequence continuous (on A) if and only if f is ε-δ continuous.*

Definition 13.3 (Continuity in metric spaces)**.** From now on we can use the terms "ε-δ continuous at x_0" and "sequence continuous at x_0" interchangeably for functions between metric spaces and we will simply speak about **continuity of** f **at** x_0. □

Theorem 13.2 (Neighborhood characterization of continuity)**.** *Let (X, d_1) and (Y, d_2) be two metric spaces. Let $A \subseteq X$, $x_0 \in A$, and let $f(\cdot) : A \to Y$ be a mapping from A to Y. Then*

f is continuous at x_0 if and only if for any neighborhood $V_{f(x_0)}$ of $f(x_0)$, there exists a neighborhood U_{x_0} of x_0 in the metric space (X, d_1), such that

$$f(U_{x_0} \cap A) \subseteq V_{f(x_0)}\,. \tag{13.6}$$

Equivalently, (13.6) can be stated in terms of the subspace $(A, d_1\big|_{A\times A})$ as follows.

for any neighborhood $V_{f(x_0)}$ of $f(x_0)$ there exists a neighborhood $U^A_{x_0}$ of x_0 in the metric space $(A, d_1\big|_{A\times A})$ such that

$$f(U^A_{x_0}) \subseteq V_{f(x_0)}. \tag{13.7}$$

Theorem 13.3 (Rules of arithmetic for continuous real–valued functions)**.**
Given is a metric space (X, d). Let the functions

$$f(\cdot), g(\cdot), f_1(\cdot), f_2(\cdot), f_3(\cdot), \dots, f_n(\cdot) : A \longrightarrow \mathbb{R}$$

all be continuous at $x_0 \in A \subseteq X$. Then

- ***(a)*** *Constant functions are continuous everywhere on A.*
- ***(b)*** *The product $fg(\cdot) : x \mapsto f(x)g(x)$ is continuous at x_0. Specifically, $\alpha f(\cdot) : x \mapsto \alpha \cdot f(x)$ where $\alpha \in \mathbb{R}$ is continuous at x_0. In particular ($\alpha = -1$) the function $-f(\cdot) : x \mapsto -f(x)$ is continuous at x_0.*
- ***(c)*** *The sum $f + g(\cdot) : x \mapsto f(x) + g(x)$ is continuous at x_0.*
- ***(d)*** *If $g(x_0) \neq 0$ then the quotient $f/g(\cdot) : x \mapsto f(x)/g(x)$ is continuous at x_0.*
- ***(e)*** *Any linear combination [20] $\sum_{j=0}^{n} a_j f_j(\cdot) : x \mapsto \sum_{j=0}^{n} a_j f_j(x)$ is continuous in x_0.*

Definition 13.4 (Continuity for topological spaces)**.** Given are two topological spaces $(X, \mathfrak{U}_1)$ and $(Y, \mathfrak{U}_2)$. Let $A \subseteq X$, $x_0 \in A$ and let $f : A \to Y$ be a mapping from A to Y.

We say that f is **continuous at x_0** if the following is true:

For any neighborhood $V_{f(x_0)}$ of $f(x_0)$, there exists a neighborhood U_{x_0} of x_0 in the topological space $(X, \mathfrak{U}_1)$, such that

$$f(U_{x_0} \cap A) \subseteq V_{f(x_0)}\,. \tag{13.8}$$

Equivalently, continuity at x_0 can be stated in terms of the subspace $(A, \mathfrak{U}_{\mathbf{1}_A})$ as follows.

For any neighborhood $V_{f(x_0)}$ of $f(x_0)$ there is a neighborhood $U^A_{x_0}$ of x_0 in $(A, \mathfrak{U}_{\mathbf{1}_A})$ such that

$$f(U^A_{x_0}) \subseteq V_{f(x_0)}. \tag{13.9}$$

We say that f is **continuous** if f is continuous at a for all $a \in A$. □

Proposition 13.1 (“f^{-1}(open) = open” continuity).
Let $(X, \mathfrak{U})$ and $(Y, \mathfrak{V})$ be two topological spaces and let $f : X \to Y$. Then

- *f is continuous (on X) $\Leftrightarrow$ All preimages $f^{-1}(V)$ of open $V \subseteq Y$ are open in X.*

Proposition 13.2 (The composition of continuous functions is continuous).
Let $(X, \mathfrak{U})$, $(Y, \mathfrak{V})$ and $(Z, \mathfrak{W})$ be topological spaces.
Let $f : X \to Y$ be continuous at $x_0 \in X$ and $g : Y \to Z$ continuous at $f(x_0)$.

- *Then the composition $g \circ f : X \to Y$ is continuous at x_0.*

Proposition 13.3 (continuity of constant functions).
Let $(X, \mathfrak{U})$ and $(Y, \mathfrak{V})$ be topological spaces and $y_0 \in Y$.

- *Then the constant function $f : x \mapsto y_0$ is continuous.*

Proposition 13.4 (continuity of the identity mapping).
Let $(X, \mathfrak{U})$be a topological space and let

$$id_X : X \to X; \qquad x \mapsto x$$

be the identity function on X. Then id_X is continuous.

Proposition 13.5. *Let d be the standard Euclidean metric and let d' be the discrete metric on the set $\mathbb{R}$ of all real numbers. Let*

$$f : (\mathbb{R}, d') \to (\mathbb{R}, d); \quad x \mapsto x \qquad \text{and} \qquad g : (\mathbb{R}, d) \to (\mathbb{R}, d'); \quad x \mapsto x$$

both be the identity function on $\mathbb{R}$. Then,

- *f is continuous at every point of $\mathbb{R}$*
- *g is not continuous anywhere on $\mathbb{R}$.*

Remark 13.1.

(a) All statements about continuity proven for topological spaces are also true for the special case of metric spaces.

(b) One may assume for statements involving continuity of a function f between metric spaces (X, d) and (Y, d') or between topological spaces$(X, \mathfrak{U})$ and $(Y, \mathfrak{V})$ that f is defined on all of X rather than assuming more generally that f is defined (only) on some arbitrary subset A of X.
The general case of $f : A \to Y$ is then covered for metric spaces by replacing (X, d) with $(A, d\big|_{A \times A})$ (we deal with $f : (A, d\big|_{A \times A}) \to (Y, d')$), and it is covered for topological spaces by replacing $(X, \mathfrak{U})$ with $(A, \mathfrak{U}_A)$ (we deal with $f : (A, \mathfrak{U}_A) \to (Y, \mathfrak{V})$), just as long as the proof does not make use of a property of X which its subset A does not satisfy. □

13.1.2 Uniform Continuity

Definition 13.5 (Uniform continuity of functions)**.** Let (X, d_1), (Y, d_2) be metric spaces and let A be a subset of X. A function

$$f(\cdot) : A \to Y \qquad \text{is called } \textbf{uniformly continuous}$$

if, for any $\varepsilon > 0$, there exists a (possibly very small) $\delta > 0$ such that

$$d_2(f(x) - f(y)) \ < \varepsilon \quad \text{for any } x, y \in A \text{ such that } d_1(x, y) \ < \ \delta. \tag{13.10}$$ □

13.1.3 Continuity of Linear Functions

Lemma 13.1. *Let $f : (V, \|\cdot\|) \to (W, \|\cdot\|)$ be a linear function between two normed vector spaces. Let*

$$\begin{aligned} a &:= \sup\{ \| f(x) \| : x \in V, \|x\| = 1\}, \\ b &:= \sup\{ \| f(x) \| : x \in V, \|x\| \leq 1\}, \\ c &:= \sup\{ \frac{\| f(x) \|}{\|x\|} : x \in V, x \neq 0\}. \end{aligned}$$

Then, $a \ = \ b \ = \ c$.

Definition 13.6 (norm of linear functions)**.** ★

Let $f : (V, \|\cdot\|) \to (W, \|\cdot\|)$ be a linear function between two normed vector spaces. We denote the quantity $a = b = c$ from lemma 13.1 by $\|f\|$, i.e.,

$$\begin{aligned} \|f\| &= \sup\{ \| f(x) \| : x \in V, \|x\| = 1\} \\ &= \sup\{ \| f(x) \| : x \in V, \|x\| \leq 1\} \\ &= \sup\{ \frac{\| f(x) \|}{\|x\|} : x \in V, x \neq 0\}. \end{aligned} \tag{13.11}$$

$\|f\|$ is called the **norm of the linear function** f.

□

Theorem 13.4 (Continuity criterion for linear functions)**.**
Let $f : (V, \|\cdot\|) \to (W, \|\cdot\|)$ be a linear function between two normed vector spaces. Then the following are equivalent.

(A) f *is continuous at* $x = 0$,
(B) f *is continuous in all points of* V,
(C) f *is uniformly continuous on* V,
(D) $\|f\| < \infty$.

Moreover, such a continuous linear function satisfies the inequality

(13.12) $$\left| f(x) \right| \leq \|f\| \cdot \|x\|, \quad \text{for all } x \in V.$$

Theorem 13.5 ($\|f\|$ is a norm)**.** ★ *Let*

(13.13) $$\mathscr{C}_{lin}(V,W) := \mathscr{C}_{lin}\big((V,\|\cdot\|),(W, \,|\cdot|\,)\big) := \{f : V \to W \; : f \text{ is linear and continuous}\}.$$

Then, $\mathscr{C}_{lin}(V,W)$ *is a vector space and*

(13.14) $$f \mapsto \|f\| = \sup\{\left| f(x) \right| : \|x\| = 1\}$$

defines a norm on $\mathscr{C}_{lin}(V,W)$.

