# Chapter 1
# Sets, Structures, Numbers

**Abstract** In this chapter we shall introduce most of the background needed to develop the foundations of mathematical analysis. We start with sets and algebraic structures. Then the real numbers are defined axiomatically. This in turn allows one to define natural numbers, integers, rational numbers, and irrational numbers as well as to derive fundamental properties of these numbers. Next, we study representations of real numbers. Then we turn our attention to mappings and the numerosity of sets. In particular, it is shown that there are many more real numbers than rational numbers. Furthermore, we introduce linear spaces. The chapter is concluded by defining the complex numbers.

## 1.1 Sets and Algebraic Structures

It is not possible to include here the foundations of set theory. The interested reader is referred to the extensive literature, e.g., Kunen [106] and Hrbacek and Jech [93]. So the approach taken here is naïve set theory.

We briefly recall some basics. Following Cantor we define a *set* to be any collection of definite, distinct objects of our perception or of our thought.

The objects of a set are called *elements*. If $M$ is a set and $x$ belongs to $M$, then we write $x \in M$; if $x$ does not belong to $M$, then we write $x \notin M$.

Note that this definition is *not* adequate for a formal development of set theory. So, we have to be careful.

We shall define sets either *extensionally*, i.e., by listing all the elements in the set; for example, $M =_{\mathrm{df}} \{a, b, c, d\}$, or *intensionally*, i.e., by providing a particular property $P$ that all the elements must fulfill; for example, we then write $M =_{\mathrm{df}} \{x \mid x \text{ satisfies } P\}$.

Furthermore, we shall use the following: By $\emptyset$ we denote the *empty set*, i.e., the set which contains no elements.

Moreover, we need the following definition:

© Springer International Publishing Switzerland 2016

W. Römisch and T. Zeugmann, *Mathematical Analysis and the Mathematics of Computation*, DOI 10.1007/978-3-319-42755-3_1

**Definition 1.1.** Let $M$ and $N$ be any sets. Then we write

(1) $N \subseteq M$ if $x \in N$ implies that $x \in M$ (*subset*);
(2) $N = M$ if $N \subseteq M$ and $M \subseteq N$, otherwise we write $N \neq M$;
(3) $N \subset M$ if $N \subseteq M$ and $N \neq M$ (*proper subset*);
(4) $M \cup N =_{df} \{x \mid x \in M \text{ or } x \in N\}$ (*union*);
(5) $M \cap N =_{df} \{x \mid x \in M \text{ and } x \in N\}$ (*intersection*);
(6) $M \setminus N =_{df} \{x \mid x \in M \text{ and } x \notin N\}$ (*difference*);
(7) $C_M(N) =_{df} M \setminus N$ (*complement* of $N$ with respect to $M$), where we have to assume that $N \subseteq M$.

*Example 1.1.* Let $M = \{a, b, c, d\}$ and let $N = \{b, c\}$. Then $b \in N$ and $b \in M$ as well as $c \in N$ and $c \in M$. We conclude that $N \subseteq M$, i.e., $N$ is a subset of $M$. Since $a \in M$ but $a \notin N$, we see that $N \subset M$, i.e., $N$ is a proper subset of $M$. Consequently, $M \subseteq N$ does *not* hold. If $M \subseteq N$ is not true then we write $M \nsubseteq N$. Furthermore, it is easy to see that $M \cup N = \{a, b, c, d\}$, $M \cap N = \{b, c\}$, $M \setminus N = \{a, d\}$, and $C_M(N) = \{a, d\}$.

**Theorem 1.1.** *Let $M, N$, and $S$ be any sets. Then the following properties are satisfied:*

(1) $M \subseteq M$;
(2) $M \subseteq N$ *and* $N \subseteq S$ *implies* $M \subseteq S$;
(3) $M \cap N \subseteq M \subseteq M \cup N$;
(4) $\emptyset \subseteq M$, $M \setminus M = \emptyset$, $M \setminus \emptyset = M$, *and* $M \setminus N \subseteq M$;
(5) *the union is associative and commutative, i.e.,* $(M \cup N) \cup S = M \cup (N \cup S)$ *and* $M \cup N = N \cup M$, *respectively;*
(6) *the intersection is associative and commutative, i.e.,*

$$(M \cap N) \cap S = M \cap (N \cap S) \text{ and } M \cap N = N \cap M, \text{ respectively.}$$

**Theorem 1.2.** *Let $M, N$, and $S$ be any sets. Then the following properties are satisfied:*

(1) $M \cap (N \cup S) = (M \cap N) \cup (M \cap S)$, *and*
$M \cup (N \cap S) = (M \cup N) \cap (M \cup S)$ (*distributive laws*);
(2) $M \cap N = M \setminus (M \setminus N)$;
(3) *let* $N \subseteq M$ *and* $S \subseteq M$; *then we have* $C_M(N \cup S) = C_M(N) \cap C_M(S)$ *and* $C_M(N \cap S) = C_M(N) \cup C_M(S)$ (*De Morgan's laws*).

We do not prove Theorems 1.1 and 1.2 here but leave the proofs of these theorems as an exercise.

Furthermore, we shall use the following notations: Let $\mathcal{S}$ be any finite or infinite collection of sets; then we set

$$\bigcup_{S \in \mathcal{S}} S =_{df} \{x \mid \text{ there is an } S \in \mathcal{S} \text{ such that } x \in S\}\ ,$$

$$\bigcap_{S \in \mathcal{S}} S =_{df} \{x \mid \text{ for all } S \in \mathcal{S} \text{ we have } x \in S\}\ .$$

If we have sets $S_1$, $S_2$, $S_3$, ... then we also use the notations $\bigcup\limits_{i=1}^{n} S_i$, and $\bigcap\limits_{i=1}^{n} S_i$, as well as $\bigcup\limits_{i=1}^{\infty} S_i$, and $\bigcap\limits_{i=1}^{\infty} S_i$.

Let $S$ be any set, then we write $\wp(S)$ to denote the set of all subsets of $S$. We call $\wp(S)$ the *power set* of $S$.

*Example 1.2.* Let us consider the set $S = \{a, b, c\}$. Then

$$\wp(S) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\} .$$

**Exercise 1.1.** *Let* $S_1 = \{1, 2, 3, 4\}$*, and let* $S_2 = \emptyset$*. Compute* $\wp(S_1)$ *and* $\wp(S_2)$*.*

Following Kuratowski [107] we define a set of two elements, where a first element is determined, and call it an *ordered pair*; that is, we set

$$(a, b) =_{df} \{\{a\}, \{a, b\}\} .$$

Note that this definition is adequate, since it allows one to show the characteristic property an ordered pair has to fulfill, i.e.,

$$(a, b) = (c, d) \quad \text{iff} \quad a = c \text{ and } b = d .$$

The definition of an ordered pair can be easily generalized to ordered triples, or more generally, ordered $n$-tuples, which we shall denote by $(a, b, c)$ and $(x_1, \ldots, x_n)$, respectively.

Let $M$ and $N$ be any sets. We define the *product* $M \times N$ of $M$ and $N$ by

$$M \times N =_{df} \{(m, n) \mid m \in M \text{ and } n \in N\} .$$

It is also called the *Cartesian product* of $M$ and $N$. Note that $M \times \emptyset = \emptyset$ by definition. Let $S_1, \ldots, S_n$ be any sets, then their $n$*-fold product* is the set

$$\bigtimes_{i=1}^{n} S_i =_{df} \{(s_1, \ldots, s_n) \mid s_i \in S_i \text{ for all } i = 1, \ldots, n\} .$$

Next, we turn our attention to *algebraic structures.* An algebraic structure is a non-empty set on which one or more *operations* are defined along with some axioms that must be satisfied.

**Definition 1.2 (Group).** Let $G \neq \emptyset$ be any set, and let $\circ\colon G \times G \to G$ be any binary operation. We call $(G, \circ)$ a *group* if

(1) $(a \circ b) \circ c = a \circ (b \circ c)$ for all $a$, $b$, $c \in G$ (i.e., $\circ$ is *associative*);
(2) there is a *neutral element* $e \in G$ such that $a \circ e = e \circ a = a$ for all $a \in G$;
(3) for every element $a \in G$ there exists an *inverse element* $b \in G$ such that $a \circ b = b \circ a = e$.
(4) A group is called an *Abelian group* if $\circ$ is also commutative, i.e., $a \circ b = b \circ a$ for all $a$, $b \in G$.

Commutative groups are called Abelian groups in honor of Niels Henrik Abel.

**Exercise 1.2.** *Show that the neutral element* $\mathsf{e}$ *and the inverse elements defined above are uniquely determined.*

**Definition 1.3 (Field).** Let $\mathsf{F} \neq \emptyset$ be any set containing two distinguished elements 0 and 1, where $0 \neq 1$, and let $\circ, *\colon \mathsf{F} \times \mathsf{F} \to \mathsf{F}$ be two binary operations. We call $(\mathsf{F}, \circ, *)$ a *field* if

(1) $(\mathsf{F}, \circ)$ is an Abelian group (with neutral element 0);
(2) $(\mathsf{F} \setminus \{0\}, *)$ is a group (with neutral element 1);
(3) the following *distributive laws* are satisfied:

$$\begin{aligned}\mathsf{a} * (\mathsf{b} \circ \mathsf{c}) &= (\mathsf{a} * \mathsf{b}) \circ (\mathsf{a} * \mathsf{b}),\\ (\mathsf{a} \circ \mathsf{b}) * \mathsf{c} &= (\mathsf{a} * \mathsf{c}) \circ (\mathsf{b} * \mathsf{c})\ .\end{aligned}$$

(4) A field $(\mathsf{F}, \circ, *)$ is said to be *Abelian* (or *commutative*) if $\mathsf{a} * \mathsf{b} = \mathsf{b} * \mathsf{a}$ for all $\mathsf{a}$, $\mathsf{b} \in \mathsf{F} \setminus \{0\}$ holds.

We refer to 0 as the *neutral element* and to 1 as the *identity element.*

The following theorem provides some fundamental properties of fields: Note that "iff" is used as an abbreviation for "if and only if."

**Theorem 1.3.** *Let* $(\mathsf{F}, \circ, *)$ *be any field. Then we have*

(1) $\mathsf{a} * 0 = 0 * \mathsf{a} = 0$ *for all* $\mathsf{a} \in \mathsf{F}$*;*
(2) $\mathsf{a} * \mathsf{b} = 0$ *iff* $\mathsf{a} = 0$ *or* $\mathsf{b} = 0$*;*
(3) *for all* $\mathsf{a}$, $\mathsf{b} \in \mathsf{F}$*,* $\mathsf{a} \neq 0$ *there is precisely one* $\mathsf{x} \in \mathsf{F}$ *such that* $\mathsf{a} * \mathsf{x} = \mathsf{b}$*.*

*Proof.* First, we show that $\mathsf{a} * 0 = 0$. Let $\mathsf{a} \in \mathsf{F}$ be arbitrarily fixed. Note that $\mathsf{a} \circ 0 = \mathsf{a}$ (Property (2) of Definition 1.2) and consider

$$\begin{aligned}\mathsf{a} * \mathsf{a} &= \mathsf{a} * (\mathsf{a} \circ 0) = (\mathsf{a} * \mathsf{a}) \circ (\mathsf{a} * 0)\\ (\mathsf{a} * \mathsf{a}) \circ (\mathsf{a} * \mathsf{a})_{inv} &= (\mathsf{a} * \mathsf{a}) \circ (\mathsf{a} * 0) \circ (\mathsf{a} * \mathsf{a})_{inv}\\ 0 &= (\mathsf{a} * \mathsf{a}) \circ (\mathsf{a} * \mathsf{a})_{inv} \circ (\mathsf{a} * 0)\\ 0 &= (\mathsf{a} * 0)\ ,\end{aligned}$$

where $(\mathsf{a} * \mathsf{a})_{inv}$ denotes the inverse of $(\mathsf{a} * \mathsf{a})$ with respect to $\circ$.

The first equation above used the distributive law, the second line applied the fact that every element in $\mathsf{F}$ has an inverse with respect to $\circ$, the third line used that $\circ$ is commutative, and the last line the fact that 0 is the neutral element.

The second part $0 * \mathsf{a} = 0$ can be shown analogously.

The sufficiency of Property (2) follows from Property (1).

For the necessity assume that $\mathsf{a} * \mathsf{b} = 0$ and $\mathsf{a} \neq 0$. Let $\overline{\mathsf{a}}$ be the inverse of $\mathsf{a}$ with respect to $*$. Then we have

$$\begin{aligned} \overline{a} * 0 &= \overline{a} * (a * b) = (\overline{a} * a) * b \\ 0 &= 1 * b = b \ , \end{aligned}$$

where we used Property (1), the associativity of $*$, and the fact that 1 is the identity element.

To show Property (3) we note that $x = (\overline{a} * b)$ is a solution of $a * x = b$ (recall that $a \neq 0$). So, it remains to show that the solution is uniquely determined.

Suppose that there are two solutions $x_1$ and $x_2$ in $F$. Then we have

$$\begin{aligned} a * x_1 &= b = a * x_2 \\ \overline{a} * (a * x_1) &= \overline{a} * (a * x_2) \\ (\overline{a} * a) * x_1 &= (\overline{a} * a) * x_2 \\ 1 * x_1 &= 1 * x_2 \\ x_1 &= x_2 \ , \end{aligned}$$

where we again used the associativity of $*$ and the property of the identity element that $1 * x = x * 1 = x$ for all $x \in F$. ∎

Next we turn our attention to relations.

**Definition 1.4 (Binary Relation).** Let $S \neq \emptyset$ be any set. Then any subset $R \subseteq S \times S$ is called a *binary relation* over S.

**Definition 1.5 (Order Relation).** Let $S \neq 0$ be any set, and let $\leqslant$ be a binary relation over $S$. We call $\leqslant$ an *order relation* if the following axioms are satisfied:

(1) $x \leqslant x$ for all $x \in S$ (*reflexivity*);
(2) $x \leqslant y$ and $y \leqslant z$ implies $x \leqslant z$ for all $x, y, z \in S$ (*transitivity*);
(3) $x \leqslant y$ and $y \leqslant x$ implies $x = y$ for all $x, y \in S$ (*antisymmetry*).

We call $(S, \leqslant)$ an *ordered set* if $\leqslant$ is an order relation.

**Definition 1.6.** Let $(S, \leqslant)$ be any ordered set, and let $A \subseteq S$. We say that $A$ is *bounded from above* if there is a $c \in S$ such that $a \leqslant c$ for all $a \in A$. The element $c$ is said to be an *upper bound* of $A$.

The terms *bounded from below* and *lower bound* are similarly defined.

Let $A$ be bounded from above and let

$$B =_{df} \{c \mid c \in S \text{ and } c \text{ is an upper bound of } A\} \ .$$

If there is an $s \in B$ such that $s \leqslant c$ for all $c \in B$ then we call $s$ the *least upper bound* of $A$ or the *supremum* of $A$ and denote it by $\sup A$. If $s \in A$ then we call it the *maximum* of $A$ and write $\max A$.

The terms *greatest lower bound*, *infimum*, $\inf A$, *minimum*, and $\min A$ are similarly defined.

## 1.2 The Real Numbers

We introduce real numbers by using an axiomatic approach.

**Definition 1.7.** A set $\mathbb{R}$ is called a *set of the real numbers* if there are two operations $+\colon \mathbb{R} \to \mathbb{R}$ and $\cdot : \mathbb{R} \to \mathbb{R}$ (called addition and multiplication, respectively) and an order relation $\leqslant$ over $\mathbb{R}$ such that the following axioms are satisfied:

(1) $(\mathbb{R}, +, \cdot\,)$ is an Abelian field;
(2) $(\mathbb{R}, \leqslant)$ satisfies also the following properties:

(i) For all $x, y \in \mathbb{R}$ we have $x \leqslant y$ or $y \leqslant x$;
(ii) for all $x, y \in \mathbb{R}$ with $x \leqslant y$ we have $x + z \leqslant y + z$ for all $z \in \mathbb{R}$;
(iii) $0 \leqslant x$ and $0 \leqslant y$ implies $0 \leqslant x \cdot y$ for all $x, y \in \mathbb{R}$.

(3) For all $A \subseteq \mathbb{R}$, if $A \neq \emptyset$, and if $A$ is bounded from above then $\sup A \in \mathbb{R}$ exists.

We shall use these axioms given in Definition 1.7 to derive all the properties of the real numbers that are relevant for the development of our theories.

It should be noted that we *postulate* the existence of a non-empty set $\mathbb{R}$ that satisfies the axioms of Definition 1.7.

It should also be noted that $\mathbb{R}$ is *not* uniquely determined by the axioms given. But the different models of $\mathbb{R}$ differ only in properties that are not interesting for the analysis.

For example, let us consider the set $\mathbb{R}' =_{\mathrm{df}} \{(r, 0) \mid r \in \mathbb{R}\}$ and let us define $+$ by setting $(a, 0) + (b, 0) =_{\mathrm{df}} (a + b, 0)$ as well as $\cdot$ by setting $(a, 0) \cdot (b, 0) =_{\mathrm{df}} (a \cdot b, 0)$. Then $\mathbb{R}'$ also satisfies the axioms of Definition 1.7 provided $\mathbb{R}$ does.

Note that one can derive the axioms of Definition 1.7 from axioms of axiomatic set theory and from axioms for the natural numbers. This was done by Cantor and Dedekind among others. This approach was fundamental for the historical development of the analysis, but here it suffices to know that it can be done. Below we shall also touch on the so-called *Dedekind cuts.*

Clearly, 0 and 1 are the neutral element and the identity element of $\mathbb{R}$, respectively. We write $-a$ to denote the additive inverse of $a$ for any $a \in \mathbb{R}$. The multiplicative inverse of $a$ for any $a \in \mathbb{R}\setminus\{0\}$ is denoted by $a^{-1}$ (or $1/a$). We frequently omit the multiplication dot, i.e., we write $ab$ instead of $a \cdot b$.

Some further notations are needed. We write $x < y$ if $x \leqslant y$ and $x \neq y$. For $a, b \in \mathbb{R}$ with $a < b$ we use

$$\begin{aligned} [a, b] &=_{\mathrm{df}} \{x \mid x \in \mathbb{R} \text{ and } a \leqslant x \leqslant b\} &&(\textit{closed interval});\\ ]a, b[ &=_{\mathrm{df}} \{x \mid x \in \mathbb{R} \text{ and } a < x < b\} &&(\textit{open interval});\\ [\mathrm{a,b}[ &=_{\mathrm{df}} \{x \mid x \in \mathbb{R} \text{ and } a \leqslant x < b\} &&(\textit{half-open interval});\\ ]\mathrm{a,b}] &=_{\mathrm{df}} \{x \mid x \in \mathbb{R} \text{ and } a < x \leqslant b\} &&(\textit{half-open interval}). \end{aligned}$$

We define the *absolute value* $|x|$ of $x \in \mathbb{R}$ as follows:

$$|x| =_{\mathtt{df}} \begin{cases} x, & \text{if } x \geqslant 0\ ; \\ -x, & \text{if } x < 0\ . \end{cases} \tag{1.1}$$

Note that $-x \leqslant |x| \leqslant x$ for all $x \in \mathbb{R}$. So, the absolute value is a function that maps the real numbers to the non-negative numbers. Figure 1.1 shows the graph of this function. For a formal definition of what is meant by a function, we refer the reader to Section 1.6.

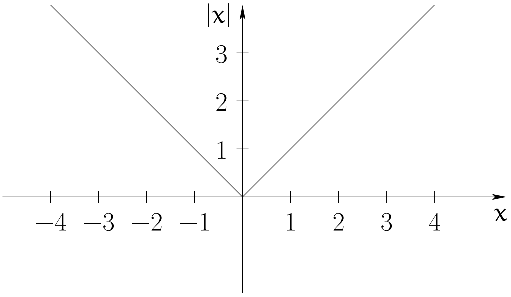

Fig. 1.1: The graph of the function $|x|$

We continue with some properties of the real numbers that can be derived from the axioms given in Definition 1.7.

**Proposition 1.1.** *For all* $a$, $b \in \mathbb{R}$ *the following properties are satisfied:*

(1) $ab > 0$ *iff* $(a > 0$ *and* $b > 0)$ *or* $(a < 0$ *and* $b < 0)$*;*
(2) $a < b$ *implies* $a < \frac{1}{2}(a + b) < b$.
(3) *For the absolute value we have*

(i) $|a| \geqslant 0$ *and* $|a| = 0$ *iff* $a = 0$*;*
(ii) $|ab| = |a|\,|b|$*;*
(iii) $|a + b| \leqslant |a| + |b|$ (*triangle inequality*)*;*
(iv) $|\,|a| - |b|\,| \leqslant |a - b|$.

*Proof.* Necessity. Let $ab > 0$ and suppose that $a > 0$ and $b < 0$. Taking into account that $-b = -b$ and $b + (-b) = 0$ (by the definition of the additive inverse), we obtain $0 < -b$.

By Theorem 1.3 we have $a \cdot 0 = 0$ for all $a \in \mathbb{R}$. Thus,

$$\begin{aligned} a(b + (-b)) &= 0 \\ ab + a(-b) &= 0 \quad \text{(distributive law)}\ . \end{aligned}$$

Next, we add $(-ab)$ on both sides and obtain

$$\begin{aligned} ab + a(-b) + (-ab) &= 0 + (-ab) \\ ab + (-ab) + a(-b) &= -ab \quad \text{(commutative law)} \\ a(-b) &= -ab\ . \end{aligned}$$

Since $0 < a$, $0 < -b$, and $a(-b) = -ab$, we therefore get by Axiom (2), Part (iii) that $0 < a(-b) = -ab$, and consequently $ab < 0$, a contradiction.

The sufficiency is a direct consequence of Axiom (2), Part (iii).

We show Property (2). Let $a < b$. Since 1 is the identity element, we directly get $a = 1 \cdot a$, and thus, by distributivity, $a + a = (1 + 1)a = 2a$. By Axiom (2), Part (ii) we conclude

$$\begin{aligned} a + a &< a + b \\ 2a &< a + b \\ 0 &< (a + b) + (-2a) \\ 0 &< \frac{1}{2}((a + b) + (-2a)) \quad \text{(Axiom (2), Part (iii))} \\ 0 &< \frac{1}{2}(a + b) + (-a) \\ a &< \frac{1}{2}(a + b) \quad \text{(Axiom (2), Part (ii))}\ . \end{aligned}$$

The right-hand side is shown analogously.

Finally, we prove Property (3). We only show the triangle inequality here; the rest is left as an exercise.

The definition of the absolute value gives $a \leqslant |a|$ and $b \leqslant |b|$ as well as $-a \leqslant |a|$ and $-b \leqslant |b|$. So by Axiom (2), Part (ii) we get

$$a + b \leqslant |a| + |b|\ , \tag{1.2}$$

$$(-a) + (-b) \leqslant |a| + |b|\ . \tag{1.3}$$

Therefore, if $a + b \geqslant 0$, then the definition of the absolute value implies that $|a + b| = a + b \leqslant |a| + |b|$ by (1.2).

Furthermore, if $a + b < 0$ then we use $-(a + b) = (-a) + (-b)$, and thus, by Inequality (1.3) we obtain

$$\begin{aligned} 0 < -(a + b) &\leqslant |a| + |b| \\ |a + b| &\leqslant |a| + |b|\ . \end{aligned}$$

This completes the proof of Proposition 1.1. ∎

Proposition 1.1 directly allows for the following corollary:

**Corollary 1.1.** *For all* $a \in \mathbb{R}$ *with* $a \neq 0$ *we have*

(1) $aa > 0$*;*

(2) *in particular,* $0 < 1$ *and* $a > 0$ *iff* $1/a > 0$.

*Proof.* Property (1) is a direct consequence of Proposition 1.1, Property (1). Since $1 \cdot 1 = 1$ and by definition $1 \neq 0$, we have $0 < 1$. Finally, $a \cdot (1/a) = 1 > 0$, and so the rest is directly due to Proposition 1.1, Property (1). ∎

We continue with further properties of the real numbers that can be derived from the axioms given in Definition 1.7.

**Theorem 1.4.** *Let* $A$, $B \subseteq \mathbb{R}$ *be non-empty sets such that* $a \leqslant b$ *for all* $a \in A$ *and all* $b \in B$*. Then there is a* $c \in \mathbb{R}$ *such that* $a \leqslant c \leqslant b$ *for all* $a \in A$ *and all* $b \in B$*.*

*Proof.* By assumption, $A \neq \emptyset$ and bounded from above (every $b \in B$ is an upper bound). Thus, by Axiom (3) we know that $c =_{df} \sup A \in \mathbb{R}$ exists. Hence, $a \leqslant c$ for all $a \in A$. Since $\sup A$ is the least upper bound, we must have $c \leqslant b$ for all $b \in B$. ∎

Theorem 1.4 allows for the following corollary:

**Corollary 1.2.** *Let* $A$, $B \subseteq \mathbb{R}$ *be any non-empty sets such that* $a < b$ *for all* $a \in A$ *and all* $b \in B$ *and* $A \cup B = \mathbb{R}$*. Then there exists a uniquely determined* $c \in \mathbb{R}$ *such that* $a \leqslant c \leqslant b$ *for all* $a \in A$ *and all* $b \in B$*.*

*Proof.* By Theorem 1.4 the existence of a $c$ with the desired properties is clear. Suppose there are $c_1$ and $c_2$ such that $a \leqslant c_i \leqslant b$, $i = 1, 2$, for all $a \in A$ and all $b \in B$.

Without loss of generality let $c_1 < c_2$.

Then $\sup A \leqslant c_1 < c_2 \leqslant b$ for all $b \in B$. Consequently, $c_1 \notin B$ and $c_2 \notin A$. Thus, we must have $c_1 \in A$ and $c_2 \in B$.

Therefore, by Proposition 1.1, Property (2), we directly obtain

$$\begin{aligned} & c_1 = \sup A < \frac{1}{2}(c_1 + c_2) < c_2 \leqslant b \qquad \text{for all } b \in B \\ & \frac{1}{2}(c_1 + c_2) \notin A \cup B = \mathbb{R} \ , \end{aligned}$$

a contradiction to Axiom (1) ($\mathbb{R}$ is a field). ∎

Sets $A, B$ fulfilling the assumptions of Corollary 1.2 are called a *Dedekind cut* and usually written as $(A|B)$. That is, for any two such sets $A, B$, there is precisely one point $c \in \mathbb{R}$, the so-called *cut*. That means the reals do not have any “gap.”

In fact, Dedekind [43] used such cuts $(A|B)$, where $A, B \subseteq \mathbb{Q}$ and $A \cup B = \mathbb{Q}$ (here $\mathbb{Q}$ denotes the set of all rational numbers), to introduce the real numbers based on the axiomatic definition of the natural numbers.

**Remarks.** We shall proceed here in the opposite direction; i.e., we shall define the set of all natural numbers $\mathbb{N}$ as a particular subset of the set of all real numbers.

After having defined the natural numbers, it is easy to define the rational numbers.

## 1.3 Natural Numbers, Rational Numbers, and Real Numbers

Next, we want to define the natural numbers and then the integers and rational numbers. We need the following:

**Definition 1.8 (Inductive Set).** A set $M \subseteq \mathbb{R}$ is said to be *inductive* if

(1) $1 \in M$;
(2) $x \in M$ implies $x + 1 \in M$.

Obviously, there are many inductive sets. In particular, the set $\mathbb{R}$ itself is inductive.

**Exercise 1.3.** *Prove that* $\{x \mid x \in \mathbb{R},\ x = 1 \text{ or } 2 \leqslant x\}$ *is inductive.*

**Definition 1.9 (Natural Numbers).** Let $\mathcal{M}$ be the family of all inductive subsets of $\mathbb{R}$, i.e., let $\mathcal{M} = \{M \mid M \subseteq \mathbb{R},\ M \text{ is inductive}\}$.

The set $\mathbb{N} =_{df} \bigcap_{M \in \mathcal{M}} M$ is said to be the *set of all natural numbers.*

By its definition, $\mathbb{N}$ is the smallest inductive set contained in $\mathbb{R}$. Also, by definition, we have $0 \notin \mathbb{N}$. We set $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Using the natural numbers, it is easy to define the following:

**Definition 1.10 (Integers, Rational Numbers).** We define the following sets:

(1) The set $\mathbb{Z} =_{df} \{x \mid x \in \mathbb{N}_0 \ \text{ or } \ -x \in \mathbb{N}\}$ is said to be the *set of all integers*;
(2) the set $\mathbb{Q} =_{df} \{x \mid \text{ there are } p, q \text{ with } p \in \mathbb{Z},\ q \in \mathbb{Z} \setminus \{0\} \text{ and } x = p/q\}$ is said to be the *set of all rational numbers.*

Note that by their definition we already know that $\mathbb{Z}$ and $\mathbb{Q}$ are subsets of $\mathbb{R}$.

Since $\mathbb{N}$ is the smallest inductive set, we directly get the *principle of induction.* That is, if $M \subseteq \mathbb{N}$ and $M$ is inductive, then $M = \mathbb{N}$ must hold.

So, if we want to show that an assertion $A(n)$ is true for all numbers $n \in \mathbb{N}$, then we can proceed as follows:

1. Show $A(1)$ holds.
2. Assume $A(n)$ holds. Show that $A(n)$ implies $A(n+1)$.

The principle of induction is also very useful to define mathematical objects $O(n)$. One defines $O(1)$, and then one continues to define $O(n+1)$ by using $O(n)$.

We continue with some examples of inductive definitions.

*The factorial function*

$$1! =_{\mathrm{df}} 1 \ , \tag{1.4}$$
$$(n+1)! =_{\mathrm{df}} n!(n+1) \ . \tag{1.5}$$

*Finite sums*

$$\sum_{i=1}^{1} a_i =_{\mathrm{df}} a_1 \ , \tag{1.6}$$
$$\sum_{i=1}^{n+1} a_i =_{\mathrm{df}} \sum_{i=1}^{n} a_i + a_{n+1} \ , \text{ where } a_i \in \mathbb{R}\,,\ i = 1,\ldots,n+1 \ . \tag{1.7}$$

Analogously, one defines $\prod_{i=1}^{n} a_i$ (*finite products*), where again $a_i \in \mathbb{R}$ for all $i = 1,\ldots,n+1$ or $a^n$ (*powers*) for $a \in \mathbb{R}$ and *exponents* $n \in \mathbb{N}$.

We continue with fundamental properties of natural numbers. First, we show that they are closed with respect to addition and multiplication. Furthermore, we prove that $n+1$ is the *successor* of $n$ for all $n \in \mathbb{N}$.

**Theorem 1.5.** *The following properties hold:*

(1) *If* $m$, $n \in \mathbb{N}$ *then* $m+n \in \mathbb{N}$ *and* $mn \in \mathbb{N}$*;*
(2) *for all* $n \in \mathbb{N}$ *we have* $]n, n+1[\, \cap \mathbb{N} = \emptyset$.

*Proof.* The proof of Assertion (1) is by induction. Let $m \in \mathbb{N}$ be arbitrarily fixed and let $A(n)$ be the assertion that $m+n \in \mathbb{N}$.

Then $A(1)$ is true, since $\mathbb{N}$ is inductive.

Assume $A(n)$ is true. We show $A(n+1)$ holds.

Since $A(n)$ is true, we have $m+n \in \mathbb{N}$. Furthermore, $\mathbb{N}$ is inductive and thus $(m+n)+1 \in \mathbb{N}$. Since addition is *associative*, we conclude

$$(m+n)+1 = m+(n+1) \quad \text{and so} \quad m+(n+1) \in \mathbb{N} \ .$$

Consequently, $A(n+1)$ holds, and so $A(n)$ is true for all $n \in \mathbb{N}$.

The part $mn \in \mathbb{N}$ for all $m$, $n \in \mathbb{N}$ is left as an exercise.