13.2 Function Sequences and Infinite Series

13.2.1 Convergence of Function Sequences

Definition 13.7 (Pointwise convergence of function sequences)**.** Let X be a nonempty set, (Y,d) a metric space and let $f_n(\cdot) : X \to Y$ and $f(\cdot) : X \to Y$ be functions on X $(n \in \mathbb{N})$. Let $A \subseteq X$ be a nonempty subset of X.

We say that $f_n(\cdot)$ **converges pointwise** or, simply, **converges** to $f(\cdot)$ on A and we write $f_n(\cdot) \to f(\cdot)$ on A as $n \to \infty$, or simply $f_n(\cdot) \to f(\cdot)$ on A, if

(13.15) $$f_n(x) \to f(x) \text{ as } n \to \infty \text{ for all } x \in A.$$

We omit the phrase “on A” if it is clear how A is defined, in particular if $A = X$. □

Definition 13.8 (Uniform convergence of function sequences)**.** Let X be a nonempty set, (Y,d) a metric space, let $f_n(\cdot) : X \to Y$ and $f(\cdot) : X \to Y$ be functions on X $(n \in \mathbb{N})$, and let $A \subseteq X$.

We say that $f_n(\cdot)$ **converges uniformly** to $f(\cdot)$ on A and we write

(13.16) $$f_n(\cdot) \xrightarrow{uc} f(\cdot) \text{ on } A$$[21]

if, for each $\varepsilon > 0$ (no matter how small), there exists an index n_0 which can be chosen once and for all, <u>independently of the specific argument x</u>, such that

(13.17) $$d(f_n(x), f(x)) < \varepsilon \quad \text{for all } x \in A \quad \text{and } n \geq n_0.$$

We omit the phrase “on A” if it is clear how A is defined, in particular if $A = X$. □

Proposition 13.6 (Uniform convergence is $\|\cdot\|_\infty$ convergence)**.**
The following is true for any a nonempty set X and $f_n, f \in \mathscr{B}(X,\mathbb{R})$:

$$f_n(\cdot) \overset{uc}{\to} f(\cdot) \quad \Leftrightarrow \quad f_n(\cdot) \overset{\|\cdot\|_\infty}{\to} f(\cdot), \quad \text{i.e.,}$$
$$f_n(\cdot) \overset{uc}{\to} f(\cdot) \quad \Leftrightarrow \quad f_n \text{ converges to } f \text{ in the metric space } \big(\mathscr{B}(X,\mathbb{R}), d_{\|\cdot\|_\infty}(\cdot,\cdot)\big)\,.$$

Definition 13.9 (Norm and metric of uniform convergence). ⋆
We also call the sup–norm on $\mathscr{B}(X,\mathbb{R})$ the **norm of uniform convergence** on X and its associated metric $d_{\|\cdot\|_\infty}(\cdot,\cdot)$ the **metric of uniform convergence** on X. □

Theorem 13.6 (Uniform limits of continuous functions are continuous)**.**
Let (X,d_1) and (Y,d_2) be metric spaces and let $f_n(\cdot) : X \to Y$ and $f(\cdot) : X \to Y$ be functions on X $(n \in \mathbb{N})$. Let $x_0 \in X$ and let $V \subseteq X$ be a neighborhood of x_0. Assume the following:
- ***(a)*** *The functions $f_n(\cdot)$ are continuous at x_0 for all n.*
- ***(b)*** *$f_n(\cdot) \overset{uc}{\to} f(\cdot)$ on V.*

Then, f is continuous at x_0

Proposition 13.7. ⋆ *Let $f : [0,1] \to \mathbb{R}$ be one of the functions*

$$1 : x \mapsto 1; \qquad id : x \mapsto x; \qquad id^2 : x \mapsto x^2; \qquad (0 \leq x \leq 1).$$

Then,
$$B_n^f(\cdot) \overset{uc}{\to} f(\cdot) \text{ on } [0,1] \text{ as } n \to \infty.$$

Proposition 13.8. *Let X be a nonempty set, (Y,d) a metric space and let $f_n, f : X \to Y$ $(n \in \mathbb{N})$. Then*

f is the uniform limit of the function sequence $(f_n)_n$
*$\Leftrightarrow$ there exists a sequence $\delta_n \geq 0$ such that **1)** $\delta_n \to 0$ as $n \to \infty$, and **2)** $d\big(f_n(x), f(x)\big) \leq \delta_n$ for all $x \in X$ and $n \in \mathbb{N}$.*

13.2.2 Infinite Series

[21]Note that the notation " $f_n(\cdot) \overset{uc}{\to} f(\cdot)$ " is not very widely used.

Proposition 13.9 (Convergence criteria for series)**.**
A series $s := \sum a_k$ *of real numbers converges if and only if for all* $\varepsilon > 0$ *there exists* $n_o \in \mathbb{N}$ *such that one of the following is true:*

(13.18a) $$\left|\sum_{k=n}^{\infty} a_k\right| < \varepsilon \qquad \text{for all } n \geq n_0$$

(13.18b) $$\left|\sum_{k=n}^{m} a_k\right| < \varepsilon \qquad \text{for all } m, n \geq n_0$$

Corollary 13.1. *If a series* $\sum a_j$ *converges then* $\lim\limits_{n\to\infty} a_n = 0$.

Corollary 13.2 (Dominance criterion for series)**.**
Let $N \in \mathbb{N}$ *and let* $\sum a_j$ *and* $\sum b_j$ *be two series such that* $|b_k| \leq a_k$ *for all* $k \geq N$.
It follows that if $\sum a_k$ *converges, then* $\sum b_k$ *converges.*
Moreover, if $|b_k| \leq a_k$ *for all* $k \in \mathbb{N}$*, then,* $\left|\sum_{k=1}^{\infty} b_j\right| \leq \sum_{k=1}^{\infty} a_j$

Definition 13.10 (Finite permutations). ★ Let $N \in \mathbb{N}$. A **permutation** of $[N]$ is a bijection

$$\pi(\cdot) : [N] \to [N]; \qquad j \mapsto \pi(j)\,.$$

As usual

$$\pi^{-1}(\cdot) : [N] \to [N]; \qquad \pi(j) \mapsto \pi^{-1}\pi(j) = j,$$

denotes the inverse function of $\pi(\cdot)$. We recall that it associates with each image $\pi(j)$ the unique argument j, which is mapped by $\pi(\cdot)$ to $\pi(j)$. It is customary to write

i_1 instead of $\pi(1)$, i_2 instead of $\pi(2)$, ..., i_j instead of $\pi(j)$, ... □

Definition 13.11 (Permutations of $\mathbb{N}$)**.** A **permutation** of $\mathbb{N}$ is a bijective function

$$\pi(\cdot) : \mathbb{N} \to \mathbb{N}; \qquad j \mapsto \pi(j).$$ □

Proposition 13.10. *Let (a_n) be a sequence of nonnegative real numbers. Exactly one of the following is true:*

(a) *Either the series $\sum a_n$ converges (to a finite number). In that case,*

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\pi(n)} \quad \textit{for } \textbf{\textit{any}} \textit{ permutation } \pi(\cdot) \textit{ of } \mathbb{N}.$$

(b) *Or the series $\sum\limits_{n=1}^{\infty} a_n$ has limit ∞. In that case, it is true for **any** permutation $\pi(\cdot)$ of $\mathbb{N}$, that the reordered series $\sum\limits_{n=1}^{\infty} a_{\pi(n)}$ also has limit ∞.*

Definition 13.12 (absolutely convergent series)**.** A series $\sum a_j$ is **absolutely convergent**, if the corresponding series $\sum |a_j|$ of its absolute values converges. ☐

Proposition 13.11. *Let $\sum a_k$ be an absolutely convergent series. Then $\sum a_k$ converges and*

$$|\sum_{k=1}^{\infty} a_k| \leq \sum_{k=1}^{\infty} |a_k|. \tag{13.19}$$

Theorem 13.7. *Let $\sum a_k$ be an absolutely convergent series. Let $\pi : \mathbb{N} \to \mathbb{N}$ be a permutation of $\mathbb{N}$, i.e., the series $\sum b_k$ with $b_k := a_{\pi(k)}$ is a rearrangement of the series $\sum a_k$. Then $\sum b_k$ converges and has the same limit as $\sum a_k$. (Note that $\sum a_k$ converges according to Proposition 13.11.)*

Proposition 13.12. *Let $\sum a_n$ be an absolutely convergent series. Let $(a_{n_k})_k$ be a subsequence of $(a_n)_n$. Then, $\sum a_{n_k}$ converges absolutely.*

Remark 13.2. Assume that $\sum a_n$ is absolutely convergent. Let $n_1 < n_2 < \dots$ be a subsequence of all natural numbers and let $J := \{n_j : j \in \mathbb{N}\}$.

- Then we write $\sum\limits_{j \in J} a_{n_j} := \sum\limits_{j=1}^{\infty} a_{n_j}$.
- In particular, we write $\sum\limits_{j \in \mathbb{N}} a_j := \sum\limits_{j=1}^{\infty} a_j$, for the full sequence $n_j = j$ of indices. ☐

Definition 13.13 (conditionally convergent series)**.** A series $\sum a_j$ is called **conditionally convergent**, if it is convergent but not absolutely convergent. □

Definition 13.14 (Alternating Series)**.** ⋆

A series $\sum a_j$ is called an **alternating series** if it is of the form $\sum(-1)^j a_j$ with either all terms a_j being strictly positive or all of them being strictly negative. □

Proposition 13.13 (Leibniz Test for Alternating Series)**.**
Let $a_1 \geq a_2 \geq \ldots \downarrow 0$ be a nonincreasing sequence which decreases to zero.
Then, the alternating series $\sum(-1)^k a_k$ converges.