We continue with Assertion (2). To prove Assertion (2), it suffices to show that $M =_{\mathrm{df}} \{n \mid n \in \mathbb{N},\ ]n, n+1[\, \cap \mathbb{N} = \emptyset\}$ is inductive.

To see that $1 \in M$, consider

$$\begin{aligned} ]1,2[\, \cap \mathbb{N} &\subseteq\, ]1,2[\, \cap \{x \mid x \in \mathbb{R},\ x = 1 \text{ or } 2 \leqslant x\} \\ &= \,]1,2[\, \cap\ (\{1\} \cup \{x \mid x \in \mathbb{R},\ 2 \leqslant x\}) = \emptyset \ . \end{aligned}$$

Assume $n \in M$. We have to show that $n+1 \in M$. Suppose the converse, i.e., there is an $m$ such that $m \in\, ]n+1, n+2[\, \cap \mathbb{N}$. So,

$$n+1 < m \ < \ n+2 \ . \tag{1.8}$$

On the other hand, the set $\widetilde{M} = \{m \mid m \in \mathbb{N},\ m-1 \in \mathbb{N}_0\}$ is clearly inductive, i.e., $\widetilde{M} = \mathbb{N}$. Hence, $m \geqslant 2$ implies $m - 1 \in \mathbb{N}$. But by (1.8) we have $n < m - 1 < n + 1$, implying that $n \notin M$, a contradiction. ∎

Definition 1.9 and Theorem 1.5 *justify* the identification of $\mathbb{N}$ with the set $\{1, 2, 3, \ldots\}$. But we still do not know whether or not $\mathbb{N}$ is bounded from above. The negative answer is given below.

**Theorem 1.6 (Archimedes).** *The set* $\mathbb{N}$ *is not bounded from above.*

*Proof.* Suppose the converse. Then, by Axiom (3) (cf. Definition 1.7), we know that $s =_{\text{df}} \sup \mathbb{N} \in \mathbb{R}$ exists.

Consider $s - 1$. Clearly, $s - 1$ is *not* an upper bound for $\mathbb{N}$. But then there must be an $n \in \mathbb{N}$ such that $s - 1 < n$, which in turn implies that $s < n + 1$. Since $\mathbb{N}$ is inductive, we have $n + 1 \in \mathbb{N}$, a contradiction to $s = \sup \mathbb{N}$. ∎

Note that Property (2) of Corollary 1.3 is usually called the *Archimedean property* or the *axiom of Archimedes*, while Property (1) is named after Eudoxus.

**Corollary 1.3.**

(1) *For all* $\varepsilon > 0$*,* $\varepsilon \in \mathbb{R}$*, there is an* $n \in \mathbb{N}$ *such that* $1/n < \varepsilon$*.*
(2) *For all* $x, y \in \mathbb{R}$ *with* $x > 0$ *and* $y \geqslant 0$ *there is an* $n \in \mathbb{N}$ *such that* $y \leqslant n \cdot x$*.*

*Proof.* To show Assertion (1), suppose the converse. Then there is an $\varepsilon_0$ such that $1/n \geqslant \varepsilon_0$ for all $n \in \mathbb{N}$. Hence, $n \leqslant 1/\varepsilon_0$ for all $n \in \mathbb{N}$ and so $\mathbb{N}$ is bounded from above, a contradiction to Theorem 1.6.

Let $x, y \in \mathbb{R}$ be arbitrarily fixed such that the assumptions of Property (2) hold. We consider $y/x \in \mathbb{R}$. By Theorem 1.6 there is an $n \in \mathbb{N}$ with $y/x \leqslant n$ and consequently $y \leqslant n \cdot x$. ∎

Next we show that every non-empty set of natural numbers possesses a minimal element.

**Theorem 1.7.** *Let* $\emptyset \neq M \subseteq \mathbb{N}$*; then* $M$ *possesses a minimal element.*

*Proof.* Suppose to the contrary that $M$ does not contain a minimal element. We consider the following set $K$ defined as:

$$K =_{\text{df}} \{k \mid k \in \mathbb{N} \text{ and } k < n \text{ for all } n \in M\} .$$

It suffices to show that $K$ is inductive, since then we know that $K = \mathbb{N}$, which in turn implies $M = \emptyset$, a contradiction.

We indirectly show that $1 \in K$. If $1 \notin K$ then $1 \in M$ and $1$ is minimal for $M$.

Now, let $k \in K$. We have to show that $k + 1 \in K$. Suppose $k + 1 \notin K$. Then there is an $m_1 \in M$ such that $m_1 \leqslant k + 1$. But by our supposition $M$ does not have a minimal element. So there must be an $m_2 \in M$ with $m_2 < m_1 \leqslant k + 1$. By Theorem 1.5, Assertion (2), we conclude $m_2 \leqslant k$, a contradiction to $k \in K$. ∎

We continue by showing that $\mathbb{Q}$ is *dense* in $\mathbb{R}$.

**Theorem 1.8.** *For every* $r \in \mathbb{R}$ *and every* $\varepsilon > 0$ *there exists a* $q \in \mathbb{Q}$ *such that* $q \in \,]r - \varepsilon, r + \varepsilon[$.

*Proof.* Let $\varepsilon > 0$ be arbitrarily fixed. We distinguish the following cases:

*Case* 1. $r \geqslant 0$.

By Corollary 1.3, Assertion (1), there is an $n \in \mathbb{N}$ such that $1/n < \varepsilon$. Consider the set $M =_{df} \{m \mid m \in \mathbb{N},\ m > n \cdot r\}$. By Corollary 1.3, Assertion (2), we know that $M \neq \emptyset$. By Theorem 1.7, $M$ contains a minimal element $m \in M$. So,

$$m > n \cdot r \quad \text{(definition of } M\text{)}\,, \tag{1.9}$$

$$m - 1 \leqslant n \cdot r \quad \text{(since } m \text{ is minimal)}\,. \tag{1.10}$$

We define $q =_{df} m/n$. Then by Inequalities (1.9) and (1.10) we obtain

$$\frac{m-1}{n} \leqslant r \ < \ \frac{m}{n}\ , \tag{1.11}$$

$$q - \frac{1}{n} \leqslant r \ < \ q\ . \tag{1.12}$$

Hence, $r - \varepsilon < r < q$ because of $\varepsilon > 0$ and by the right-hand side of (1.12). Moreover, $q \leqslant r + 1/n$ by the left-hand side of (1.12). So $q \leqslant r + 1/n < r + \varepsilon$, and we are done.

*Case* 2. $r < 0$.

Then $-r > 0$, and by Case 1, there is a $q \in \mathbb{Q}$ with $-r - \varepsilon < q < -r + \varepsilon$. Consequently, $r - \varepsilon < -q < r + \varepsilon$, and the theorem is shown. ∎

Our next goal is to show that $\mathbb{Q} \subset \mathbb{R}$ (cf. Theorem 1.11). In order to achieve this result and several related ones, we need some preparations. First, it is sometimes useful to extend the domain of an inductive definition to the set $\mathbb{N}_0$. Since $0 + 1 = 1$, we see that 1 is the successor of 0. So this extension is well in line with our previous inductive definitions. Let us exemplify this for the factorial function (cf. (1.4) and (1.5)). We replace the induction basis (cf. (1.4)) by defining $0! =_{df} 1$ and leave (1.5) unchanged. So we have to check that $1! = 1$. By the induction step now we have

$$1! = (0+1)! \ = \ 0!(0+1) \ = \ 1 \cdot 1 \ = \ 1\ ,$$

and thus our extension of the factorial function is well defined.

Furthermore, it is also very useful to extend the definition of powers to all exponents $n \in \mathbb{N}_0$. So we set $a^0 =_{df} 1$ and $a^{n+1} = a \cdot a^n$. A quick check shows that $a^1 = a$ and so this extension is also well defined for all $a \in \mathbb{R} \setminus \{0\}$. So this leaves the problem of whether or not it is meaningful to define $0^0$. We refer the reader to Knuth [103] for a short historical survey concerning the debate around this question. After the debate stopped, there was apparently the general consensus around that $0^0$ should be undefined. Though we agree with this general conclusion, there are many places where it is convenient to define $0^0 =_{\mathrm{df}} 1$, including the setting here (see Theorem 1.9 below) and whenever one deals with series (see Chapter 2). Therefore, unless stated otherwise, we shall assume that $0^0 = 1$.

Next, we define for all $k, n \in \mathbb{Z}$ the so-called *binomial coefficients*, i.e.,

$$\binom{n}{k} =_{\mathrm{df}} \begin{cases} \dfrac{n(n-1)\cdots(n-k+1)}{k!}\,, & \text{if } k \geqslant 0\ ; \\ 0\,, & \text{if } k < 0\ . \end{cases} \tag{1.13}$$

The symbol $\binom{n}{k}$ is read "$n$ choose $k$." Note that $\binom{0}{0} = 1$, since for $k = 0$ we have the empty product, i.e., a product of no factors, in the numerator, which is conventionally defined to be 1.

Also note that $\binom{n}{n} = 1$ for all $n \in \mathbb{N}_0$, but $\binom{n}{n} = 0$ for $n \in \mathbb{Z} \setminus \mathbb{N}_0$. We should memorize this fact.

It is useful to memorize the following formulae for all $n \in \mathbb{Z}$:

$$\binom{n}{0} = 1\ , \quad \binom{n}{1} = n\ , \quad \binom{n}{2} = \frac{n(n-1)}{2}\ . \tag{1.14}$$

Definition 1.13 can be recast for the case that $k, n \in \mathbb{N}_0$ and $n \geqslant k$. That is, we can multiply the numerator and the denominator of (1.13) by $(n-k)!$ and obtain

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}\ . \tag{1.15}$$

This formula directly yields a nice symmetry property, i.e., we can change $k$ to $n-k$ for $k, n \in \mathbb{N}_0$ and $n \geqslant k$. Looking again at Definition 1.13 we see that this symmetry is also satisfied if $n \in \mathbb{N}_0$ and $k \in \mathbb{Z}$. Therefore, for all $n \in \mathbb{N}_0$ and $k \in \mathbb{Z}$ we have

$$\binom{n}{k} = \binom{n}{n-k}\ . \tag{1.16}$$

Note that this symmetry *fails* for $n \in \mathbb{Z} \setminus \mathbb{N}_0$. To see this, let us consider

$$\binom{-1}{k} = \frac{(-1)(-2)\cdots(-k)}{k!} = (-1)^k \tag{1.17}$$

for all $k \in \mathbb{N}_0$. So for $k = 0$ we have $\binom{-1}{0} = 1$ but $\binom{-1}{-1-0} = 0$ by Definition 1.13. If $k \in \mathbb{Z} \setminus \mathbb{N}_0$ then $\binom{-1}{k} = 0$ (cf. Definition 1.13) and $\binom{-1}{-1-k} = (-1)^{-1-k} \neq 0$. Consequently, we have shown that

$$\binom{-1}{k} \neq \binom{-1}{-1-k} \quad \text{for all } k \in \mathbb{Z}\ . \tag{1.18}$$

We leave it as an exercise to show that (1.15) is also false for all other negative integers.

Below we shall also need the so-called *addition formula* for binomial coefficients, i.e., for all $n \in \mathbb{N}$ and all $k \in \mathbb{Z}$ we have

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} . \tag{1.19}$$

*Proof.* We distinguish the following cases:

*Case* 1. $k \in \mathbb{N}$.

By Definition 1.13, the definition of the factorial function, and by using the identities $n - 1 - k + 1 = n - k$ and $n - 1 - (k - 1) + 1 = n - k + 1$ we directly obtain

$$\begin{aligned}\binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)\cdots(n-k)}{k!} + \frac{(n-1)\cdots(n-k+1)}{(k-1)!} \\ &= \frac{(n-1)\cdots(n-k+1)(n-k)}{k!} \\ &\quad + \frac{(n-1)\cdots(n-k+1)k}{k!} \\ &= \frac{(n-1)\cdots(n-k+1)(n-k+k)}{k!} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!} = \binom{n}{k} ,\end{aligned}$$

and the addition formula is shown for $k \in \mathbb{N}$.

*Case* 2. $k \in \mathbb{Z} \setminus \mathbb{N}$.

Then we have $k \leqslant 0$. If $k < 0$ then, by Definition 1.13, we see that all binomial coefficients in the addition formula are zero, and thus the addition formula is shown. If $k = 0$ then $\binom{n-1}{0} = \binom{n}{0} = 1$ and $\binom{n-1}{-1} = 0$. Thus, the addition formula is shown for $k \in \mathbb{Z} \setminus \mathbb{N}$.

Case 1 and 2 together imply the addition formula as stated in Equation (1.19). ∎

Figure 1.2 shows some special values of the binomial coefficients. These values are the beginning of *Pascal's triangle* (named after Blaise Pascal, who published it in 1655). In China this triangle is known as *Yang-Hui's triangle.* Yang Hui published it already in 1261. We refer the reader to Edwards [52] for the history of Pascal's triangle. This triangle is easily obtained. We write the first and last entry as 1 in every row, then the addition formula tells us that the remaining numbers in a row are obtained by adding the number just above the desired number and the number to its left. The numbers in Pascal's triangle satisfy many identities. We refer the reader to Graham, Knuth, and Patashnik [71] for further information.

| $n$ | $\binom{n}{0}$ | $\binom{n}{1}$ | $\binom{n}{2}$ | $\binom{n}{3}$ | $\binom{n}{4}$ | $\binom{n}{5}$ | $\binom{n}{6}$ | $\binom{n}{7}$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 1 | | | | | | | |
| 1 | 1 | 1 | | | | | | |
| 2 | 1 | 2 | 1 | | | | | |
| 3 | 1 | 3 | 3 | 1 | | | | |
| 4 | 1 | 4 | 6 | 4 | 1 | | | |
| 5 | 1 | 5 | 10 | 10 | 5 | 1 | | |
| 6 | 1 | 6 | 15 | 20 | 15 | 6 | 1 | |
| 7 | 1 | 7 | 21 | 35 | 35 | 21 | 7 | 1 |

Fig. 1.2: Pascal's triangle

Assertion (1) of the following theorem explains where the name binomial coefficient comes from: They get their name from the binomial theorem.

**Theorem 1.9.** *The following assertions hold:*

(1) *For all* $a$, $b \in \mathbb{R}$ *and all* $n \in \mathbb{N}$ *we have*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \quad \text{(binomial theorem)} \ .$$

(2) *For all* $a \in \mathbb{R}$, $a > -1$ *and all* $n \in \mathbb{N}$ *we have* $(1+a)^n \geqslant 1 + na$ (called Bernoulli's inequality).[1]

(3) *For all* $a \in [0,1]$ *and all* $n \in \mathbb{N}$ *we have* $(1+a)^n \leqslant 1 + (2^n - 1)a$.

*Proof.* Assertion (1) is shown inductively. For the induction basis let us look at the cases $n = 0$ and $n = 1$. We directly see that

$$\begin{aligned}(a+b)^0 &= 1a^0b^0 \ = \ 1 \\ (a+b)^1 &= 1a^1b^0 + 1a^0b^1 = a + b \ ,\end{aligned}$$

and the induction basis is shown. Also, we have seen why our convention to set $a^0 = 1$ for all $a \in \mathbb{R}$ is really meaningful.

Next, we assume the induction hypothesis for $n$ and perform the induction step from $n$ to $n+1$; that is, we have to show that

$$(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k \ .$$

This is done as follows: We use the inductive definition of powers, then we apply the distributive law in lines three to six below. Thus, we obtain

[1] Named after Jakob I. Bernoulli

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k \quad \text{(by the induction hypothesis)} \\
&= a\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k + b\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k \quad \text{(distributive law)} \\
&= a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=0}^{n-1}\binom{n}{k}a^{n-k}b^{k+1} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=1}^{n}\binom{n}{k-1}a^{n-k+1}b^k + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\left[\binom{n}{k} + \binom{n}{k-1}\right]a^{n-k+1}b^k + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}a^{n-k+1}b^k + b^{n+1} \quad \text{(by (1.19))} \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^{n+1-k}b^k\ .
\end{aligned}$$

Therefore, we have shown Assertion (1).

We show Assertion (2) inductively. The case $n = 1$ is obvious. Assume the inequality for $n$. For $n + 1$ we obtain

$$\begin{aligned}
(1+a)^{n+1} &= (1+a)^n(1+a) \\
&\geqslant (1+na)(1+a) \quad \text{(by the induction hypothesis} \\
&\qquad\qquad\qquad\qquad\quad \text{and Axiom 2, (iii))}\ .
\end{aligned}$$

Since $a > -1$, we have $1 + a > 0$, and the distributive law gives

$$(1+na)(1+a) = 1 + na + a + na^2 \geqslant 1 + (n+1)a\ ,$$

where we used that $a^2 \geqslant 0$ and $n > 0$ and Axiom 2, (iii). Hence, Assertion (2) is shown.

Assertion (3) is shown analogously. Let $a \in [0, 1]$. Assertion (3) is obviously true for $n = 1$. We assume that $(1 + a)^n \leqslant 1 + (2^n - 1)a$.

For the induction step consider

$$\begin{aligned}
(1+a)^{n+1} &= (1+a)^n(1+a) \\
&\leqslant (1+(2^n-1)a)(1+a) \\
&\leqslant 1 + 2^n a + (2^n-1)a^2 \\
&\leqslant 1 + (2\cdot 2^n - 1)a\ ,
\end{aligned}$$

where the last step holds, since $a^2 \leqslant a$. ∎

The binomial theorem also provides a combinatorial meaning of the binomial coefficients. If we expand the power $(a+b)^n$ in its $n$ factors then the distributive law tells us that we have to take from each factor either $a$ or $b$. So the number of terms with $n-k$ factors of $a$ and $k$ factors of $b$ is given by the coefficient of $a^{n-k}b^k$; i.e., it is $\binom{n}{k}$.

**Exercise 1.4.** *Prove the following: For all* $n \in \mathbb{N}_0$ *we have*

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n \ ,$$

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

**Exercise 1.5.** *Show the following identity for the binomial coefficients holds: For all* $k, n \in \mathbb{Z}$ *we have*

$$\binom{n}{k} = (-1)^k \binom{k-n-1}{k} \ .$$

## 1.4 Roots

This is a good place to show the following theorem which establishes the existence and uniqueness of solutions of equations having the form $x^n = a$, where the right-hand side is any positive real number:

**Theorem 1.10.** *Let* $a \in \mathbb{R}$, $a > 0$ *and* $n \in \mathbb{N}$. *Then there exists a unique number* $x \in \mathbb{R}$ *such that* $x > 0$ *and* $x^n = a$.

*Proof.* Let $a \in \mathbb{R}$, $a > 0$, and let $n \in \mathbb{N}$ be arbitrarily fixed. We consider the set $S_a =_{df} \{y \mid y \in \mathbb{R},\ y \geqslant 0,\ y^n \leqslant a\}$. Clearly, $0 \in S_a$ and so $S_a \neq \emptyset$.

Now, let $y \in S_a$. Then, by the definition of $S_a$ we have

$$y^n \leqslant a < a + 1 \leqslant (1+a)^n \ ,$$

and thus $y \leqslant 1 + a$. Consequently, $S_a$ is also bounded from above. By Axiom (3) we conclude that $x =_{df} \sup S_a \in \mathbb{R}$ exists. Since $\min\{1, a\} \in S_a$, we also have $x > 0$.

*Claim* 1. $x^n = a$.

Suppose that $x^n < a$. Let $\eta =_{df} a - x^n > 0$ and

$$\varepsilon =_{df} \min\left\{x,\ \frac{\eta}{(2^n - 1)x^{n-1}}\right\} \ .$$

Then we have

$$\begin{aligned}(x+\varepsilon)^n &= x^n\left(1+\frac{\varepsilon}{x}\right)^n\\ &\leqslant x^n\left(1+(2^n-1)\cdot\frac{\varepsilon}{x}\right) \quad \text{(Theorem 1.9, Ass. (3))}\\ &= x^n+(2^n-1)\varepsilon x^{n-1}\leqslant x^n+\eta = a\ .\end{aligned}$$

But this implies that $(x+\varepsilon)\in S_a$, a contradiction to $x=\sup S_a$.

So, we must have $x^n \geqslant a$.

Next, suppose that $x^n > a$. By Corollary 1.3 there exists an $m \in \mathbb{N}$ such that $1/x < m$. This implies that $-1/(mx) > -1$, and hence Theorem 1.9, Assertion (2) is applicable. We obtain

$$\begin{aligned}\left(x-\frac{1}{m}\right)^n &= x^n\left(1-\frac{1}{mx}\right)^n\\ &\geqslant x^n\left(1-\frac{n}{mx}\right) \qquad \text{(Theorem 1.9, Ass. (2))}\ .\end{aligned}$$

Now, we choose $m$ large enough such that $m > \max\left\{\frac{1}{x},\ \frac{nx^{n-1}}{x^n-a}\right\}$. Thus,

$$\left(x-\frac{1}{m}\right)^n \geqslant x^n\left(1-\frac{n}{mx}\right)\geqslant x^n\left(1-\frac{x^n-a}{x^n}\right) = a\ .$$

Consequently, for any $y\in S_a$ we have $y^n\leqslant a\leqslant (x-1/m)^n$. Therefore, we conclude that $y\leqslant x-1/m$.

But this means that $x-1/m$ is an upper bound for $S_a$, and so we have a contradiction to $x=\sup S_a$.

Putting this all together, we must have $x^n=a$. ∎

We call the $x$ satisfying the properties of Theorem 1.10 the $n$*th root of* $a$. It is denoted by $\sqrt[n]{a}$ or $a^{1/n}$.

Note that we have just proved the *existence* of $n$th roots. So far, we have no idea how to *compute* them.

On the positive side, Theorem 1.10 allows us to define powers of positive real numbers $a$ for rational exponents. This is done as follows: We have already defined $a^n$ for $n\in\mathbb{N}_0$. Now, we extend this definition as follows:

$$a^{-n} =_{\text{df}} \frac{1}{a^n} \quad \text{for all } n\in\mathbb{N}\ .$$

So, $a^p$ is defined for all $p\in\mathbb{Z}$.

Next, let $r\in\mathbb{Q}$. Then there are $p, q\in\mathbb{Z}$, $q\neq 0$, such that $r=p/q$. Without loss of generality, we can assume $q\in\mathbb{N}$. We define

$$a^r =_{\text{df}} \sqrt[q]{a^p}\ . \tag{1.20}$$

**Exercise 1.6.** *Show the definition of* $a^r$ *to be independent of the choice of the representation of* $r$.

Now we are ready to show the desired separation result, i.e., $\mathbb{Q} \subset \mathbb{R}$.

**Theorem 1.11.** *The rational numbers are a proper subset of the real numbers. In particular,* $\sqrt{2} \notin \mathbb{Q}$.

*Proof.* It suffices to show $\sqrt{2} \notin \mathbb{Q}$. Suppose the converse, i.e., $\sqrt{2} \in \mathbb{Q}$. Then we can directly conclude that $M = \{m \mid m \in \mathbb{N},\ m \cdot \sqrt{2} \in \mathbb{N}\} \neq \emptyset$.

By Theorem 1.7 there exists an $m_0 \in M$ such that $m_0 \leqslant m$ for all $m \in M$. Therefore, we know that $m_0\sqrt{2} \in \mathbb{N}$ and because of $1 < \sqrt{2} < 2$, we also have $\ell_0 =_{df} m_0\sqrt{2} - m_0 \in \mathbb{N}$ and $\ell_0 < m_0$ (note that $\sqrt{2} < 2$ implies $\sqrt{2} - 1 < 1$).

On the other hand, $\ell_0\sqrt{2} = 2m_0 - m_0\sqrt{2} \in \mathbb{N}$. So, $\ell_0 \in M$ and $\ell_0 < m_0$, a contradiction to the choice of $m_0$.

Consequently, $M = \emptyset$ and thus $\sqrt{2} \notin \mathbb{Q}$. ∎

The elements of $\mathbb{R} \setminus \mathbb{Q}$ are called *irrational numbers.*

Theorem 1.11 directly allows for the following corollary:

**Corollary 1.4.** *The following assertions hold:*

(1) *Between any two different (rational) real numbers there is always a (rational) real number;*

(2) *between any two different (rational) real numbers there is always an (irrational) rational number.*

*Proof.* Assertion (1) follows from Proposition 1.1 (Assertion (2)).

To show Assertion (2), let $a$, $b \in \mathbb{R}$, $a < b$. Now, we apply Theorem 1.8 to $(a + b)/2$ for $\varepsilon = (b - a)/8$. Thus, there is a $q \in \mathbb{Q}$ such that $a < q < b$.

Finally, let $a$, $b \in \mathbb{Q}$, $a < b$. Then $(1/\sqrt{2})a < (1/\sqrt{2})b$. As just shown, there exists a $q \in \mathbb{Q}$ such that $(1/\sqrt{2})a < q < (1/\sqrt{2})b$. Without loss of generality let $q \neq 0$. Hence, $a < \sqrt{2}q < b$, and as in the proof of Theorem 1.11 one easily verifies that $\sqrt{2}q \notin \mathbb{Q}$. ∎

## 1.5 Representations of the Real Numbers

Next we ask how we may represent the real numbers. The following lemma is needed to prepare the corresponding result: It introduces the technique of *nested intervals.*

**Lemma 1.1.** *Let* $k \in \mathbb{N}_0$, $m \in \mathbb{N}$, $m \geqslant 2$, *and* $z_i \in \{0, \ldots, m-1\}$ *for all* $i \in \mathbb{N}$. *Then there exists a uniquely determined* $x \in \mathbb{R}$, $x \geqslant 0$ *such that* $x \in \bigcap_{n \in \mathbb{N}} [a_n, b_n]$, *where* $a_n = \sum_{i=1}^{n} z_i m^{k-i}$ *and* $b_n = a_n + m^{k-n}$ *for all* $n \in \mathbb{N}$.

*Proof.* By definition we have $0 \leqslant a_n < b_n$ for all $n \in \mathbb{N}$. Furthermore, by construction we know that $a_n \leqslant a_{n+1}$ and

$$\begin{aligned} b_{n+1} = \sum_{i=1}^{n+1} z_i m^{k-i} + m^{k-(n+1)} &= a_n + (z_{n+1}+1)m^{k-(n+1)} \\ &\leqslant a_n + m \cdot m^{k-(n+1)} = b_n \quad \text{for all } n \in \mathbb{N} \ . \end{aligned}$$

Consequently, $\{a_n \mid n \in \mathbb{N}\}$ is bounded from above by $b_1$. We define

$$x =_{df} \sup\{a_n \mid n \in \mathbb{N}\} \ .$$

Therefore, we already have $x \geqslant 0$ and $a_n \leqslant x$ for all $n \in \mathbb{N}$.

*Claim* 1. $x \leqslant b_n$ *for all* $n \in \mathbb{N}$.

Suppose the converse, i.e., there exists an $n_*$ such that $b_{n_*} < x$. Then we have $a_n < b_n \leqslant b_{n_*} < x$ for all $n \geqslant n_*$. Consequently, we conclude that $a_n < b_{n_*} < \frac{1}{2}(x + b_{n_*}) < x$ for all $n \geqslant n_*$, and so $x$ cannot be the least upper bound of $\{a_n \mid n \in \mathbb{N}\}$, a contradiction. This shows Claim 1.

We conclude that $x \in \bigcap_{n\in\mathbb{N}}[a_n, b_n]$.

*Claim* 2. $x$ *is uniquely determined.*

Suppose the converse, i.e., there are $x, y \in \bigcap_{n\in\mathbb{N}}[a_n, b_n]$ with $x \neq y$, where without loss of generality, $x < y$. So $a_n \leqslant x < y \leqslant a_n + m^{k-n}$ and thus $0 < y - x \leqslant b_n - a_n = m^{k-n}$ for all $n \in \mathbb{N}$. By Corollary 1.3 there is an $n_0 \geqslant 2$ such that $\frac{1}{n_0} < y - x$. Now, let $n \geqslant k + n_0$. Then we have

$$\begin{aligned} \frac{1}{2^{n_0}} < \frac{1}{n_0} \ &< \ y - x \ \leqslant \ m^{k-n} \\ &= \frac{1}{m^{n-k}} \ \leqslant \ \frac{1}{2^{n-k}} \ \leqslant \ \frac{1}{2^{n_0}} \ , \end{aligned}$$

a contradiction. ∎

**Theorem 1.12.** *Let* $m \in \mathbb{N}$, $m \geqslant 2$, *and* $x \in \mathbb{R}$, $x \geqslant 0$. *Then there exist* $k \in \mathbb{N}_0$ *and* $z_i \in \{0, \ldots, m-1\}$ *for all* $i \in \mathbb{N}$ *such that* $x \in \bigcap_{n\in\mathbb{N}}[a_n, b_n]$, *where* $a_n =_{df} \sum_{i=1}^{n} z_i m^{k-i}$ *and* $b_n =_{df} a_n + m^{k-n}$ *for all* $n \in \mathbb{N}$.

*Proof.* Let $x \in \mathbb{R}$, $x \geqslant 0$, and $m \in \mathbb{N}$, $m \geqslant 2$, be arbitrarily fixed. Consider the set $K =_{df} \{n \mid n \in \mathbb{N},\ x < m^n\}$. Theorem 1.7 implies that $k =_{df} \min K$ exists.

Let $n \in \mathbb{N}$ be any natural number such that $\frac{x}{m-1} < n$. Then we know that $x < n(m-1)$. Using Theorem 1.9, Assertion (2), we obtain

$$\begin{aligned} m^n &= (m-1+1)^n = (1+m-1)^n \\ &\geqslant 1 + n(m-1) > 1 + x > x \ . \end{aligned}$$

Hence, we can conclude that $n \in K$.

We define $\ell_1 =_{\mathrm{df}} \min\left\{n \mid n \in \mathbb{N},\ \dfrac{x}{m^{k-1}} < n\right\}$ and $z_1 =_{\mathrm{df}} \ell_1 - 1$. By construction we have $z_1 \in \mathbb{N}_0$ and

$$z_1 = \ell_1 - 1 \leqslant \frac{x}{m^{k-1}} < \ell_1 = z_1 + 1 \ .$$

Note that $x/m^{k-1} < m$. Consequently, $z_1 < m$ and

$$z_1 m^{k-1} = a_1 \leqslant x < (z_1 + 1)m^{k-1} = b_1 \ ,$$

and so $x \in [a_1, b_1]$ (and the first interval has been constructed).

So it suffices to iterate the construction. Consider $(x - z_1 m^{k-1})/m^{k-2} \in \mathbb{R}$. Then there is a $z_2$ such that

$$0 \leqslant z_2 \leqslant \frac{x - z_1 m^{k-1}}{m^{k-2}} \ < \ z_2 + 1 \ , \text{ and } z_2 \leqslant m - 1 \ .$$

Consequently, an easy calculation gives

$$a_2 = z_1 m^{k-1} + z_2 m^{k-2} \leqslant x < z_1 m^{k-1} + z_2 m^{k-2} + m^{k-2} = b_2 \ ,$$

and therefore $x \in [a_2, b_2]$.

Let us perform the induction step formally. Then we define

$$\ell_j =_{\mathrm{df}} \min\left\{\frac{1}{m^{k-j}}\left(x - \sum_{i=1}^{j-1} z_i m^{k-i}\right)\right\} \quad \text{and } z_j =_{\mathrm{df}} \ell_j - 1 \ .$$

As above, we directly obtain

$$\begin{aligned} z_j &\leqslant \frac{1}{m^{k-j}}\left(x - \sum_{i=1}^{j-1} z_i m^{k-i}\right) < z_j + 1 \quad \text{and thus} \\ z_j m^{k-j} &\leqslant \left(x - \sum_{i=1}^{j-1} z_i m^{k-i}\right) < z_j m^{k-j} + m^{k-j} \ , \quad \text{and so} \\ a_j &\leqslant x < a_j + m^{k-j} = b_j \ . \end{aligned}$$

By the induction hypothesis, $a_{j-1} \leqslant x < b_j$, and therefore

$$z_j \leqslant \frac{1}{m^{k-j}}(x - a_{j-1}) < \frac{1}{m^{k-j}}(b_{j-1} - a_{j-1}) = m \ .$$

Hence, the theorem is shown. ∎

As the proof of Theorem 1.12 shows, $k$ and all $z_i \in \{0, \ldots, m-1\}$, $i \in \mathbb{N}$, are uniquely determined. By Lemma 1.1 and its proof, we can express $x$ as

$$x = z_1 \cdots z_k . z_{k+1} z_{k+2} \cdots$$

i.e., we obtain the so-called $m$-*representation* of $x$.