Theorem 13.8 (Riemann's Rearrangement Theorem)**.**
Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha \leq \beta$. and let the series $\sum a_k$ be conditionally convergent.
Then a rearrangement $\sum b_k$ of $\sum a_k$ exists such that

$$\liminf_{n\to\infty} \sum_{k=1}^{n} b_k = \alpha \qquad \text{and} \qquad \limsup_{n\to\infty} \sum_{k=1}^{n} b_k = \beta\,.$$

Corollary 13.3. *Let the series $\sum a_k$ be conditionally convergent and let $\alpha \in \mathbb{R}$.*
Then, a rearrangement $\sum b_k$ of $\sum a_k$ exists such that

$$\lim_{n\to\infty} \sum_{k=1}^{n} b_k = \alpha.$$

Corollary 13.4. *Let $\sum a_k$ be a convergent series with limit $\alpha \in \mathbb{R}$ such that $\sum b_k = \alpha$, for each rearrangement.*
Then $\sum a_k$ converge absolutely.

Corollary 13.5 (Dichotomy for convergent series)**.** *Let series $\sum a_k$ be a convergent series. Then*

(a) *either all rearrangements of $\sum a_k$ converge to the same limit,*
(b) *or, for any $\alpha \in \mathbb{R}$, there is a rearrangement of $\sum a_k$ which converges to α.*

14 Compactness

14.1 ε-Nets and Total Boundedness

Definition 14.1 (ε–nets)**.** Let $\varepsilon > 0$. Let (X, d) be a metric space and $A \subseteq X$. let $G \subseteq X$ be a subset of X with the following property:

(14.1) For each $x \in A$ there exists $g \in G$ such that $x \in N_\varepsilon(g)$, i.e., $\bigcup\limits_{g \in G} N_\varepsilon(g) \supseteq A$.

In other words, the points of G form a "grid" or "net" fine enough so that no matter what point x of A you choose, you can always find a "grid point" g with distance less than ε to x, because that is precisely the meaning of $x \in N_\varepsilon(g)$.

We call G an ε**–net** or ε**–grid** for A and we call $g \in G$ a **grid point** of the net. □

Definition 14.2 (Total boundedness)**.** Let (X, d) be a metric space and let A be a subset of X. We say that A is **totally bounded** if, for each $\varepsilon > 0$, there exists a finite(!) ε–grid for A. □

Proposition 14.1 (ε–nets in $\mathbb{R}^n$)**.** ⋆ *Let (X, d) be $\mathbb{R}^n$ with the Euclidean metric.*

(A) *Let $\varepsilon > 0$. Then the set*

$$\varepsilon\mathbb{Z}^n = \{\varepsilon\vec{z} : \vec{z} \in \mathbb{Z}^n\} = \{(\varepsilon z_1, \dots, \varepsilon z_n) : z_j \in \mathbb{Z} \text{ for } j = 1, \dots, n\}$$

is an $(\varepsilon\sqrt{n})$–net for (any subset of) $\mathbb{R}^n$.

(B) *Let A be a bounded set in $\mathbb{R}^n$ and $\varepsilon > 0$. Then there is $k \in \mathbb{N}$ and $g_1, \dots g_k \in \varepsilon\mathbb{Z}^n$ such that*

$$A \subseteq N_\varepsilon(g_1) \cup N_\varepsilon(g_2) \cup \dots, \cup N_\varepsilon(g_k),$$

i.e., A is covered by finitely many ε–neighborhoods of points in the $(\varepsilon/\sqrt{n})$–grid $\varepsilon\mathbb{Z}^n$.

Theorem 14.1. *Bounded subsets of $\mathbb{R}^n$ are totally bounded.*

Proposition 14.2. *Let $A \subseteq \mathbb{R}^n$ be bounded and let $(x_n)_n$ be a sequence such that $x_n \in A$ for all n. Then there exists a subsequence x_{n_j} which is Cauchy.*

Theorem 14.2. *Let A be a totally bounded subset of a metric space (X, d). Let $(x_n)_n$ be a sequence such that $x_n \in A$ for all n. Then there exists a subsequence x_{n_j} which is Cauchy.*

Proposition 14.3. *Totally bounded subsets of metric spaces are bounded.*

Corollary 14.1. *If $A \subseteq \mathbb{R}^n$, then* $\quad$ *A is bounded* $\Leftrightarrow$ *A is totally bounded.*

Theorem 14.3. *Let A be a subset of a metric space (X, d) such that each sequence in A contains a Cauchy subsequence. Then A is totally bounded.*

Corollary 14.2. *Let A be a subset of a metric space (X, d). Then,*

A is totally bounded $\Leftrightarrow$ *every sequence in A possesses a Cauchy subsequence.*

14.2 Sequence Compactness

Definition 14.3 (Sequence compactness)**.** Let (X, d) be a metric space and let $A \subseteq X$. We say that A is **sequence compact** or **sequentially compact** if it has the following property: Given any sequence (a_n) of elements of A, there exists $a \in A$ and a subset

$$n_1 < n_2 < \ldots < n_j < \ldots \quad \text{of indices such that} \quad a = \lim_{j \to \infty} a_{n_j},$$

In other words, there exists a subsequence [22](a_{n_j}) which converges to a. ☐

Proposition 14.4 (Sequence compactness implies total boundedness)**.**
Let (X, d) be a metric space and let A be a sequentially compact subset of X.
Then A is totally bounded.

Proposition 14.5 (Sequence compact implies completeness)**.**
Let (X, d) be a metric space and let A be a sequence compact subset of X.
Then A is complete, i.e., any Cauchy sequence (x_{n_j}) in A converges to a limit $L \in A$.

Theorem 14.4 (Sequence compact $\Leftrightarrow$ totally bounded and complete)**.**
Let A be a subset of a metric space (X, d).
Then, A is sequence compact if and only if A is totally bounded and complete.

[22]See Definition 5.22 on p.48.

Theorem 14.5 (Sequence compact sets are closed and bounded)**.**
Let A be sequence compact subset of a metric space (X, d). Then A is a bounded and closed set.

A subset of a metric space is sequentially compact
$\Leftrightarrow$ it is totally bounded and complete
$\Rightarrow$ it is bounded and closed. $\square$

Theorem 14.6. *A subset of $\mathbb{R}^n$ is sequentially compact*
$\Leftrightarrow$ it is totally bounded and complete
$\Leftrightarrow$ it is bounded and closed.

14.3 Open Coverings and the Heine–Borel Theorem

Definition 14.4 (Coverings and open coverings)**.** Let X be an arbitrary nonempty set and $A \subseteq X$. Let $U_i \in X$ $(i \in I)$ such that $A \subseteq \bigcup_{i \in I} U_i$. We call such a family a **covering** of A.

A **finite subcovering** of a covering $(U_i)_{i \in I}$ of the set A is a finite collection

$$U_{i_1}, \dots, U_{i_n} \quad (i_j \in I \quad \text{for } 1 \leq j \leq n) \quad \text{such that} \quad A \subseteq U_{i_1} \cup U_{i_2} \cup \dots \cup U_{i_n}. \tag{14.2}$$

Assume in addition that X is a topological space, e.g., a normed vector space or a metric space. If all members U_i are open then we call $(U_i)_{i \in I}$ an **open covering** of A.

We also write **cover**, **finite subcover**, **open cover** instead of covering, finite subcovering, open covering $\square$

Definition 14.5 (Compact sets)**.** Let $(X, \mathfrak{U})$ be a topological space and $K \subseteq X$.

- We call K **compact**, if K possesses the **"extract finite open subcovering" property**:

Given any **open** covering $(U_i)_{i \in I}$ of K, one can extract a finite subcovering. In other words, there is $n \in \mathbb{N}$ and indices

$$i_1, i_2, \dots, i_n \in I \qquad \text{such that} \quad A \subseteq \bigcup_{j=1}^{n} U_{i_j}. \ \square$$

Example 14.1. Here are some simple examples.

(a) Any finite topological space is compact.
(b) Any topological space that only contains finitely many open sets is compact. In particular a set with the indiscrete topology [23] is compact
(c) A space with the discrete metric [24] is compact if and only if it is finite.

[23] See Definition 12.14 on p.122
[24] See Definition 12.3 on p.117

Theorem 14.7 (Compact metric spaces are sequence compact).
Let (X, d) be a compact metric space. Then X is sequence compact.

Proposition 14.6. *Let (X, d) be a sequence compact metric space. Let $(U_i)_{i \in I}$ be an open cover of X. Then, one can find for $(U_i)_{i \in I}$ a number $\rho > 0$ which possesses the following property:*

- *For each $x \in X$ there exists $i \in I$ such that $N_\rho(x) \subseteq U_i$.*

Theorem 14.8. *Sequence compact metric spaces are compact.*

Theorem 14.9 (Sequence compactness coincides with compactness in metric spaces).
Let (X, d) be a metric space and let A be a subset of X. Then,

A is sequence compact $\Leftrightarrow$ A is compact, i.e.,
A is sequence compact $\Leftrightarrow$ every open cover of A possesses a finite subcover.