In our proof of Theorem 1.12 we have always ensured that $a_j \leqslant x < b_j$. We should note that the proof goes through *mutatis mutandis* if we always ensure that $a_j < x \leqslant b_j$. Lemma 1.1 guarantees that we obtain the same $x$.

If $m = 10$ and $m = 2$ then we refer to the resulting representation as *decimal representation* and *binary representation*, respectively. Interestingly, in the first case we obtain for $x = 1/2$ and $m = 10$ the representation $0.50\cdots$ and in the second case $0.4999\cdots$. So, the $m$-representation of $x$ is in *this* sense not uniquely determined.

Note that in the second case we have a representation which is called a *repeating decimal* (or *recurring decimal*), since at some point it becomes *periodic*. Here by periodic we mean that there is some finite sequence of digits that is repeated indefinitely. In the case of $0.4999\cdots$ this sequence has length 1 and is 9 and we then write $0.4\overline{9}$. Another example is $1/3 = 0.\overline{3}$ (read as $0.\overline{3}$ repeating) and a more complicated one is $1/7 = 0.\overline{142857}$. A decimal representation with repeating final 0 is called *terminating* before these zeros, e.g., we just write 0.5 in the first case. Note that terminating representations and repeating decimal representations represent rational numbers. Of course, these considerations generalize to $m$-representations. For example, if we have a terminating decimal representation, e.g., 0.125, then we can directly write this number as a fraction, i.e., we have $0.125 = 125/1000$. So, while this case is clear, it may be less obvious to find integers $p, q$, $q \neq 0$, if a repeating decimal is given. For example, assume we are given $0.\overline{125}$ and we want to express it as $p/q$. The reader is encouraged to think about this problem. We shall solve it in Section 2.10 (see Example 2.19).

**Exercise 1.7.** *Determine the two decimal representations of* 1.

**Exercise 1.8.** *Determine a binary representation of* $1/7$. *Find out whether or not it is uniquely determined.*

## 1.6 Mappings and Numerosity of Sets

Let $X$ and $Y$ be any sets. The notion of a *mapping* is central for mathematical analysis and many other branches of mathematics. Thus, we continue by defining it.

**Definition 1.11.** Let $X$ and $Y$ be any sets.

(1) We call any $F \subseteq X \times Y$ a *mapping* on $X$ into $Y$.
(2) For every $x \in X$ we call $F(x) =_{\mathrm{df}} \{y \mid y \in Y,\ (x, y) \in F\}$ the *value of* $F$ *in* $x$.
(3) We call $\mathrm{dom}(F) =_{\mathrm{df}} \{x \mid x \in X,\ F(x) \neq \emptyset\}$ the *domain* of $F$.
(4) We call $\mathrm{range}(F) =_{\mathrm{df}} \{y \mid y \in Y,$ there is an $x \in X$ with $(x, y) \in F\}$ the *range* of $F$.
(5) If $(x, y) \in F$ then we call $x$ the *preimage* of $y$ and $y$ the *image* of $x$ under $F$.

Of particular importance are mappings that assign to every $x \in X$ precisely one element $y \in Y$. We call such mappings *definite*. Note that we shall mainly deal with definite mappings. Thus we refer to them frequently just as mappings, or *functions*, and sometimes call them *operators* or *functionals*. Roughly speaking, in this book, functions map numbers or tuples of numbers to numbers and/or to tuples of numbers. Operators map functions to functions and functionals map functions to numbers.

For definite mappings we often use the following notations:
$F\colon \operatorname{dom}(F) \to Y$ and $Fx = y$ provided $(x, y) \in F$.
If $\operatorname{dom}(F) = X$ then we say that $F$ is a mapping *from* $X$ into $Y$.
If $\operatorname{range}(F) = Y$ then we say that $F$ is a mapping on $X$ *to* (or *onto*) $Y$.
By $F(X, Y)$ we denote the set of definite mappings from $X$ into $Y$.
For the sake of illustration we include some examples.

*Example 1.3.* Let $X$ be any set, $Y = X$, and $I_X =_{df} \{(x, x) \mid x \in X\}$. Then we call $I_X$ the *identity mapping*. Note that $\operatorname{dom}(I_X) = X$ and that $\operatorname{range}(I_X) = X$.

*Example 1.4.* Let $X$ and $Y$ be any sets. We call the mappings
$pr_1\colon X \times Y \to X\ ,\ pr_1(x, y) =_{df} x$ and $pr_2\colon X \times Y \to Y\ ,\ pr_2(x, y) =_{df} y$
the first and second *projection* of $X \times Y$.

**Exercise 1.9.** *Consider the following mapping* $F\colon \mathbb{N} \to \mathbb{N}$ *defined as:*

$$F =_{df} \{(n, 2n) \mid n \in \mathbb{N}\}\ .$$

*Determine* $\operatorname{dom}(F)$ *and* $\operatorname{range}(F)$.

Next, we define important properties of mappings.

**Definition 1.12.** Let $X$ and $Y$ be any sets, and let $F\colon X \to Y$ be any mapping.

(1) We call $F$ *injective* if $F$ is definite and if $Fx_1 = Fx_2$ implies that $x_1 = x_2$ for all $x_1, x_2 \in \operatorname{dom}(F)$.
(2) We call $F$ *surjective* if $\operatorname{range}(F) = Y$.
(3) We call $F$ *bijective* if $F$ is injective and surjective.
(4) Let $A \subseteq X$; then we call $F(A) =_{df} \bigcup_{x \in A} F(x)$ the *image* of $A$ with respect to $F$.
(5) We call $F^{-1} =_{df} \{(y, x) \mid (y, x) \in Y \times X,\ (x, y) \in F\}$ the *inverse* mapping of $F$.
(6) Let $B \subseteq Y$; then $F^{-1}(B) =_{df} \{x \mid x \in X,\ \text{there is a } y \in B \text{ with } (x, y) \in F\}$ is called the *preimage* of $B$ with respect to $F$.

**Exercise 1.10.** *Consider again the mapping* $F$ *defined in Exercise 1.9.*

(1) *Determine whether or not* $F$ *is injective;*
(2) *determine whether or not* $F$ *is surjective.*

Note that $I_X$ is bijective for every non-empty set $X$. The projection functions $pr_1$ and $pr_2$ (see Example 1.4) are surjective but *not injective*.

**Definition 1.13 (Restriction, Continuation).** Let $\mathsf{X}$ and $\mathsf{Y}$ be any sets, and let $\mathsf{F}\colon \mathsf{X} \to \mathsf{Y}$ be any mapping.

(1) Let $\mathsf{A} \subseteq \mathsf{X}$; then we call $\mathsf{F}|_{\mathsf{A}} =_{\mathrm{df}} \{(x, y) \mid (x, y) \in \mathsf{F},\ x \in \mathsf{A}\}$ the *restriction* of $\mathsf{F}$ to $\mathsf{A}$.

(2) A mapping $\hat{\mathsf{F}} \subseteq \mathsf{X} \times \mathsf{Y}$ is said to be a *continuation* of $\mathsf{F}$ if $\mathrm{dom}(\mathsf{F}) \subseteq \mathrm{dom}(\hat{\mathsf{F}})$ and $\mathsf{F}(x) = \hat{\mathsf{F}}(x)$ for all $x \in \mathrm{dom}(\mathsf{F})$.

For example, our definition of the binomial coefficients (1.13) provides a mapping $\binom{\cdot}{\cdot}\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{N}_0$, i.e., $\mathsf{X} = \mathbb{Z} \times \mathbb{Z}$, $\mathsf{Y} = \mathbb{N}_0$, and $\mathsf{F}(n, k) = \binom{n}{k}$. Consequently, our recast Equation (1.15) is a restriction of this mapping to $\mathsf{A} = \{(n, k) \mid n,\ k \in \mathbb{N}_0,\ n \geqslant k\}$.

The following exercise summarizes several properties that are occasionally needed:

**Exercise 1.11.** *Let* $\mathsf{X}$ *and* $\mathsf{Y}$ *be any sets, let* $\mathsf{F}\colon \mathsf{X} \to \mathsf{Y}$ *be any mapping from* $\mathsf{X}$ *into* $\mathsf{Y}$*, and let* $\mathsf{A}, \tilde{\mathsf{A}} \subseteq \mathsf{X}$ *and* $\mathsf{B} \subseteq \mathsf{Y}$*. Then the following properties are satisfied:*

(1) $\mathsf{A} \neq \emptyset$ *iff* $\mathsf{F}(\mathsf{A}) \neq \emptyset$*;*
(2) $\mathsf{A} \subseteq \tilde{\mathsf{A}}$ *implies* $\mathsf{F}(\mathsf{A}) \subseteq \mathsf{F}(\tilde{\mathsf{A}})$*;*
(3) $\mathsf{F}(\mathsf{A} \cap \tilde{\mathsf{A}}) \subseteq \mathsf{F}(\mathsf{A}) \cap \mathsf{F}(\tilde{\mathsf{A}})$ (*equality holds if* $\mathsf{F}^{-1}$ *is definite*)*;*
(4) $\mathsf{F}(\mathsf{A} \cup \tilde{\mathsf{A}}) = \mathsf{F}(\mathsf{A}) \cup \mathsf{F}(\tilde{\mathsf{A}})$*;*
(5) $\mathsf{F}(\mathsf{A}) = pr_2(\mathsf{F} \cap (\mathsf{A} \times \mathsf{Y}))$ *and* $\mathsf{F}^{-1}(\mathsf{B}) = pr_1(\mathsf{F} \cap (\mathsf{X} \times \mathsf{B}))$*;*
(6) $\mathrm{dom}(\mathsf{F}) = \mathrm{range}(\mathsf{F}^{-1})$ *and* $\mathrm{dom}(\mathsf{F}^{-1}) = \mathrm{range}(\mathsf{F})$*;*
(7) *if* $\mathsf{F}$ *is definite then* $\mathsf{F}(\mathsf{F}^{-1}(\mathsf{B})) = \mathsf{B} \cap \mathrm{range}(\mathsf{F})$*;*
(8) *if* $\mathsf{F}$ *is injective then* $\mathsf{F}^{-1}(\mathsf{F}(\mathsf{A})) = \mathsf{A}$*;*
(9) $pr_1^{-1}(\mathsf{A}) = \mathsf{A} \times \mathsf{Y}$ *and* $pr_2^{-1}(\mathsf{B}) = \mathsf{X} \times \mathsf{B}$*;*
(10) $\mathsf{Z} \subseteq pr_1(\mathsf{Z}) \times pr_2(\mathsf{Z})$ *for every* $\mathsf{Z} \subseteq \mathsf{X} \times \mathsf{Y}$*.*

**Definition 1.14 (Composition).** Let $\mathsf{X}$, $\mathsf{Y}$, $\mathsf{Z}$ be any sets and let $\mathsf{F} \subseteq \mathsf{X} \times \mathsf{Y}$ and $\mathsf{G} \subseteq \mathsf{Y} \times \mathsf{Z}$ be any mappings. The mapping

$$\begin{aligned} \mathsf{G} \circ \mathsf{F} \ =_{\mathrm{df}}\ \mathsf{GF} =_{\mathrm{df}} \{(x, z) \mid (x, z) \in \mathsf{X} \times \mathsf{Z},\ &\text{there is a } y \in \mathsf{Y} \\ &\text{such that } (x, y) \in \mathsf{F},\ (y, z) \in \mathsf{G}\} \end{aligned}$$

is called the *composition* of $\mathsf{F}$ and $\mathsf{G}$.

The following proposition establishes fundamental basic properties of the composition of mappings:

**Proposition 1.2.** *Let* $\mathsf{X}$, $\mathsf{Y}$, $\mathsf{Z}$, $\mathsf{U}$ *be any sets and let* $\mathsf{F} \subseteq \mathsf{X} \times \mathsf{Y}$*,* $\mathsf{G} \subseteq \mathsf{Y} \times \mathsf{Z}$*, and* $\mathsf{H} \subseteq \mathsf{Z} \times \mathsf{U}$ *be any mappings. Then we have:*

(1) $(\mathsf{G} \circ \mathsf{F})(\mathsf{A}) = \mathsf{G}(\mathsf{F}(\mathsf{A}))$ *for every* $\mathsf{A} \subseteq \mathsf{X}$*;*
(2) $\mathrm{dom}(\mathsf{G} \circ \mathsf{F}) \subseteq \mathrm{dom}(\mathsf{F})$ *and equality holds if* $\mathrm{range}(\mathsf{F}) \subseteq \mathrm{dom}(\mathsf{G})$*;*
(3) $\mathsf{H} \circ (\mathsf{G} \circ \mathsf{F}) = (\mathsf{H} \circ \mathsf{G}) \circ \mathsf{F}$*;*

(4) $(G \circ F)^{-1} = F^{-1} \circ G^{-1}$.

*Proof.* To show Property (1) we consider any $A \subseteq X$. Then

$$\begin{aligned}(G \circ F)(A) &= \{z \mid z \in Z,\ \text{there is an } x \in A \text{ with } (x,z) \in G \circ F\}\\ &= \{z \mid z \in Z,\ \text{there are } x \in A,\ y \in Y \text{ with}\\ &\quad (x,y) \in F,\ (y,z) \in G\}\\ &= \{z \mid z \in Z,\ \text{there is a } y \in F(A) \text{ with } (y,z) \in G\}\\ &= G(F(A))\ ,\end{aligned}$$

and Property (1) is shown.

The proof of Properties (2) and (3) is left as an exercise.

To show Property (4) consider any $(z,x) \in (G \circ F)^{-1}$. Then $(x,z) \in G \circ F$, and thus there is a $y$ with $(x,y) \in F$ and $(y,z) \in G$. Consequently, $(y,x) \in F^{-1}$ and $(z,y) \in G^{-1}$. But this means that $(z,x) \in F^{-1} \circ G^{-1}$ and therefore we have $(G \circ F)^{-1} \subseteq F^{-1} \circ G^{-1}$.

The opposite inclusion is shown analogously. ∎

**Definition 1.15 (Family, Sequence).** Let $L$ and $X$ be non-empty sets. A (definite) mapping $\mathcal{F}\colon L \to X$ is often called a *family* of elements of $X$ with the index set $L$. We denote it by $\mathcal{F} = (x_\lambda)_{\lambda \in L}$ or just by $(x_\lambda)_{\lambda \in L}$.

If $L \subseteq \mathbb{N}_0$ then we call the family $(x_\lambda)_{\lambda \in L}$ a *sequence*.

Let $L' \subseteq L$; then we call the restriction of $\mathcal{F}\colon L \to X$ to $L'$ a *subfamily* of $\mathcal{F}$. In the case of sequences we then speak about a *subsequence*.

**Remark.** Note that we clearly distinguish between the family $\mathcal{F} = (x_\lambda)_{\lambda \in L}$ and the range of it, i.e., $\text{range}(\mathcal{F}) = \{x_\lambda \mid \lambda \in L\}$.

The notion of a family emphasizes the order and frequency of the elements.

Now we are in a position to deal with the numerosity of sets.

**Definition 1.16 (Cantor [29]).** Let $X$ and $Y$ be any sets.

(1) We say that $X$ and $Y$ are *equinumerous* (or have *the same cardinality*) if there is a bijection from $X$ to $Y$. Then we write $X \sim Y$.
(2) A set $X$ is said to be *finite* if $X = \emptyset$ or there is an $n \in \mathbb{N}$ such that $X \sim \{m \mid m \in \mathbb{N},\ m \leqslant n\}$.
(3) A set $X$ is said to be *countable* if $X \sim \mathbb{N}$.
(4) A set $X$ is said to be *at most countable* if $X$ is finite or countable.
(5) A set $X$ is said to be *uncountable* if it is not countable and not finite.

If a set is not finite then we also say it is *infinite*.

Note that equinumerosity is an *equivalence relation*. Let $S$ be any non-empty set. Then a binary relation $\sim$ is said to be an *equivalence relation* if it is reflexive, transitive, and symmetric. We say that $\sim$ is *symmetric* if $a \sim b$ implies $b \sim a$ for all $a, b \in S$.

**Proposition 1.3.**

(1) *For every non-empty finite set* $X$ *there is precisely one* $m \in \mathbb{N}$ *such that* $X \sim \{n \mid n \in \mathbb{N},\ n \leqslant m\}$.
(2) *Every infinite set contains a countable subset.*
(3) *A set* $X$ *is infinite iff there is a set* $Y \subset X$ *such that* $Y \sim X$.
(4) *Every countable set is not finite.*

*Proof.* We leave the proof of Property (1) as an exercise.

To show Property (2), let $X$ be an infinite set. Then there is an $x_1 \in X$ such that $X \setminus \{x_1\} \neq \emptyset$. We continue inductively. So for $n \in \mathbb{N}$ there must be an $x_{n+1} \in X \setminus \{x_1, \ldots, x_n\}$ with $X \setminus \{x_1, \ldots, x_{n+1}\} \neq \emptyset$, since otherwise $X$ would be finite. Consequently, for $\widetilde{X} =_{df} \{x_n \mid n \in \mathbb{N}\}$ we have $\widetilde{X} \subseteq X$ and $\widetilde{X}$ is countable.

We continue with Property (3). For the sufficiency, assume that there is a $Y \subset X$ such that $Y \sim X$. Then, by Property (1) we conclude that $X$ is not finite.

For the necessity, we distinguish the following cases:

*Case* 1. $X$ is countable.

Then the set $X = \{x_n \mid n \in \mathbb{N}\}$ is equinumerous to $Y = \{x_n \mid n \in \mathbb{N},\ n \geqslant 2\}$, since the mapping $x_i \mapsto x_{i+1}$, $i \in \mathbb{N}$, is a bijection.

*Case* 2. $X$ is not finite and uncountable.

By Property (2) we know that $X$ contains a countable subset, and we are back to Case 1. The details are left as an exercise.

To show Property (4), we suppose the converse. Then $\mathbb{N}$ would be finite. Hence there must be an $m \in \mathbb{N}$ such that $\mathbb{N} \sim \{n \mid n \in \mathbb{N},\ n \leqslant m\}$, a contradiction to (3) and Theorem 1.6 (Archimedes). ∎

**Lemma 1.2.** *Let* $X \subseteq \mathbb{R}$ *be any non-empty and finite set. Then there are uniquely determined* $a, b \in X$ *such that* $a \leqslant x \leqslant b$ *for all* $x \in X$.

*Proof.* We show the existence of $a$ and $b$ by induction on $n \in \mathbb{N}$, where we may assume $X \sim \{m \mid m \in \mathbb{N},\ m \leqslant n\}$ (cf. Definition 1.16).

For $n = 1$ we have $X = \{x_1\}$ and so $a = b =_{df} x_1$ satisfy the lemma.

The induction step is from $n$ to $n+1$. Assume $X \sim \{m \mid m \in \mathbb{N},\ m \leqslant n+1\}$ then $X = \{x_1, \ldots, x_n, x_{n+1}\}$. We apply the induction hypothesis to $X \setminus \{x_{n+1}\}$. Hence, there are uniquely determined $\widetilde{a}$, $\widetilde{b} \in X \setminus \{x_{n+1}\}$ with $\widetilde{a} \leqslant x \leqslant \widetilde{b}$ for all $x \in X \setminus \{x_{n+1}\}$. We set $a =_{df} \min\{\widetilde{a}, x_{n+1}\}$ and $b =_{df} \max\left\{\widetilde{b}, x_{n+1}\right\}$, and the lemma is shown. ∎

Next we ask whether or not the set of all rational numbers is countable. To answer this question some preparations are necessary. In particular, we shall prove that $\mathbb{N} \times \mathbb{N}$ is countable. To establish this result we need the *Gaussian summation formula.* It says that

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \ . \tag{1.21}$$

This is shown inductively. The induction basis is for $n = 1$. We thus have

$$\begin{aligned}\sum_{i=1}^{1} i &= 1 \quad \text{(by (1.6))}\\ &= \frac{1(1+1)}{2} \ ,\end{aligned}$$

and the induction basis is shown.

Next, we assume the induction hypothesis for $n$, i.e., $\sum\limits_{i=1}^{n} i = \frac{n(n+1)}{2}$, and have to perform the induction step from $n$ to $n+1$. We obtain

$$\begin{aligned}\sum_{i=1}^{n+1} i &= \sum_{i=1}^{n} i + (n+1) \quad \text{( by (1.7))}\\ &= \frac{n(n+1)}{2} + (n+1) \quad \text{(by the induction hypothesis)}\\ &= \frac{n(n+1)}{2} + \frac{2(n+1)}{2} \ = \ \frac{n^2+n+2n+2}{2}\\ &= \frac{(n+1)(n+2)}{2} \ ,\end{aligned}$$

where the last two steps have been performed by using the distributive laws (cf. Definition 1.3, Part (3)).

It should be noted that the inductive proof given above is formally correct and sufficient to establish the Gaussian summation formula. However, the proof does not tell us anything about how the Formula (1.21) might have been found. So let us elaborate this point. We may write the numbers to be summed up in two ways, i.e., in increasing order and in decreasing order (cf. Figure 1.3). Then we take the sum in each column which is always $n+1$.

$$\begin{array}{ccccc} 1 & 2 & \ldots & n-1 & n\\ n & n-1 & \ldots & 2 & 1\\ \hline n+1 & n+1 & \ldots & n+1 & n+1 \end{array}$$

Fig. 1.3: The numbers from 1 to $n$ in increasing order and in decreasing order

Since we have $n$ columns, the sum of the two rows is therefore $n(n+1)$. Taking into account that every number appears exactly twice, the desired sum is just $n(n+1)/2$. However, this formula was known long before Carl Friedrich Gauss rediscovered it at the age of nine by using the technique displayed in

Figure 1.3 when his teacher requested all pupils to sum the numbers from 1 to 100 as reported by Wolfgang Sartorius von Waltershausen [154, Page 12, 13].

If one has made such a discovery, one should try to figure out whether or not the technique used generalizes to related problems. Thus the reader is encouraged to determine the sum of the first $n$ odd numbers and the sum of the first $n$ even numbers, i.e., to solve the following exercise:

**Exercise 1.12.** *Determine the following sums and prove inductively the results obtained:*

(1) $\sum_{i=1}^{n}(2i-1)$, *and*

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

A more challenging problem is to add consecutive powers. A flavor of this problem is provided by the following exercise:

**Exercise 1.13.** *Prove inductively the following formulae:*

(1) $\sum_{i=0}^{n} i^2 = n(n+1)(2n+1)/6$, *and*

(2) $\sum_{i=0}^{n} i^3 = \left(\sum_{k=0}^{n} k\right)^2$ .

**Theorem 1.13 (Cantor [28, 29]).**

(1) *Every subset of* $\mathbb{N}$ *is at most countable.*
(2) *The set* $\mathbb{N} \times \mathbb{N}$ *is countable.*

*Proof.* Let $X \subseteq \mathbb{N}$ be an infinite set. We show that $X$ is countable. Therefore, we define a mapping $f\colon \mathbb{N} \to X$ by setting $f(n) =_{df} x_n$ for all $n \in \mathbb{N}$, where $x_1 =_{df} \min X$ and $x_{n+1} =_{df} \min(X \setminus \{x_1, \ldots, x_n\})$. By Theorem 1.7, this definition is admissible. It remains to prove $f$ is bijective.

By the inductive definition of $f$ we have $x_i < x_{i+1}$ for all $i \in \mathbb{N}$. Thus, $f$ is injective.

In order to see that $f$ is surjective, let $a \in X$ be arbitrarily fixed. If $a = x_1$ then $f(1) = a$. Let $a > x_1$, and $m =_{df} \max\{n \mid n \in \mathbb{N},\ x_n < a\}$ (cf. Lemma 1.2). Then $f(m+1) = a$ and so for every $a \in X$ there is an $n \in \mathbb{N}$ such that $f(n) = a$, and Property (1) is shown.

To show Property (2) let us arrange $\mathbb{N} \times \mathbb{N}$ in an array as shown in Figure 1.4, where row $x$ contains all pairs $(x, y)$, i.e., having $x$ in the first component and $y = 1, 2, 3 \ldots$.

The resulting bijection $c$ is shown in Figure 1.5; that is, we arrange all these pairs in a sequence starting

$$(1,1),\ (1,2),\ (2,1),\ (1,3),\ (2,2),\ (3,1),\ (1,4),\ (2,3),\ldots\ . \tag{1.22}$$

$$\begin{array}{llllll}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & \ldots \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & \ldots \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & \ldots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & \ldots \\
(5,1) & \ldots & & & & \\
\ldots & \ldots & & & &
\end{array}$$

Fig. 1.4: A two-dimensional array representing $\mathbb{N} \times \mathbb{N}$

| $m \backslash n$ | 1 | 2 | 3 | 4 | 5 | 6... |
|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 4 | 7 | 11 | ↙ |
| 2 | 3 | 5 | 8 | 12 | ↙ | |
| 3 | 6 | 9 | 13 | ↙ | | |
| 4 | 10 | 14 | ↙ | | | |
| 5 | 15 | ↙ | | | | |
| 6 | ↙ | | | | | |
| ⋮ | | | | | | |

Fig. 1.5: The bijection $c$

In this order, all pairs $(m, n)$ appear before all pairs $(m', n')$ if and only if $m + n < m' + n'$. So they are arranged in order of incrementally growing component sums. The pairs with the same component sum are ordered by the first component, starting with the smallest one. That is, pair $(1, 1)$ is the only one in the first segment, and pair $(m, n)$, $m + n > 2$, is located in the segment

$$(1, m+n-1),\ (2, m+n-2),\ \ldots,\ (m, n),\ \ldots,\ (m+n-1, 1)\ . \tag{1.23}$$

Note that there are $m+n-1$ many pairs having the component sum $m+n$. Thus, in front of pair $(1, m+n-1)$ in the Sequence 1.22 we have $m+n-2$ many segments containing a total of $1 + \cdots + (m+n-2)$ many pairs. Using Equation (1.21) we formally define the desired bijection $c \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ as

$$\begin{aligned}
c(m, n) = m + \sum_{i=1}^{m+n-2} i \ &= \ x + \frac{(m+n-2)(m+n-1)}{2} \qquad \text{(by (1.21))} \\
&= \frac{(m+n)^2 - m - 3n + 2}{2}\ . 
\end{aligned} \tag{1.24}$$

Note that we start counting with 1 in the Sequence (1.22), since otherwise we would not obtain a bijection (see Figure 1.5). Consequently, we have shown that $\mathbb{N} \times \mathbb{N}$ is countable. ∎

The bijection $\mathsf{c}$ is called *Cantor's pairing function*. For more information concerning this pairing function we refer the reader to [196].

Theorem 1.13 allows for several further results. The first one is recommended as an exercise.

**Exercise 1.14.** *Show that for every fixed* $k \in \mathbb{N}$, $k > 2$, *there is a bijection* $c_k \colon \mathbb{N}^k \to \mathbb{N}$, *where* $\mathsf{N}^k = \bigtimes_{i=1}^{k} \mathbb{N}$.

Furthermore, now we are in a position to show the following result:

**Theorem 1.14.**

(1) *Let* $X$ *and* $Y$ *be at most countable sets. Then* $X \cup Y$ *is at most countable.*
(2) *If* $X$ *is a countable set and* $f$ *a mapping from* $X$ *onto* $Y$ *then the set* $Y$ *is at most countable.*
(3) *Let* $L$ *be an index set which is at most countable. Furthermore, assume that for all* $\lambda \in L$ *the sets* $X_\lambda$ *are at most countable. Then the set* $\widetilde{X} = \bigcup_{\lambda \in L} X_\lambda$ *is at most countable.*
(4) *Let* $X$ *be any uncountable set and let* $Y \subseteq X$ *be at most countable. Then* $X$ *and* $X \setminus Y$ *are equinumerous.*

*Proof.* Property (1) is obvious if both $X$ and $Y$ are finite or if one set is finite and the other set is countable. So, it remains to consider the case where both sets $X$ and $Y$ are countable. Now, if $X \setminus Y$ is finite then we are again done. Otherwise, it is easy to see that $X \setminus Y$ is countable, too.

Therefore, let $X \setminus Y = \{x_1, x_2, x_3, \ldots\}$ and $Y = \{y_1, y_2, y_3, \ldots\}$. Then we define $f(2n) = y_n$ and $f(2n-1) = x_n$ for all $n \in \mathbb{N}$. Clearly, $f$ is a bijection between $\mathbb{N}$ and $X \cup Y$, and Property (1) is shown.

Property (2) is shown as follows: Since the set $X$ is countable, we can write the set $X$ as $X = \{x_1, x_2, x_3, \ldots\}$. Consider the mapping $g \colon \mathbb{N} \to Y$ defined as $g(n) =_{df} f(x_n)$ for all $n \in \mathbb{N}$. By assumption we conclude that $g$ is surjective. Hence, we can define the following mapping $h \colon Y \to \mathbb{N}$, where

$$h(y) =_{df} \min\{n \mid n \in \mathbb{N},\ g(n) = y\} \qquad \text{for all } y \in Y\ .$$

Note that by construction we have $S_y =_{df} \{n \mid n \in \mathbb{N},\ g(n) = y\} \neq \emptyset$ for every $y \in Y$. Therefore, by Theorem 1.7 we know that $S_y$ possesses a minimal element for every $y \in Y$ and so $h$ is well defined. Furthermore, by construction we have $g(h(y)) = y$ for all $y \in Y$ and we know that $h$ is definite.

We claim that the mapping $h$ is injective. Let $y_1, y_2 \in Y$ be any elements such that $h(y_1) = h(y_2)$. We have to show that $y_1 = y_2$. Suppose to the contrary that $y_1 \neq y_2$. Without loss of generality we can assume that $y_1 < y_2$. Let $n_i =_{df} h(y_i)$, $i = 1, 2$, then we directly obtain that

$$y_1 = g(n_1)\ <\ y_2\ =\ g(n_2)\ .$$

Since $h(y_1) = h(y_2)$, we also have $g(n_1) = g(h(y_1)) = g(h(y_2)) = g(n_2)$, a contradiction. Consequently, $h$ is injective and thus a bijection from $Y$ to $h(Y)$.

Since $h(Y) \subseteq \mathbb{N}$ we conclude that $h(Y)$ is at most countable (cf. Theorem 1.13, Assertion (1)). Therefore, $Y$ is at most countable, too, and Property (2) is shown.

We continue with Property (3). If $L$ is finite then we can directly apply Property (1) a finite amount of times and thus $\widetilde{X}$ is at most countable. If $L$ is countable then we can write $L = \{\lambda_1, \lambda_2, \lambda_3, \ldots\}$. By assumption we know that for every $\lambda_n \in L$ the set $X_{\lambda_n}$ is either finite or countable. Hence, there is a surjective mapping $m \mapsto x^{(m)}_{\lambda_n}$ such that $X_{\lambda_n} = \{x^{(1)}_{\lambda_n}, \ldots, x^{(m)}_{\lambda_n}, \ldots\}$. So we define a mapping $F\colon \mathbb{N} \times \mathbb{N} \to \widetilde{X}$ by setting

$$F(m, n) =_{df} x^{(m)}_{\lambda_n} \quad \text{for all } m, n \in \mathbb{N}\ .$$

By construction the mapping $F$ is surjective. Since $\mathbb{N} \times \mathbb{N}$ is countable (cf. Theorem 1.13, Assertion (2)), we conclude by Assertion (1) of Theorem 1.13 that $\widetilde{X}$ is at most countable.