Theorem 14.10 (Heine–Borel Theorem).
A subset of Euclidean space $\mathbb{R}^n$ is compact $\Leftrightarrow$ this set is closed and bounded.

14.4 Continuous Functions and Compact Spaces

Theorem 14.11 (Closed subsets of compact topological spaces are compact).
Let A be a closed subset of a compact topological space $(X, \mathfrak{U})$. Then A is a compact subspace.

Corollary 14.3 (Closed subsets of compact metric spaces are compact).
Let A be a closed subset of a compact metric space (X, d). Then $(A, d\big|_{A \times A})$ is a compact subspace.

Theorem 14.12 (Continuous images of compact topological spaces are compact).
Let $(X, \mathfrak{U})$ and $(Y, \mathfrak{V})$ be two topological spaces. and let $f : X \to Y$ be continuous on X.

- *If X is compact then the direct image $f(X)$ is compact.*

In other words, the topological subspace $(f(X), \mathfrak{V}_{f(X)})$ of Y is compact.

Corollary 14.4 (Continuous images of compact metric spaces are compact)**.**
Let (X, d_1) and (Y, d_2) be two metric spaces. and let $f : X \to Y$ be continuous on X.
If X is compact, then its image $f(X)$ is compact, i.e., the metric subspace $(f(X), d_2)$ of Y is compact.

Corollary 14.5. *Let $(X, \mathfrak{U})$ be a topological space and (Y, d) a metric space.*

- *If X is compact and $f : X \to Y$ is continuous, then f is bounded.*
- *In particular, any continuous function on a closed interval of real numbers is bounded.*

Corollary 14.6 (Continuous real–valued functions attain max and min on a compact domain)**.**

Let $(X, \mathfrak{U})$ be a topological space, $A \subseteq X$ a compact subspace and $f : A \to \mathbb{R}$ continuous on A.
Then there exist $x_\star, x^\star \in A$ such that

$$f(x_\star) = \min_{x \in A} f(x) \quad \text{and} \quad f(x^\star) = \max_{x \in A} f(x).$$

Theorem 14.13 (Uniform continuity on sequence compact spaces)**.**
Let (X, d_1), (Y, d_2) be metric spaces and let A be a compact subset of X. Then,

- *any continuous function $A \to Y$ is uniformly continuous on A.*

Corollary 14.7 (Uniform continuity on closed intervals)**.** *Let $a, b \in \mathbb{R}$ such that $a \leq b$.*
Any continuous real–valued function on the closed interval $[a, b]$ is uniformly continuous:
For any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon \quad \text{for all } x, y \in [a, b] \text{ such that } |x - y)| < \delta \tag{14.3}$$

15 Applications of Zorn's Lemma

15.1 More on Partially Ordered Sets

Definition 15.1. Let $(X, \preceq)$ be a POset (partially ordered set), $A \subseteq X$, and $m \in A$.

- m is called **maximal** for A iff there is no $a \in A$ such that $a \neq m$ and $m \preceq a$. m is called a **maximum** of A if $a \in A$ and $a \preceq m$ for all $a \in A$.
- m is called **minimal** for A iff there is no $a \in A$ such that $a \neq m$ and $m \succeq a$. m is called a **minimum** of A if $a \in A$ and $a \succeq m$ for all $a \in A$.

Proposition 15.1 below shows that such a maximum or minimum is unique. Thus, we may write $\max(A)$ for the maximum of A and $\min(A)$ for the minimum of A. ☐

Proposition 15.1.
Let $(X, \preceq)$ be a nonempty POset and $A \subseteq X$. If A has a maximum then it is unique.

Note 15.1 (Notes on maximal elements and maxima)**.**

(a) If $(X, \preceq)$ is not linearly ordered, then its subsets may have many maximal elements. For example, for the trivial partial ordering $x \preceq y$ if and only if $x = y$, every element is maximal. A maximum is a maximal element, but the converse is often not true.

(b) If an ordering is not specified, then we always mean set inclusion.

(c) Let $A \subseteq X$. If $m \in A$ is a maximum of A then this implies that m must be related to all other elements of A. ☐

Axiom 15.1 (Zorn's Lemma)**. Zorn's Lemma**: Let $(X, \preceq)$ be a partially ordered set with the **ZL property**:

$$\text{Every chain } C \subseteq X, \text{ possesses an upper bound } u \in X, \text{ i.e., } c \preceq u \text{ for all } c \in C. \qquad \textbf{(ZL)}$$

Then X has a maximal element. ☐

15.2 Existence of Bases in Vector Spaces

For the remainder of this chapter we assume that V is a vector space and define

$$\mathfrak{B} := \{A \subseteq V : A \text{ is linearly independent}\}. \qquad (15.1)$$

Lemma 15.1. *Every chain $\mathfrak{C}$ in $(\mathfrak{B}, \subseteq)$ possesses an upper bound.*

Theorem 15.1. *Every vector space V has a basis.*

15.3 The Cardinal Numbers are a totally ordered set

Theorem 15.2. *Let $X, Y \subseteq \Omega$. Then $\textbf{card}(X) \leq \textbf{card}(Y)$ or $\textbf{card}(Y) \leq \textbf{card}(X)$*

15.4 Extensions of Linear Functions in Arbitrary Vector Spaces

Lemma 15.2. *Let V be a vector space and let F be a (linear) subspace of V. Let $f : F \to \mathbb{R}$ be linear. Let*

$$\begin{aligned}\mathscr{G} := \{(W, f_W) : & W \text{ is a subspace of } V,\ W \supseteq F, \\ & f_W : W \to \mathbb{R} \text{ is a linear extension of } f \text{ to } W\}.\end{aligned}$$

Then the following defines a partial ordering on $\mathscr{G}$:

$$(U, f_U) \preceq (W, f_W) \ \Leftrightarrow\ V \subseteq W \text{ and } f_W\big|_U = f_U .$$

Moreover this ordering satisfies the requirements of Zorn's Lemma:
Every chain in $(\mathscr{G}, \preceq)$ possesses an upper bound (in $\mathscr{G}$).

Theorem 15.3 (Extension theorem for linear real–valued functions)**.**
Let V be a vector space and let F be a (linear) subspace of V. Let $f : F \to \mathbb{R}$ be a linear mapping. Then there is an extension of f to a linear mapping $\tilde{f} : V \to \mathbb{R}$.

Definition 15.2 (Dual vector space)**.** ⋆
Let V and W be vector spaces, and let $L : V \to W$ be linear.

(a) We call $V^* := \{f : f \text{ is a linear function } V \to \mathbb{R}\}$ the **dual** or **algebraic dual** of V.
(b) We call $L^* : W^* \to V^*$, defined by $L^*(f) := f \circ L$, i.e., $L^*(f)(x) = f(Lx)\ \forall\, x \in V$, the **dual function** or **dual mapping** of L. □

Proposition 15.2. ⋆ *For the following see Definition 11.2 (Transposed matrix) on p.104.*

(a) *V^* is a vector space, i.e., $f, g, \in V^*$ and $\alpha, \beta \in \mathbb{R} \Rightarrow \alpha f + \beta g \in V^*$.*

(b) *Since V^* is a vector space, its dual $V^{**} := \left(V^*\right)^*$ exists.*

(c) *Assume that $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$. For every linear function $L : \mathbb{R}^n \to \mathbb{R}^m$ there exists a matrix $A = \left((a_{ij})\right)$ such that for every column vector $\vec{x}$, $L(\vec{x}) = A\vec{x}$, i.e., the function value $\vec{y} = L(\vec{x})$ has coordinates $y_i = \sum_{j=1^n} a_{ij} x_j$.*

(d) *If V is a finite dimensional vector space, then there is a bijection $V \to V^*$ which is linear in both directions.* [25] *This allows us to "identify" $\left(\mathbb{R}^n\right)^*$ with $\mathbb{R}^n$, thus, the dual function of L from Definition 15.2 is a linear function $L^* : \mathbb{R}^m \to \mathbb{R}^n$ (Careful: Switched dimensions!).*

*According to part **(c)** of this remark, there exists a matrix $A^* = \left((a_{k\ell})\right)$ such that the following is true. If $\vec{y}^* \in \mathbb{R}^m$ and $\vec{x}^* \in \mathbb{R}^n$ are column vectors such that $\vec{x}^* = L^*\left(\vec{y}^*\right)$, then $\vec{x}^* = A^*\vec{y}$, the product of the matrix A^* and the column vector $\vec{y}$.*

This matrix A^* is the transpose $A^\top$ of A: $a^*_{k\ell} = a_{\ell k}$ for $k = 1, \dots n$ and $\ell = 1, \dots, m$. ☐

Theorem 15.4. ⋆ *Let $L : V \to W$ be a linear function between two vector spaces V and W. Let $L^* : W^* \to V^*$ be the associated dual function of L. Then,*

$$L \text{ is injective} \quad \Leftrightarrow \quad L^* \text{ is surjective}\,; \qquad L^* \text{ is injective} \quad \Leftrightarrow \quad L \text{ is surjective}\,.$$

Corollary 15.1. *Let $A = \left((a_{ij})\right)$ be a matrix with m rows and n columns. Then **(a)** ⇔ **(b)**, where*

(a) *The set of m linear equations in n unknowns $\vec{x} = (x_1, \dots, x_n)^\top$,*

$$A\,\vec{x} \;=\; \vec{y}\,,$$

has a solution $\vec{x}$ for any choice of right hand side $\vec{y} = (y_1, \dots, y_m)^\top$.

(b) *the set of n linear equations in m unknowns $\vec{\xi} = (\xi_1, \dots, \xi_m)^\top$,*

$$A^\top\,\vec{\xi} \;=\; \vec{\eta}\,,$$

has at most one solution $\vec{\xi}$ for any $\vec{\eta} = (\eta_1, \dots, \eta_n)^\top$.

15.5 The Hahn-Banach Extension Theorem ⋆

15.5.1 Sublinear Functionals

Definition 15.3 (Sublinear functionals)**.** Let V be a vector space and $p : V \to \mathbb{R}$ such that

(a) if $\lambda \in \mathbb{R}_{\geq 0}$ and $x \in V$ then $p(\lambda x) \;=\; \lambda p(x)$ (positive homogeneity)

(b) if $x, y \in V$ then $p(x + y) \;\leq\; p(x) + p(y)$ (subadditivity)

Then we call p a **sublinear functional** on V. ☐

[25] One calls such bijection which is structure compatible a **linear isomorphism** or a **vector space isomorphism**.

Proposition 15.3. *Let V be a vector space and $p : V \to \mathbb{R}$ sublinear. Let $x \in V$. Then*

- ***(a)*** $p(0) = 0$,
- ***(b)*** $-p(x) \le p(-x)$,

Example 15.1 (Norms are sublinear)**.** Let $(V, \|, \cdot\|)$ be a normed vector space.
Then the function $p(x) := \|x\|$ is sublinear.

Example 15.2 (Linear functions are sublinear)**.**
Let V be a vector space and let $f := V \to \mathbb{R}$ be a linear function. Then f is sublinear.

15.5.2 The Hahn-Banach extension theorem and its Proof

Theorem 15.5 (Hahn–Banach extension theorem)**.**
Let V be a vector space and $p : V \to \mathbb{R}$ a sublinear function.
Suppose F is a (linear) subspace of V and $f : F \to \mathbb{R}$ is a linear mapping such that $f \le p$ on F.
Then there is an extension of f to a linear map $\tilde{f} : V \to \mathbb{R}$ such that $\tilde{f} \le p$ on V.

Theorem 15.6 (Continuous extensions of continuous linear functions)**.**
Let $(V, \| \cdot \|)$ be a normed vector space. Let F be a (linear) subspace of V. Then,

- *any continuous, linear $f : F \to \mathbb{R}$ possesses a continuous, linear extension $\tilde{f} : V \to \mathbb{R}$.*

15.6 Convexity ⋆

Definition 15.4 (Concave-up and convex functions)**.** Let $-\infty \le \alpha < \beta \le \infty$ and let $I :=]\alpha, \beta[$ be the open interval of real numbers with endpoints α and β. Let $f : I \to \mathbb{R}$.

- **(a)** The **epigraph** of f is the set $epi(f) := \{(x_1, x_2) \in I \times \mathbb{R} : x_2 \ge f(x_1)\}$ of all points in the plane that lie above the graph of f.
- **(b)** f is **convex** if for any two vectors $\vec{a}, \vec{b} \in epi(f)$ the entire line segment $S := \{\lambda\vec{a} + (1-\lambda)\vec{b}\} : 0 \le \lambda \le 1$ is contained in $epi(f)$. See Figure 15.1.
- **(c)** Let f be differentiable at all points $x \in I$. Then f is **concave-up**, if the function $f' : x \mapsto f'(x) = \frac{df}{dx}(x)$ is increasing. □

Proposition 15.4 (Convexity criterion)**.** *f is convex if and only if the following is true: For any*

$$\alpha < a \le x_0 \le b < \beta$$

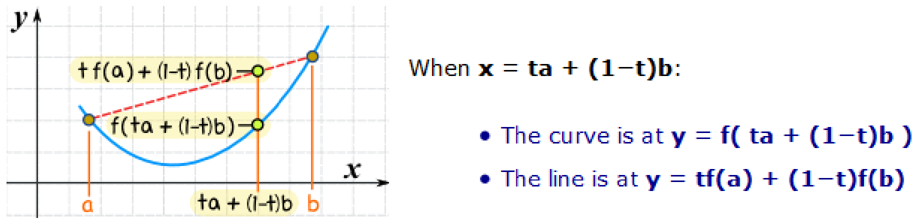


Figure 15.1: Convex function
26

let $S(x_0)$ be the unique number such that the point $(x_0, S(x_0))$ is on the line segment that connects the points $(a, f(a))$ and $(b, f(b))$. Then

(15.2) $$f(x_0) \leq S(x_0).$$

Note that any x_0 between a and b can be written as $x_0 = \lambda a + (1-\lambda)b$ for some $0 \leq \lambda \leq 1$ and that the corresponding y-coordinate $S(x_0) = S(\lambda a + (1-\lambda)b)$ on the line segment that connects $(a, f(a))$ and $(b, f(b))$ then is $S(\lambda a + (1-\lambda)b) = \lambda f(a) + (1-\lambda)f(b)$. Hence we can rephrase the above as follows:
f is convex if and only if for any $a < b$ such that $a, b \in I$ and $0 \leq \lambda \leq 1$ it is true that

(15.3) $$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b).$$

Proposition 15.5 (Convex vs concave-up). *Let $f : \mathbb{R} \to \mathbb{R}$ be concave-up. Then f is convex.*

Proposition 15.6 (Sublinear functions are convex). *Let $f : \mathbb{R} \to \mathbb{R}$ be sublinear. Then f is convex.*

16 Approximation theorems ★

16.1 The Positive, Linear Operators $f \mapsto B_n^f$

Definition 16.1 (Positive linear operators)**.** Let (X, d) be a metric space, and let $\mathscr{F}$ be a subspace of the vector space $\mathscr{F}(X, \mathbb{R})$, i.e., with any two functions $f(\cdot), g(\cdot) \in \mathscr{F}$ their sum $f + g$ also belongs to $\mathscr{F}$ and that the function λf ($\lambda \in \mathbb{R}$) also belongs to $\mathscr{F}$.
A ***linear operator*** T on $\mathscr{F}$ is a linear function [27] $T : \mathscr{F} \to \mathscr{F}$
A ***positive linear operator*** T on $\mathscr{F}$ is a linear operator on $\mathscr{F}$ with the following property:

(16.1) $$f \geq 0 \quad \Rightarrow \quad Tf \geq 0, \quad \text{i.e.,} \quad f(x) \geq 0 \text{ for all } x \in X \quad \Rightarrow \quad Tf(x) \geq 0 \text{ for all } x \in X.$$

Proposition 16.1 (Properties of positive linear operators)**.**
Let T be a positive linear operator on a subspace $\mathscr{F}$ of $\mathscr{F}(X, \mathbb{R})$. Then,

(a) *T is **monotone increasing**, i.e., for any two functions $f, g \in \mathscr{F}$ such that $f \leq g$ it is true that $T(f) \leq T(g)$. In other words,*

(16.2) $$f(x) \leq g(x) \text{ for all } x \in X \quad \Rightarrow \quad T(f)(x) \leq T(g)(x) \text{ for all } x \in X.$$

(b) *Assume that $T(|f|)$ is defined for a function $f \in \mathscr{F}$. Then $|T(f)| \leq T(|f|)$. In other words,*

(16.3) $$|T(f)(x)| \leq T(|f|)(x) \quad \text{for all } x \in X.$$

Proposition 16.2 (Linearity and positivity of Bernstein polynomial assignments)**.**

(a) *Let $f(\cdot), g(\cdot)$ be two real–valued functions on $[0, 1]$ and $\alpha, \beta \in \mathbb{R}$. Let $h : [0, 1] \to \mathbb{R}$ be defined as*

$$h := \alpha f + \beta g, \text{ i.e., } h(x) = \alpha f(x) + \beta g(x) \ (0 \leq x \leq 1).$$
$$\text{Then } B_n^h = \alpha B_n^f + \beta B_n^g, \text{ i.e., } B_n^h(x) = \alpha B_n^f(x) + \beta B_n^g(x) \ (x \in \mathbb{R}).$$

To express this more succinctly:

(16.4) $$B_n^{\alpha f + \beta g} = \alpha B_n^f + \beta B_n^g.$$

(b) *Let f be a real–valued function on $[0, 1]$ which is nonnegative, i.e., $f(x) \geq 0$ for $0 \leq x \leq 1$. Then $B_n^f(\cdot) \geq 0$ on $[0, 1]$ (but not necessarily for $x \notin [0, 1]$).*

Corollary 16.1. *Let $n \in \mathbb{N}$. Then $B_n(\cdot)$ is a positive linear operator on $\mathscr{C}([0, 1], \mathbb{R})$.*

16.2 Korovkin's First Theorem

Unless stated differently we assume the following for all of this subchapter:
a and b are two real numbers such that $a < b$, and

$$T_n(\cdot) : \mathscr{C}([a,b],\mathbb{R}) \to \mathscr{C}([a,b],\mathbb{R}); \qquad f(\cdot) \mapsto T_n^f(\cdot) = T_n(f)(\cdot)$$

is a sequence of positive linear operators on $\mathscr{C}([a,b],\mathbb{R})$. This means in particular that for each continuous real–valued function $f(\cdot)$ on $[a,b]$ the image

$$T_n^f : x \mapsto T_n^f(x)$$

is itself a continuous, real–valued function on $[a,b]$.