Finally, we show Property (4). If $Y$ is finite, the assertion is obvious. So let $Y$ be countable. Then $X \setminus Y$ must be uncountable, since otherwise, by Property (1), we would directly obtain that $(X \setminus Y) \cup Y = X$ is countable, a contradiction.

By Proposition 1.3, Assertion (2), there is a countable subset $Y_1$ of $X \setminus Y$. We set $Z =_{df} (X \setminus Y) \setminus Y_1$ and obtain

$$X = Z \cup (Y \cup Y_1) \quad \text{and} \quad X \setminus Y = Z \cup Y_1.$$

We define a mapping $f\colon X \to X \setminus Y$ by setting $f(x) = x$ for all $x \in Z$ and such that $f|_{Y \cup Y_1}\colon Y \cup Y_1 \to Y_1$ is bijective. This is possible, since both $Y \cup Y_1$ and $Y_1$ are countable.

Consequently, $f$ is bijective by construction, and so $X \sim X \setminus Y$. ∎

Theorems 1.13 and 1.14 directly allow for the following corollary:

**Corollary 1.5.** *The set of all rational numbers is countable.*

*Proof.* First, we consider the set $\{r \mid r \in \mathbb{Q},\ r > 0\}$ and the following mapping $f\colon \mathbb{N} \times \mathbb{N} \to \{r \mid r \in \mathbb{Q},\ r > 0\}$ defined as $f(m, n) =_{df} m/n$ for all $m, n \in \mathbb{N}$. By the definition of $\mathbb{Q}$ we conclude that $f$ is surjective. Since $\mathbb{N} \times \mathbb{N}$ is countable (cf. Theorem 1.13, Assertion (2)), we know by Assertion (1) of Theorem 1.13 that $\{r \mid r \in \mathbb{Q},\ r > 0\}$ is at most countable. Since $\mathbb{N} \subseteq \{r \mid r \in \mathbb{Q},\ r > 0\}$ the set $\{r \mid r \in \mathbb{Q},\ r > 0\}$ must be countable.

Analogously one shows that $\{r \mid r \in \mathbb{Q},\ r < 0\}$ is countable. Thus, we can apply Theorem 1.14, Assertion (1), twice and see that

$$\mathbb{Q} = \{r \mid r \in \mathbb{Q},\ r > 0\} \cup \{0\} \cup \{r \mid r \in \mathbb{Q},\ r < 0\}$$

is countable. ∎

So it remains to clarify whether or not the set of all real numbers and the set of all irrational numbers are countable.

**Theorem 1.15 (Cantor [28, 29]).** *The set* $]0,1[\, \subset \mathbb{R}$ *is uncountable.*

*Proof.* Note that $]0,1[$ is not finite, since $1/n \in\, ]0,1[$ for every $n \in \mathbb{N}$. Suppose the converse, i.e., $]0,1[$ is countable. Then we can write $]0,1[= \{x_1, x_2, x_3, \ldots\}$. In accordance with Theorem 1.12 we take for every $x_i$, $i \in \mathbb{N}$, its decimal representation and obtain $x_i = .z_{i1}z_{i2}z_{i3}\cdots$, where $z_{ij} \in \{0, 1, \ldots, 9\}$ for all $j \in \mathbb{N}$.

Next, we chose for all $i \in \mathbb{N}$ numbers $z_i$, where $z_i \in \{1, 2, \ldots, 8\} \setminus \{z_{ii}\}$. For $k = 1$ and $m = 10$ and the chosen $z_i$ we know by Lemma 1.1 that there exists a uniquely determined $x \in \mathbb{R}$, $x \geqslant 0$, such that $x \in \bigcup_{n\in\mathbb{N}}[a_n, b_n]$, namely $x = \sup\{\sum_{i=1}^{n} z_i 10^{k-i} \mid n \in \mathbb{N}\}$. By Theorem 1.12 this $x$ has the decimal representation $.z_1z_2z_3\cdots$.

Due to our construction we conclude that $x \in\, ]0,1[$ (this is the reason we excluded 0 and 9 as possible choices for $z_i$). Furthermore, $x \neq x_i$ for all $i \in \mathbb{N}$, since the decimal representation is unique for every real number $x \geqslant 0$. But this is a contradiction to our supposition. ∎

Theorems 1.15 and 1.14 directly yield the following corollary:

**Corollary 1.6.** *The sets* $\mathbb{R}$ *and* $\mathbb{R} \setminus \mathbb{Q}$ *are equinumerous.*

So, there are many more irrational numbers than rational ones, since the rational numbers are countable and the irrational numbers are uncountable.

**Exercise 1.15.** *Generalize Theorem 1.13 as follows. If* $A_1$ *and* $A_2$ *are at most countable then* $A_1 \times A_2$ *is at most countable.*

**Exercise 1.16.** *Show that* $\wp(\mathbb{N})$ *is uncountable.*

Note that Theorem 1.15 is in some sense much deeper than Theorem 1.13. Of course, it is very important to ask whether or not $\mathbb{N} \times \mathbb{N}$ is still countable. But once asked, it is not too difficult to establish the countability of $\mathbb{N} \times \mathbb{N}$. On the other hand, it is Theorem 1.15 that makes the subject of countability interesting, since it establishes the existence of a well-defined set which is *not* countable. This came as a big surprise. Furthermore, the proof technique used in the demonstration of Theorem 1.15 turned out to be of major importance and has found numerous applications. It is usually referred to as a *diagonalization argument* or the *diagonal method.* In the form used above it was invented by Cantor [31].

**Exercise 1.17.** *Provide a bijection* $b\colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$.

## 1.7 Linear Spaces

Next, we introduce further sets that are important for the further development of mathematical analysis.

We set $\mathbb{R}^m =_{\text{df}} \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{m \text{ times}}$ for every $m \in \mathbb{N}$.

So, every $x \in \mathbb{R}^m$ can be written as $(x_1, \ldots, x_m)$, and we refer to $x_i$ as the $i$th component of $x$.

Next, we define addition for elements of $\mathbb{R}^m$, i.e., $+\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$, where

$$x + y =_{\text{df}} (x_1 + y_1, \ldots, x_m + y_m) \quad \text{for all } x, y \in \mathbb{R}^m .$$

Note that the addition of the $i$th components, i.e., $x_i + y_i$, $i = 1, \ldots, m$, is the usual addition in $\mathbb{R}$.

Furthermore, we define a multiplication $\cdot : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ as follows: Let $\alpha \in \mathbb{R}$, and let $x \in \mathbb{R}^m$; then we set

$$\alpha \cdot x =_{\text{df}} (\alpha x_1, \ldots, \alpha x_m) .$$

Note that the multiplication of the $i$th components, i.e., $\alpha x_i$, is the usual multiplication in $\mathbb{R}$.

The following proposition summarizes basic properties:

**Proposition 1.4.** *Let $m \in \mathbb{N}$ be arbitrarily fixed.*

(1) $(\mathbb{R}^m, +)$ *is an Abelian group with neutral element* $(0, \ldots, 0)$*;*
(2) $1 \cdot x = x$ *for all* $x \in \mathbb{R}^m$*;*
(3) $(\alpha + \beta)x = \alpha \cdot x + \beta \cdot x$ *for all* $\alpha, \beta \in \mathbb{R}$ *and all* $x \in \mathbb{R}^m$*;*
(4) $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$ *for all* $\alpha \in \mathbb{R}$ *and all* $x, y \in \mathbb{R}^m$*;*
(5) $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$ *for all* $\alpha, \beta \in \mathbb{R}$ *and all* $x \in \mathbb{R}^m$*.*

The proof of Proposition 1.4 is left as an exercise.

Note that in the following we shall usually omit the multiplication dot, i.e., we shortly write $\alpha x$ instead of $\alpha \cdot x$.

Now we are in a position to define the fundamental notions of a vector space, also called a linear space and related notions such as the scalar product, Euclidean norm, and Euclidean distance.

Formally, this is done as follows:

**Definition 1.17 ($m$-Dimensional Linear Space).**

(1) We call $(\mathbb{R}^m, +, \cdot\,)$ the $m$-*dimensional linear space* or $m$-*dimensional vector space.*
(2) The mapping $\langle \cdot\,, \cdot \rangle \colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ defined as

$$\langle x, y \rangle =_{\text{df}} \sum_{i=1}^{m} x_i y_i$$

for all $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ and $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ is called the *scalar product* on $\mathbb{R}^m$.

(3) The mapping $\|\cdot\| : \mathbb{R}^m \to \mathbb{R}$ defined as $\|x\| =_{df} \langle x, x\rangle^{1/2}$ for all $x \in \mathbb{R}^m$ is said to be the *Euclidean norm* on $\mathbb{R}^m$.

(4) The number $\|x - y\|$ is called the *Euclidean distance* of $x$ and $y$ in $\mathbb{R}^m$.

(5) We call $(\mathbb{R}^m, \|\cdot\|)$ the $m$-*dimensional Euclidean space.*

We shall generalize the notion of a linear space later by using any Abelian group $(X, +)$ and by defining a multiplication of the elements of $X$ with the elements of a field $\mathbb{F}$ in a way such that the Assertions (2) through (5) of Proposition 1.4 are satisfied.

Occasionally we shall use the *canonical basis* of $\mathbb{R}^m$, which we define as follows: For $i = 1, \ldots, m$, let $e_i =_{df} (0, \ldots, 0, 1, 0, \ldots, 0)$, where the $i$th component is 1. We refer to the $e_i$ as *canonical unit vectors.* Then for all $x \in \mathbb{R}^m$ we have

$$x = \sum_{i=1}^{m} x_i e_i \ . \tag{1.25}$$

Note that we use 0 to denote the neutral element in $\mathbb{R}$ and in $\mathbb{R}^m$. So, in the first case, 0 denotes 0, while in the second case it stands for $(0, \ldots, 0) \in \mathbb{R}^m$. This is a notational overload, but it will be clear from the context what is meant.

Next we show a famous and very helpful inequality found by Cauchy [32] and in a more general form by Bunyakovsky [23]. Schwarz [166] rediscovered it without being aware of Bunyakovsky's work. It is widely known as the *Cauchy–Schwarz inequality.*

**Theorem 1.16 (Cauchy–Schwarz Inequality).** *For all* $x, y \in \mathbb{R}^m$ *we have* $|\langle x, y\rangle| \leqslant \|x\|\,\|y\|$. *Equality holds if and only if there are* $\alpha, \beta \in \mathbb{R}$ *with* $(\alpha, \beta) \neq (0, 0)$ *such that* $\alpha x + \beta y = 0$.

*Proof.* For all $x \in \mathbb{R}^m$, if $\|x\| = 0$ then $\sum_{i=1}^{m} x_i^2 = 0$. But this can only happen iff $x_i = 0$ for $i = 1, \ldots, m$ (cf. Corollary 1.1). Consequently, we see that $\sum_{i=1}^{m} x_i y_i = 0$, and thus $\langle x, y\rangle = 0$. This proves the case that $\|x\| = 0$.

Next, let $\|x\| \neq 0$, and let $\alpha, \beta \in \mathbb{R}$. By Corollary 1.1 we obtain

$$\begin{aligned} 0 &\leqslant \sum_{i=1}^{m} (\alpha x_i + \beta y_i)^2 && (1.26)\\ &= \alpha^2 \sum_{i=1}^{m} x_i^2 + 2\alpha\beta \sum_{i=1}^{m} x_i y_i + \beta^2 \sum_{i=1}^{m} y_i^2 \\ &= \alpha^2 \|x\|^2 + 2\alpha\beta\langle x, y\rangle + \beta^2 \|y\|^2 \ . && (1.27) \end{aligned}$$

We set $\alpha =_{df} -\langle x, y\rangle / \|x\|$ and $\beta = \|x\|$. Then (1.26) and (1.27) directly yield that

$$0 \leqslant \langle x, y\rangle^2 - 2\langle x, y\rangle^2 + \|x\|^2 \, \|y\|^2 \ , \tag{1.28}$$

i.e., the desired inequality.

Finally, let $\langle x, y\rangle^2 = \|x\|^2 \, \|y\|^2$. The equality is trivial if $x = 0$ or $y = 0$ and then we can choose any $\alpha, \beta \in \mathbb{R}$ with $(\alpha, \beta) \neq (0, 0)$.

So let $x \neq 0 \neq y$. Then we have $\|x\| \neq 0 \neq \|y\|$. By (1.26), we see that equality holds iff

$$\begin{aligned}
0 &= \sum_{i=1}^{m} (\alpha x_i + \beta y_i)^2 \quad \text{iff} \\
0 &= \sum_{i=1}^{m} \left( -\frac{\langle x, y\rangle}{\|x\|} x_i + \|x\| \, y_i \right)^2 \quad \text{iff} \\
0 &= \sum_{i=1}^{m} \left( -\langle x, y\rangle x_i + \|x\|^2 \, y_i \right)^2 \ .
\end{aligned}$$

Since this is a sum of squares, each summand must be 0; i.e., we must have $-\langle x, y\rangle x_i + \|x\|^2 \, y_i = 0$. Thus, $\alpha x + \beta y = 0$ with $\alpha = -\langle x, y\rangle$ and $\beta = \|x\|^2$, and so $(\alpha, \beta) \neq (0, 0)$. ∎

Now we are ready to establish the fundamental properties of the Euclidean norm in $\mathbb{R}^m$.

**Theorem 1.17.** *Let $m \in \mathbb{N}$ be arbitrarily fixed. Then for all $x$, $y \in \mathbb{R}^m$ and all $\alpha \in \mathbb{R}$ we have:*

(1) $\|x\| \geqslant 0$ *and* $\|x\| = 0$ *iff* $x = 0$*;*
(2) $\|\alpha x\| = |\alpha| \, \|x\|$*;*
(3) $\|x + y\| \leqslant \|x\| + \|y\|$*;*
(4) $|\|x\| - \|y\|| \leqslant \|x - y\|$*.*

Note that Property (3) is called the *triangle inequality* or *Minkowski's inequality* in honor of Hermann Minkowski [124], who primarily pushed the study of norms other than the Euclidean one in finite-dimensional spaces.

*Proof.* By definition we have $\|x\| = \left(\sum_{i=1}^{m} x_i^2\right)^{1/2}$, and thus Properties (1) and (2) obviously hold.

Let $x$, $y \in \mathbb{R}^m$, then Property (3) is shown as follows:

$$\begin{aligned}
\|x + y\|^2 = \sum_{i=1}^{m} (x_i + y_i)^2 \ &= \ \sum_{i=1}^{m} x_i^2 + 2 \sum_{i=1}^{m} x_i y_i + \sum_{i=1}^{m} y_i^2 \\
&= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \\
&\leqslant \|x\|^2 + 2\,|\langle x, y\rangle| + \|y\|^2 \\
&\leqslant \|x\|^2 + 2\,\|x\| \; \|y\| + \|y\|^2 \quad \text{(by Theorem 1.16)} \\
&= (\|x\| + \|y\|)^2 \ ,
\end{aligned}$$

and Property (3) is proved.

By Property (3) we have $\|x\| = \|x - y + y\| \leqslant \|x - y\| + \|y\|$. Hence,

$$\begin{aligned} |\|x\| - \|y\|| &\leqslant |\|x - y\| + \|y\| - \|y\|| \\ &= |\|x - y\|| \ = \ \|x - y\| \ , \end{aligned}$$

where the last line holds because of Property (1) and the definition of the absolute value. ∎

**Definition 1.18 (Norm).** Any functional $\|\cdot\| : \mathbb{R}^m \to \mathbb{R}$ satisfying Properties (1) through (3) of Theorem 1.17 is called a *norm*.

On $\mathbb{R}^m$ one can define many more functionals $\|\cdot\| : \mathbb{R}^m \to \mathbb{R}$ satisfying the conditions of Definition 1.18. We mention here some famous examples.

Let $p \in \mathbb{R}$, $p \geqslant 1$; then we define for all $x \in \mathbb{R}^m$ the so-called $p$-*norm* by

$$\|x\|_p =_{df} \left( \sum_{i=1}^{m} |x_i|^p \right)^{1/p} \ . \tag{1.29}$$

Note that the so far considered Euclidean norm is then $\|\cdot\|_2$.