Theorem 16.1 (Korovkin's First Theorem). *Assume that we have uniform convergence*

$$T_n^f(\cdot) \xrightarrow{uc} f(\cdot),$$

for the following three elements f of $\mathscr{C}([a,b],\mathbb{R})$:

$$\begin{aligned} 1(\cdot) :& x \mapsto 1 && \text{the constant function } 1, \\ id(\cdot) :& x \mapsto x && \text{the identity on } [a,b], \\ id^2(\cdot) :& x \mapsto x^2. \end{aligned}$$

Then $T_n^f \xrightarrow{uc} f$ for all $f \in \mathscr{C}([a,b],\mathbb{R})$.

16.3 The Weierstrass Approximation Theorem

Proposition 16.3 (Weierstrass Approximation Theorem on $[0,1]$). *Any continuous real–valued function on the unit interval $[0,1]$ can be uniformly approximated by a sequence of polynomials.*

Lemma 16.1. *Let $n \in [0,\infty[_{\mathbb{Z}}$, $\alpha_j, m, b \in \mathbb{R}$, $\alpha_n \neq 0$.*

Let $p : \mathbb{R} \to \mathbb{R}$ be a polynomial $p(x) = \sum\limits_{j=0}^{n} \alpha_j x^j$, and let $\varphi : \mathbb{R} \to \mathbb{R}$ be defined as $\varphi(x) = mx + b$.

Then $p \circ \varphi : \mathbb{R} \to \mathbb{R}$; $x \mapsto \sum\limits_{j=0}^{n} \alpha_j (mx+b)^j$ is a polynomial.

Proposition 16.4. *Let $A \subseteq \mathbb{R}$, $\varphi : \mathbb{R} \to \mathbb{R}$ defined as $\varphi(x) = mx + b$ $(m, b \in \mathbb{R})$. Further, let $f_n, f \in \mathscr{C}(\varphi(A),\mathbb{R})$, $n \in \mathbb{N}$. (Thus, f_n and f are continuous functions on $\varphi(A) = \{\varphi(x) : x \in A\}$.) Assume further that $f_n \xrightarrow{uc} f$ on $\varphi(A)$. Then $f_n \circ \varphi \xrightarrow{uc} f \circ \varphi$, on A.*

Corollary 16.2. *Let* $a, b \in \mathbb{R}$ *such that* $a < b$. *Let* $\varphi : [a, b] \to [0, 1]$; $\varphi(x) := \dfrac{x-a}{b-a}$. *Then*

(a) φ *is a bijection* $[a, b] \xrightarrow{\sim} [0, 1]$.

(b) *If* $h_n, h \in \mathscr{C}([0, 1], \mathbb{R})$ $(n \in \mathbb{N})$ *such that* $h_n \xrightarrow{uc} h$ *on* $[0, 1]$, *then* $h_n \circ \varphi \xrightarrow{uc} h \circ \varphi$ *on* $[a, b]$.

Theorem 16.2 (Weierstrass Approximation Theorem)**.** *Let* $a, b \in \mathbb{R}$ *such that* $a < b$. *Then any continuous real–valued function on* $[a, b]$ *can be uniformly approximated by a sequence of polynomials.*

17 Construction of the Number Systems ★

17.1 The Peano Axioms

Definition 17.1 (Set of nonnegative integers)**.**
We define the set $\mathbb{N}_0$ (the nonnegative integers) axiomatically as follows:

Ax.1 There is an element "0" contained in $\mathbb{N}_0$.
Ax.2 There is a function $\sigma : \mathbb{N}_0 \to \mathbb{N}_0$ such that
- **Ax.2.1** σ is injective,
- **Ax.2.2** $0 \notin \sigma(\mathbb{N}_0)$ (range of σ),
- **Ax.2.3** Induction axiom: Let $U \subseteq \sigma(\mathbb{N}_0)$ such that **(a)** $0 \in U$, **(b)** If $n \in U$ then $\sigma(n) \in U$. It then follows that $U = \mathbb{N}_0$.

We define $\mathbb{N} := \mathbb{N}_0 \setminus \{0\}$. ☐

Definition 17.2 (Iterative function composition)**.** Let $X \neq \emptyset$ and $f : X \to X$.
We use the induction axiom above to define f^n for an arbitrary function $f : X \to X$:
(a) $f^0 := \mathrm{id}_X : x \mapsto x$, **(b)** $f^1 := f$, **(c)** $f^2 := f \circ f$ (function composition), **(c)** $f^{\sigma(n)} := f \circ f^n$. ☐

Proposition 17.1. *f^n is defined for all $n \in \mathbb{N}_0$.*

Definition 17.3 (Addition and multiplication on $\mathbb{N}_0$)**.** Let $m, n \in \mathbb{N}_0$. Let

$$m + n := \sigma^n(m), \tag{17.1}$$
$$m \cdot n := (\sigma^m)^n(0). \tag{17.2}$$
☐

Proposition 17.2. *Addition and multiplication satisfy all commonly known rules of arithmetic, such as*

$m + n = n + m$	*commutativity of addition*
$k + (m + n) = (k + m) + n$	*associativity of addition*
$m \cdot n = n \cdot m$	*commutativity of multiplication*
$k \cdot (m \cdot n) = (k \cdot m) \cdot n$	*associativity of multiplication*
$k \cdot (m + n) = k \cdot m + k \cdot n$	*distributivity of addition*
$n \cdot 1 = 1 \cdot n = n$	*neutral element for multiplication*

Here, 1 is defined as $1 = \sigma(0)$.

Definition 17.4 (Order relation $m < n$ on $\mathbb{N}_0$)**.** Let $m, n \in \mathbb{N}_0$.

(a) We say m is less than n and we write $m < n$, if there exists $x \in \mathbb{N}$ such that $n = m + x$.
(b) We say m is less or equal than n and we write $m \leq n$, if $m < n$ or $m = n$.
(c) We say that m is greater than n and we write $m > n$, if $n < m$.
We say m is greater or equal than n and we write $m \geq n$, if $n \leq m$. ☐

Proposition 17.3. *"$<$" and "$\leq$" satisfy all commonly known rules, such as*

- *Trichotomy of the order relation: Let $m, n \in \mathbb{N}_0$. Then exactly one of the following is true:*

$$m < n, \qquad m = n, \qquad m > n.$$

17.2 Constructing the Integers from $\mathbb{N}_0$

Definition 17.5 (Integers as equivalence classes)**.** We define the following equivalence relation $(m_1, n_1) \sim (m_2, n_2)$ on the cartesian product $\mathbb{N}_0 \times \mathbb{N}_0$:

$$(m_1, n_1) \sim (m_2, n_2) \;\Leftrightarrow\; m_1 + n_2 = n_1 + m_2 \tag{17.3}$$

We write $\mathbb{Z} := \{[(m, n)] : m, n \in \mathbb{N}_0\}$. In other words, $\mathbb{Z}$ is the set of all equivalence classes with respect to the equivalence relation (17.3).
We "embed" $\mathbb{N}_0$ into $\mathbb{Z}$ with the following injective function $e : \mathbb{N}_0 \to \mathbb{Z}$: e(m) := [(m,0)].
From this point forward we do not distinguish between $\mathbb{N}_0$ and its image $e(\mathbb{N}_0) \subseteq \mathbb{Z}$ and we do not distinguish between $\mathbb{N}$ and its image $e(\mathbb{N}) \subseteq \mathbb{Z}$. In particular we do not distinguish between the two zeros 0 and $[(0, 0)]$ and between the two ones 1 and $[(1, 0)]$.
Finally we write $-n$ for the integer $[(0, n)]$. ☐

Proposition 17.4 (Trichotomy of the integers)**.** *Let $z \in \mathbb{Z}$. Then exactly one of the following is true:*

(a) *$z \in \mathbb{N}$, i.e., $z = [(m, 0)]$ for some $m \in \mathbb{N}$*
(b) *$-z \in \mathbb{N}$, i.e., $z = [(0, n)]$ for some $n \in \mathbb{N}$*
(c) *$z = 0$.* ☐

Definition 17.6 (Addition, multiplication and subtraction on $\mathbb{Z}$)**.**
Let $[(m_1, n_1)]$ and $[(m_2, n_2)] \in \mathbb{Z}$. We define

$$-[(m_1, n_1)] := [n_1, m_1], \tag{17.4}$$
$$[(m_1, n_1)] + [(m_2, n_2)] := [(m_1 + m_2, n_1 + n_2)] \tag{17.5}$$
$$[(m_1, n_1)] \cdot [(m_2, n_2)] := [(m_1 m_2 + n_1 n_2, m_1 n_2 + n_1 m_2)] \tag{17.6}$$

We write $[(m_1, n_1)] - [(m_2, n_2)]$ ("$[(m_1, n_1)]$ minus $[(m_2, n_2)]$") as an abbreviation for $[(m_1, n_1)] + \big(- [(m_2, n_2)]\big)$.

We write $[(m_1, n_1)] < [(m_2, n_2)]$ if $[(m_2, n_2)] - [(m_1, n_1)] \in \mathbb{N}$, i.e., if there is $k \in \mathbb{N}$ such that $[(m_2, n_2)] - [(m_1, n_1)] = [(k, 0)]$. We then say that $[(m_1, n_1)]$ is less than $[(m_2, n_2)]$.

We write $[(m_1, n_1)] \leq [(m_2, n_2)]$ if $[(m_1, n_1)] < [(m_2, n_2)]$ or if $[(m_1, n_1)] = [(m_2, n_2)]$ and we then say that $[(m_1, n_1)]$ is less than or equal to $[(m_2, n_2)]$.

We write $[(m_1, n_1)] > [(m_2, n_2)]$ if $[(m_2, n_2)] < [(m_1, n_1)]$ and we then say that $[(m_1, n_1)]$ is greater than $[(m_2, n_2)]$.

We write $[(m_1, n_1)] \geq [(m_2, n_2)]$ if $[(m_2, n_2)] \leq [(m_1, n_1)]$ and we then say that $[(m_1, n_1)]$ is greater than or equal to $[(m_2, n_2)]$.

We write $\mathbb{Z}_{\geq 0}$ for the set of all integers z such that $z \geq 0$ and $\mathbb{Z}_{\neq 0}$ for the set of all integers z such that $z \neq 0$. You should convince yourself that $\mathbb{Z}_{\geq 0} = \mathbb{N}_0$. ☐

Proposition 17.5. *Let* $m, n \in \mathbb{N}_0$. *Then*

$$[(m,n)] + [(0,0)] = [(0,0)] + [(m,n)] = [(m,n)], \tag{17.7}$$

$$\big(-[(m,n)]\big) + [(m,n)] = [(m,n)] + \big(-[(m,n)]\big) = [0,0] \tag{17.8}$$

$$[(m,n)] \cdot [(1,0)] = [(1,0)] \cdot [(m,n)] = [(m,n)], \tag{17.9}$$

i.e., $[(0,0)]$ *becomes the neutral element with respect to addition,* $[(1,0)]$ *becomes the neutral element with respect to multiplication and* $-[(m,n)]$ *becomes the additive inverse of* $[(m,n)]$.

17.3 Constructing the Rational Numbers from $\mathbb{Z}$

Definition 17.7 (Fractions as equivalence classes)**.** We define the following equivalence relation $(p, q) \sim (r, s)$ on the cartesian product $\mathbb{Z} \times \mathbb{Z}_{\neq 0}$:

$$(p,q) \sim (r,s) \;\Leftrightarrow\; p \cdot s = q \cdot r \tag{17.10}$$

We write $\mathbb{Q} := \{[(p,q)] : p, q \in \mathbb{Z} \text{ and } q \neq 0\}$. In other words, $\mathbb{Q}$ is the set of all equivalence classes with respect to the equivalence relation (17.10).

We "embed" $\mathbb{Z}$ into $\mathbb{Q}$ with the injective function $e : \mathbb{Z} \to \mathbb{Q}$ defined as $e(z) := [(z, 1)]$. ☐

Definition 17.8 (Addition, multiplication, subtraction and division in $\mathbb{Q}$)**.**

Let $[(p_1, q_1)]$ and $[(p_2, q_2)] \in \mathbb{Q}$. We define

$$-[(p_1, q_1)] := [(-p_1, q_1)], \tag{17.11}$$
$$[(p_1, q_1)] + [(p_2, q_2)] := [(p_1 q_2 + q_1 p_2, n_1 n_2)] \tag{17.12}$$
$$[(p_1, q_1)] - [(p_2, q_2)] := [(p_1, q_1)] + \big(- [(p_2, q_2)]\big) \tag{17.13}$$
$$[(p_1, q_1)] \cdot [(p_2, q_2)] := [(p_1 p_2, q_1 q_2)] \tag{17.14}$$
$$[(p_1, q_1)]^{-1} := [(1,1)]/[(p_1, q_1)] \ := \ [(q_1, p_1)] \text{ (if } p_1 \neq 0), \tag{17.15}$$
$$[(p_1, q_1)]/[(p_2, q_2)] := [(p_1 q_2, q_1 p_2)] \ = \ [(p_1, q_1)] \cdot [(p_2, q_2)]^{-1} \text{ (if } p_2 \neq 0) \ \square \tag{17.16}$$

Proposition 17.6 (Trichotomy of the rationals). *Let $x \in \mathbb{Q}$. Then exactly one of the following is true:*

(a) *Either **(a)** $x > 0$, i.e., $x = [(p, q)]$ for some $p, q \in \mathbb{N}$,*
(b) *$-x > 0$, i.e., $x = [(-p, q)]$ for some $p, q \in \mathbb{N}$,*
(c) *$x = 0$.* $\square$

17.4 Constructing the Real Numbers via Dedekind Cuts

Definition 17.9 (Dedekind cuts). (Rudin, def.1.4)
We call a subset $\alpha \subseteq \mathbb{Q}$ a **cut** or **Dedekind cut** if it satisfies the following:

(a) $\alpha \neq \emptyset$ and $\alpha^{\complement} \neq \emptyset$
(b) Let $p, q \in \mathbb{Q}$ such that $p \in \alpha$ and $q < p$. Then $q \in \alpha$.
(c) α does not have a max: $\forall p \in \alpha \ \exists q \in \alpha$ such that $p < q$.

Given a cut α, let $p \in \alpha$ and $q \in \alpha^{\complement}$. We call p a **lower number** of the cut α and we call q an **upper number** of α. $\square$

Theorem 17.1. *(Rudin thm.1.5)*
Let $\alpha \subseteq \mathbb{Q}$ be a cut. Let $p \in \alpha, q \in \alpha^{\complement}$. Then $p < q$.

Theorem 17.2. *(Rudin thm.1.6)*
Let $r \in \mathbb{Q}$. Let $r^\star := \{p \in \mathbb{Q} : p < r\}$. Then $r^\star$ is a cut, and $r = \min\big((r^\star)^{\complement}\big)$.

Definition 17.10 (Rational cuts). Let $r \in \mathbb{Q}$. The cut

$$r^\star \ = \ \{p \in \mathbb{Q} : \ p \ < \ r\}$$

from the previous theorem is called the **rational cut** associated with r. $\square$

Definition 17.11 (Ordering Dedekind cuts). (Rudin def.1.9) Let α, β be two cuts.
We say $\alpha < \beta$ if $\alpha \subsetneq \beta$ (strict subset) and we say $\alpha \leq \beta$ if $\alpha < \beta$ or $\alpha = \beta$, i.e., $\alpha \subseteq \beta$. $\square$

Proposition 17.7 (Trichotomy of the cuts). *(Rudin thm.1.10)*
Let α, β be two cuts. Then either $\alpha < \beta$ or $\alpha > \beta$ or $\alpha = \beta$.

Theorem 17.3 (Addition of two cuts). *(Rudin thm.1.12) Let α, β be two cuts and let*

$$\alpha + \beta \ := \{a + b : a \in \alpha, b \in \beta\}.$$

Then the set of all cuts is an abelian group with this operation. In other words, $+$ is commutative and associative with a neutral element (which turns out to be $0^\star$, the rational cut corresponding to $0 \in \mathbb{Q}$) and a suitably defined cut $-\alpha$ for a given cut α which satisfies $\alpha + (-\alpha) = (-\alpha) + \alpha = 0^\star$
Having defined negatives $-\alpha$ for all cuts we then also can define their absolute values

$$|\alpha| \ := \begin{cases} \alpha & \text{if } \alpha \geq 0^\star, \\ -\alpha & \text{if } \alpha < 0^\star. \end{cases}$$

Theorem 17.4 (Multiplication of two cuts). *Let $\alpha \geq 0^\star, \beta \geq 0^\star$ be two nonnegative cuts. Let*

$$\alpha \cdot \beta \ := \begin{cases} \{q \in \mathbb{Q} : q < 0\} \cup \{ab : a \in \alpha, b \in \beta\} & \text{if } \alpha \geq 0^\star, \beta \geq 0^\star, \\ -|\alpha| \cdot |\beta| & \text{if } \alpha < 0^\star, \beta \geq 0^\star \text{ or } \alpha \geq 0^\star, \beta < 0^\star, \\ |\alpha| \cdot |\beta| & \text{if } \alpha < 0^\star, \beta < 0^\star. \end{cases}$$

Then the set $\alpha \cdot \beta$ is a cut, called the product of α and β.
It can be proved that for each cut $\alpha \neq 0^\star$ there is a cut α^{-1} uniquely defined by the equation $\alpha \cdot \alpha^{-1} = 1^\star$.

Theorem 17.5 (The set of all cuts forms a field).
Let Let $\mathbb{R}$ be the set of all cuts. Then $\mathbb{R}$ satisfies axioms 8.1 - 8.5 of B/G:

Addition and multiplication are both commutative and associative and the law of distributivity $\alpha \cdot (\beta + \gamma) \;=\; \alpha \cdot \beta + \alpha \cdot \gamma$ *holds.*
The cut $0^\star$ *is the neutral element for addtition and the cut* $1^\star$ *is the neutral element for multiplication.*
$-\alpha$ *is the additive inverse of any cut* α *and* α^{-1} *is the multiplicative inverse of* $\alpha \neq 0^\star$*.*
Further the set $\mathbb{R}_{>0} := \{\alpha \in \mathbb{R} : \alpha > 0^\star\}$ *satisfies B/G axiom 8.26.*

Theorem 17.6. *(Rudin thm.1.29)*
Let $\alpha, \beta \in \mathbb{R}$ *and let* $\alpha < \beta$*. Then there exists* $q \in \mathbb{Q}$ *such that* $\alpha < q^\star < \beta$

Theorem 17.7. *(Rudin thm.1.30) Let* $\alpha \in \mathbb{R}, p \in \mathbb{Q}$*. Then* $p \in \alpha \;\Leftrightarrow\; p^\star < \alpha$*, i.e.,* $p^\star \subsetneq \alpha$

Theorem 17.8 (Dedekind's Theorem)**.** *(Rudin thm.1.32)*
Let $\mathbb{R} = A \biguplus B$ *a partitioning of* $\mathbb{R}$ *such that*

(a) $A \neq \emptyset$ *and* $B \neq \emptyset$
(b) $\alpha \in A, \beta \in B \;\Rightarrow\; \alpha < \beta \;$ *(i.e.,* $\alpha \subsetneq \beta$*).*

Then there exists a unique cut $\gamma \in \mathbb{R}$ *such that if* $\alpha \in A$ *then* $\alpha \leq \gamma$ *and if* $\beta \in B$ *then* $\gamma \leq \beta$*.*

Corollary 17.1. *Let* $\mathbb{R} = A \uplus B$ *be a partitioning of* $\mathbb{R}$ *such that*

(a) $A \neq \emptyset$ *and* $B \neq \emptyset$
(b) $\alpha \in A, \beta \in B \;\Rightarrow\; \alpha < \beta \;$ *(i.e.,* $\alpha \subsetneq \beta$*).*

Then either $\max(A)(= l.u.b.(A))$ *exists or* $\min(B)(= g.l.b.(B))$ *exists.*

Theorem 17.9 (Completeness theorem for $\mathbb{R}$)**.** *(Rudin thm.1.36)*
Let $\emptyset \neq E \subset \mathbb{R}$ *and assume that* E *is bounded above. Then* E *has a least upper bound.*
It is denoted by $\sup(E)$*, also l.u.b.(E).*

17.5 Constructing the Real Numbers via Cauchy Sequences

In the following we always assume that

$$i, j, k, m, n \in \mathbb{N},\; \varepsilon, p, q, r, s, p_n, p_{i,j}, \dots \in \mathbb{Q},\; x, y, z, x_n, x_{i,j}, \dots \in \mathbb{R}\,.$$

The construction of the real numbers from the rationals is done according to the following steps:

(a) def. convergence in $\mathbb{Q}$: $\lim\limits_{n\to\infty} q_n = q \Leftrightarrow \forall$ pos. $\varepsilon \in \mathbb{Q}\ \exists\ N \in \mathbb{Q}$ such that if $n \geq N$ then $|q_n - q| < \varepsilon$.

(b) def. Cauchy seqs. in $\mathbb{Q}$: $(q_n)_n$ is Cauchy $\Leftrightarrow \forall$ pos. $\varepsilon \in \mathbb{Q}\ \exists\ N \in \mathbb{Q}$ such that if $i, j \geq N$ then $|q_i - q_j| < \varepsilon$.

(c) Let $\mathscr{C} := \{$ all Cauchy sequences in $\mathbb{Q}\}$. For $(q_n)_n, (r_n)_n$ we define $(q_n)_n \sim (r_n)_n$ iff $\lim\limits_{n\to\infty}(r_n - q_n) = 0$.

(d) Let $q \in \mathbb{Q}$ and $q_n := q\ \forall\ n$. Write q for $[(q_n)_n]$.

(e) Let $\mathbb{R} := \mathscr{C}_{/\sim}$. Show that for $[(p_n)], [(q_n)] \in \mathscr{C}$ the operations $([(p_n)_n], [(q_n)_n]) \mapsto [(p_n + q_n)_n]$ and $([(p_n)_n], [(q_n)_n]) \mapsto [(p_n \cdot q_n)_n]$ are well defined (do not depend on the particular members chosen from the equivalence classes).

(f) Let $[(p_n)_n] \neq 0$ (i.e., $\lim\limits_{n} p_n \neq 0$), i.e., we may assume $p_n \neq 0$ for all n. Show $-[(q_n)_n] := [(-q_n)_n]$ and $[(p_n)_n]^{-1} := [(1/p_n)_n]$ are additive and multiplicative inverses

g1. Define $[(p_n)_n] < [(q_n)_n]$ iff $\exists\ \varepsilon > 0$ and $N \in \mathbb{N}$ such that $q_n - p_n \geq \varepsilon\ \forall\ n \geq N$.

g2. Define $[(p_n)_n] \leq [(q_n)_n]$ iff $\forall\ \varepsilon > 0\ exists\ N \in \mathbb{N}$ such that $q_n - p_n \geq -\varepsilon\ \forall\ n \geq N$.

g3. show that $[(p_n)_n] < [(q_n)_n]$ iff $[(p_n)_n] \leq [(q_n)_n]$ and $[(p_n)_n] \neq [(q_n)_n]$.

(h) Show that $(\mathbb{R}, +, \cdot, <)$ satisfies the axioms of B/G ch.8 with the exception of the completeness axiom.
Easy to see this specific item: If $[(p_n)_n] > 0$ then there is $[(q_n)_n] > 0$ such that $[(q_n)_n] < [(p_n)_n]$: choose $\varepsilon > 0$ as in **g1** (remember: $\varepsilon \in \mathbb{Q}$) and set $q_n := \varepsilon/2$.

(i) Embed $\mathbb{Q}$ into $\mathbb{R}$: $q \mapsto \bar{q} := [(q, q, q, \dots)]$.

(j) Define limits and Cauchy sequences in $\mathbb{R}$ just as in **(a)** and **(b)**.

k. Let $(q_n)_n$ be Cauchy in $\mathbb{Q}$. Prove that $\bar{q_n} \to [(q_j)_j]$

l. Let $x_n \in \mathbb{R}$ such that $(x_n)_n$ is Cauchy in $\mathbb{R}$. With a density argument we find $q_n \in \mathbb{Q}$ such that $x_n \leq \bar{q}_n \leq x_n + 1/n$. Now show that **(1)** $(q_n)_n$ is Cauchy and then **(2)** $\lim\limits_{n} x_n = [(q_n)_n]$.

m. Prove completeness according to B/G: If nonempty $A \subseteq \mathbb{R}$ is bounded above then its set of upper bounds U has a min: Let $Q_n := \{i/j : i, j \in \mathbb{Z}$ and $j \leq n\}$. Let $U_n := U \cap Q_n$. Let $u_n := \min(U_n)$ (exists because $n \cdot U_n \subset Z$ is bounded below and has a min. Easy to see that u_n is Cauchy (in $\mathbb{Q}$ and, because distance$(u_n, A) \leq 1/n$, $[(u_n)_n]$ is the least upper bound of A.

18 Other Appendices

18.1 Greek Letters

The following section lists all greek letters that are commonly used in mathematical texts. You do not see the entire alphabet here because there are some letters (especially upper case) which look just like our latin alphabet letters. For example: $A =$ Alpha $B =$ Beta. On the other hand there are some lower case letters, namely epsilon, theta, sigma and phi which come in two separate forms. This is not a mistake in the following tables!

α	alpha	θ	theta	ξ	xi	ϕ	phi
β	beta	ϑ	theta	π	pi	φ	phi
γ	gamma	ι	iota	ρ	rho	χ	chi
δ	delta	κ	kappa	ϱ	rho	ψ	psi
ϵ	epsilon	$\varkappa$	kappa	σ	sigma	ω	omega
ε	epsilon	λ	lambda	ς	sigma		
ζ	zeta	μ	mu	τ	tau		
η	eta	ν	nu	υ	upsilon		

Γ	Gamma	Λ	Lambda	Σ	Sigma	Ψ	Psi
Δ	Delta	Ξ	Xi	Υ	Upsilon	Ω	Omega
Θ	Theta	Π	Pi	Φ	Phi		

18.2 Notation

This appendix on notation has been provided because future additions to this document may use notation which has not been covered in class. It only covers a small portion but provides brief explanations for what is covered.

For a complete list check the list of symbols and the index at the end of this document.

Notation 18.1. a) If two subsets A and B of a space Ω are disjoint, i.e., $A \cap B = \emptyset$, then we often write $A \uplus B$ rather than $A \cup B$ or $A + B$. Both $A^{\complement}$ and, occasionally, $\complement A$ denote the complement $\Omega \setminus A$ of A.

b) $\mathbb{R}_{>0}$ or $\mathbb{R}^+$ denotes the interval $]0, +\infty[$, $\mathbb{R}_{\geq 0}$ or $\mathbb{R}_+$ denotes the interval $[0, +\infty[$,

c) The set $\mathbb{N} = \{1, 2, 3, \cdots\}$ of all natural numbers excludes the number zero. We write $\mathbb{N}_0$ or $\mathbb{Z}_+$ or $\mathbb{Z}_{\geq 0}$ for $\mathbb{N} \uplus \{0\}$. $\mathbb{Z}_{\geq 0}$ is the B/G notation. It is very unusual but also very intuitive. □

Definition 18.1. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of real numbers. We call that sequence **nondecreasing** or **increasing** if $x_n \leq x_{n+1}$ for all $n \in \mathbb{N}$.

We call it **strictly increasing** if $x_n < x_{n+1}$ for all $n \in \mathbb{N}$.

We call it **nonincreasing** or **decreasing** if $x_n \geq x_{n+1}$ for all n.

We call it **strictly decreasing** if $x_n > x_{n+1}$ for all $n \in \mathbb{N}$. □

References

[1] Matthias Beck and Ross Geoghegan. The Art of Proof. Springer, 1st edition, 2010.

List of Symbols

Index