For $p = 1$ we obtain $\|x\|_1 = \sum_{i=1}^{m} |x_i|$ (the *sum norm*).

Another important example is $\|x\|_\infty =_{df} \max_{i=1,\ldots,m} |x_i|$, the so-called *maximum norm*.

Note that for $m = 1$, i.e., in $\mathbb{R}$, all these norms coincide and are equal to the absolute value (cf. Proposition 1.1).

**Exercise 1.18.** *Show that the conditions of Definition 1.18 are satisfied for the functionals $\|\cdot\|_\infty$ and $\|\cdot\|_1$ defined above.*

Figure 1.6 and Figure 1.7 show the set $U_1$ and $U_\infty$ of all points in $x \in \mathbb{R}^2$ such that $\|x\|_1 = 1$ and $\|x\|_\infty = 1$, respectively. We refer to these sets as the *unit circle*.

Of course, the definition of the unit circle generalizes to any norm $\|\cdot\|$; i.e., we then define $U =_{df} \{x \mid x \in \mathbb{R}^2 \ , \ \|x\| = 1\}$.

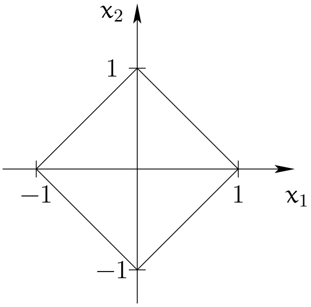

Fig. 1.6: The unit circle $U_1$

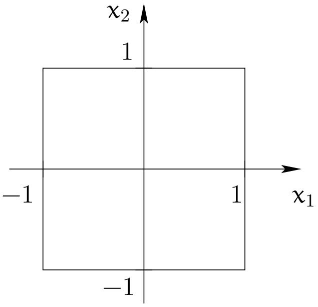

Fig. 1.7: The unit circle $U_\infty$

**Exercise 1.19.** *Draw the unit circle for the Euclidean norm.*

## 1.8 Complex Numbers

Historically, the complex numbers were introduced to extend the root function to all real numbers.

From an algebraic point of view it is interesting to ask whether or not we can define on $\mathbb{R} \times \mathbb{R}$ addition and multiplication in a way such that we obtain a *field*. The affirmative answer is provided below.

We define $+, \cdot : (\mathbb{R} \times \mathbb{R}) \times (\mathbb{R} \times \mathbb{R}) \to \mathbb{R} \times \mathbb{R}$, i.e., the operations addition and multiplication, respectively, as follows: For all $(a, b), (c, d) \in \mathbb{R} \times \mathbb{R}$ let

$$(a,b) + (c,d) =_{\mathrm{df}} (a+c, b+d) \; , \tag{1.30}$$

$$(a,b) \cdot (c,d) =_{\mathrm{df}} (ac - bd, ad + bc) \; . \tag{1.31}$$

Note that addition is defined as before (cf. Section 1.7).

We set $\mathbb{C} =_{\mathrm{df}} (\{(x,y) \mid x,\, y \in \mathbb{R}\}, +, \cdot\,)$.

**Theorem 1.18.** *The structure $\mathbb{C}$ is an Abelian field with neutral element $(0,0)$ and identity element $(1,0)$.*

*Proof.* By Proposition 1.4, we already know that $(\{(x,y) \mid x,\, y \in \mathbb{R}\}, +)$ is an Abelian group with neutral element $(0,0)$.

By its definition, the operation $\cdot$ is *commutative*. An easy calculation shows that it is also *associative*.

Furthermore, by (1.31) we directly obtain that

$$(x,y) \cdot (1,0) = (x - 0, 0 + y) = (x,y) \quad \text{for all } x, y \in \mathbb{R} \; .$$

Thus, $(1,0)$ is the identity element.

If $z = (x,y) \neq (0,0)$ then

$$\begin{aligned} (x,y) \cdot \left( \frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2} \right) &= \left( \frac{x^2}{x^2+y^2} + \frac{y^2}{x^2+y^2}, \frac{-xy}{x^2+y^2} + \frac{xy}{x^2+y^2} \right) \\ &= (1,0) \; , \end{aligned}$$

and thus the inverse element $1/z$ of $z$ exists.

It remains to show the distributive laws. Let $(x,y), (u,v)$, and $(w,z)$ be arbitrarily fixed. Then we have

$$\begin{aligned} (x,y) \cdot ((u,v) + (w,z)) &= (x,y) \cdot (u+w, v+z) \\ &= (x(u+w) - y(v+z), x(v+z) + y(u+w)) \\ &= (xu - yv, xv + yu) + (xw - yz, xz + yw) \\ &= (x,y) \cdot (u,v) + (x,y) \cdot (w,z) \; . \end{aligned}$$

The remaining distributive law is shown analogously. ∎

We call $\mathbb{C}$ the field of the *complex numbers*.

**Remarks.**

(a) All calculation rules for real numbers that result directly from the field properties of $(\mathbb{R}, +, \cdot\,)$ can be translated to $\mathbb{C}$. Whenever order is involved, special care has to be taken, since the relation $\leqslant$ is *not* defined for complex numbers.

(b) Consider the subset $\{(x,0) \mid x \in \mathbb{R}\}$ of $\mathbb{C}$. By definition we have

$$\begin{aligned}(x,0) + (y,0) &= (x+y,0)\ ,\\ (x,0)\cdot(y,0) &= (xy,0)\ .\end{aligned}$$

Thus, from the viewpoint of algebraic structures, $x \in \mathbb{R}$ and $(x,0) \in \mathbb{C}$ can be identified and so can $\mathbb{R}$ and $\{(x,0) \mid x \in \mathbb{R}\}$. In this sense, $\mathbb{C}$ is an *extension* of $\mathbb{R}$.

(c) Euler (1777) introduced $i =_{\mathrm{df}} (0,1)$ (*imaginary unit*). Using $i$ we can represent every $z = (x,y) \in \mathbb{C}$ as

$$\begin{aligned} z = (x,y) &= (x,0) + (0,y) = (x,0) + y(0,1)\\ &= (x,0) + (y,0)\cdot(0,1) = (x,0) + (y,0)\cdot i\\ &= x + yi \quad \text{(cf. (b))}\ . \end{aligned} \tag{1.32}$$

We call $x$ the *real part* and $y$ the *imaginary part* of $z$, denoted by $\mathfrak{R}(z)$ and $\mathfrak{I}(z)$, respectively. Note that for $x = 0$ and $y \neq 0$ we shall shortly write $z = yi$ instead of $z = 0 + yi$. Furthermore, we shall use $-z = (-x,-y)$ to denote the inverse of $z$ with respect to addition, since $z + (-z) = (0,0)$. We define the *complex conjugate* $\overline{z}$ of $z$; let $z = (x,y)$ then $\overline{z} =_{\mathrm{df}} (x,-y)$; i.e., the complex conjugate of $x + yi$ is $x - yi$.

We continue with further definitions that will be needed later.

Inequality of complex numbers $z_1, z_2$ is defined as follows: Let $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$, then we say that $z_1 \neq z_2$ if $x_1 \neq x_2$ or $y_1 \neq y_2$.

Consequently, for all $z_1, z_2 \in \mathbb{C}$ we have either $z_1 = z_2$ or $z_1 \neq z_2$. Note that the inequality relation is *symmetric*, but it is neither transitive nor reflexive.

Furthermore, we define *powers* of complex numbers with integer exponents as follows: For all $z \in \mathbb{C}$ and all $n \in \mathbb{N}$ we set $z^1 =_{\mathrm{df}} z$ and $z^{n+1} =_{\mathrm{df}} z \cdot z^n$. For $z \in \mathbb{C}$ such that $z \neq 0$ we define $z^{-n} =_{\mathrm{df}} 1/z^n$ for all $n \in \mathbb{N}$ as well as $z^0 = 1$.

For example, by the definition just made we see that $i^{-1}$ is the inverse of $i$. By the proof of Theorem 1.18 we already know how to compute the inverse of $i$; i.e., recalling that $i = (0,1)$, we have

$$\begin{aligned}\frac{1}{i} &= \left(\frac{0}{0^2+1^2}, \frac{-1}{0^2+1^2}\right)\\ &= (0,-1) = -i\ ,\end{aligned}$$

where we used the second convention made in Part (c) of the remarks above. Let us also compute $i^2$ by using (1.31). We directly obtain

$$i^2 = (0,1)\cdot(0,1) = (-1,0) = -1\ , \tag{1.33}$$

where we used the first convention made in Part (c) of the remarks above. Now, it is easy to see that $i^3 = -i$ and $i^4 = 1$.

Moreover, having the imaginary unit allows for a convenient way to perform multiplication and division of complex numbers. In order to see this, let $z_1 = x_1 + y_1 i$ and $z_2 = x_2 + y_2 i$. Then we have

$$\begin{aligned} z_1 \cdot z_2 &= (x_1 + y_1 i)(x_2 + y_2 i) \\ &= x_1x_2 + x_1y_2 i + y_1y_2 i^2 + y_1x_2 i \\ &= x_1x_2 - y_1y_2 + (x_1y_2 + x_2y_1)i\ , \end{aligned} \tag{1.34}$$

and for $z_2 \neq 0$ we obtain

$$\begin{aligned} \frac{z_1}{z_2} = \frac{x_1 + y_1 i}{x_2 + y_2 i} &= \frac{(x_1 + y_1 i)(x_2 - y_2 i)}{(x_2 + y_2 i)(x_2 - y_2 i)} \\ &= \frac{x_1x_2 + y_1y_2 + (x_2y_1 - x_1y_2)i}{x_2^2 + y_2^2}\ . \end{aligned} \tag{1.35}$$

The real number $|z| =_{\mathrm{df}} |(x,y)| = \left(x^2+y^2\right)^{1/2}$ (cf. Definition 1.17) is called the *absolute value* of $z$.

In order to show our next theorem, it is very helpful to show the equality

$$\overline{w\overline{z}} = \overline{w}z \quad \text{for all } w, z \in \mathbb{C}\ . \tag{1.36}$$

Using (1.34) we directly have

$$w\overline{z} = ux + vy + (vx - uy)i\ .$$

Therefore, we conclude that

$$\begin{aligned} \overline{w\overline{z}} &= ux + vy - (vx - uy)i = ux + vy + (uy - vx)i \\ &= (u - vi)\cdot(x + yi) = \overline{w}z\ , \end{aligned}$$

and the Equality (1.36) is shown.

The following theorem summarizes the important properties of the absolute value of complex numbers:

**Theorem 1.19.** *For all $w, z \in \mathbb{C}$ the following properties are satisfied:*

(1) $|z| \geqslant 0$ *and* $|z| = 0$ *iff* $z = 0$*;*
(2) $|z|^2 = z \cdot \overline{z}$*;*
(3) $|wz| = |w|\,|z|$*;*
(4) $|w + z| \leqslant |w| + |z|$*, and*

(5) $||w| - |z|| \leqslant |w - z|$.

*Proof.* By Corollary 1.1, Property (1) is obvious.

To show Property (2), let $z = x + yi$. Then we obtain

$$\begin{aligned} z \cdot \overline{z} &= (x + yi) \cdot (x - yi) \\ &= x^2 + x(-yi) + yxi + yi(-yi) \\ &= x^2 - xyi + xyi - y^2i^2 \\ &= x^2 + y^2 = |z|^2 \ , \end{aligned}$$

and Property (2) is shown.

We continue with Property (3). Let $w = u + vi$ and let $z = x + iy$. Then by (1.34) and the definition of the absolute value we have

$$\begin{aligned} |wz|^2 &= (ux - vy)^2 + (vx + uy)^2 \\ &= u^2x^2 - 2uxvy + v^2y^2 + v^2x^2 + 2uxvy + u^2y^2 \\ &= u^2x^2 + v^2y^2 + v^2x^2 + u^2y^2 \\ &= (u^2 + v^2) \cdot (x^2 + y^2) = |w|^2 \cdot |z|^2 \ . \end{aligned}$$

Thus, we conclude that $|wz| = |w|\,|z|$, and Property (3) is proved.

It remains to show the triangle inequality. First, we note that $2x = 2\sqrt{x^2}$ for all $x \in \mathbb{R}$. By Corollary 1.1 we also know that $y^2 \geqslant 0$ for all $y \in \mathbb{R}$. Hence, we conclude that $2x \leqslant 2\sqrt{x^2 + y^2}$. Using the latter inequality we directly see that for all $z \in \mathbb{C}$, where $z = x + yi$ the following holds:

$$z + \overline{z} = (x + yi) + (x - yi) = 2x \leqslant 2\sqrt{x^2 + y^2} = 2\,|z| \ . \tag{1.37}$$

Now, we apply Property (2) of Theorem 1.19 and the equality $\overline{w + z} = \overline{w} + \overline{z}$ (cf. Exercise 1.21 below), and obtain

$$\begin{aligned} |w + z|^2 &= (w + z) \cdot (\overline{w + z}) = (w + z) \cdot (\overline{w} + \overline{z}) \\ &= w\overline{w} + z\overline{w} + w\overline{z} + z\overline{z} \\ &= |w|^2 + \overline{w\overline{z}} + w\overline{z} + |z|^2 \quad \text{(by Eq. (1.36))} \\ &\leqslant |w|^2 + 2\,|w\overline{z}| + |z|^2 \quad \text{(by Eq. (1.37))} \\ &= |w|^2 + 2\,|w|\,|\overline{z}| + |z|^2 \quad \text{(by Property (3))} \\ &= |w|^2 + 2\,|w|\,|z| + |z|^2 \quad (\text{since } |\overline{z}| = |z|) \\ &= (|w| + |z|)^2 \ . \end{aligned}$$

Thus, taking the root on both sides yields $|w + z| \leqslant |w| + |z|$, and Property (4) is shown.

Finally, Property (5) is shown as in the real case (cf. Proposition 1.1). ∎

Moreover, the complex numbers can be represented as points of the *complex plane* (see Figure 1.8). It should be noted that $\varphi$ is given in *radians*.

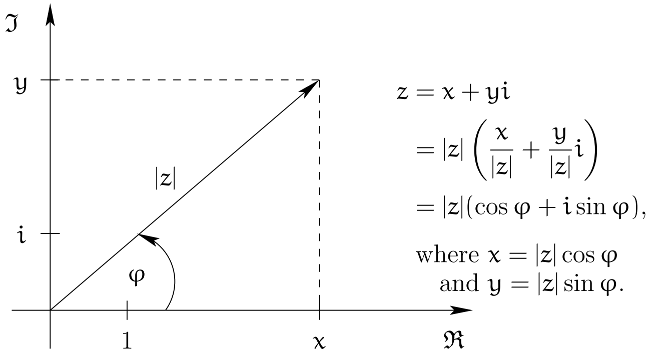
$\mathfrak{I}$
$y$
$i$
$|z|$
$\varphi$
$1$
$x$
$\mathfrak{R}$
$z = x + yi$
$= |z|\left(\frac{x}{|z|} + \frac{y}{|z|}i\right)$
$= |z|(\cos\varphi + i\sin\varphi)$,
where $x = |z|\cos\varphi$
and $y = |z|\sin\varphi$.

Fig. 1.8: The complex plane

We include Figure 1.8 at this time for the sake of illustration, since we have not defined yet what the functions sine and cosine are. This will be done later. Then we shall also see why this representation has several benefits. Note that $\varphi$ and $|z|$ uniquely determine the represented complex number provided $-\pi < \varphi \leqslant \pi$ and $z \neq 0$. Then we use $\arg(z)$ to refer to $\varphi$.

Finally, Definition 1.17 suggests to ask whether or not we can also define an $m$-dimensional complex linear space. The affirmative answer is provided below, but as we shall see some modifications are necessary.

We define $\mathbb{C}^m$ in analogue to $\mathbb{R}^m$. Addition for elements of $\mathbb{C}^m$, i.e., $+\colon \mathbb{C}^m \times \mathbb{C}^m \to \mathbb{C}^m$ is also defined analogously; that is, we set

$$w + z =_{\mathrm{df}} (w_1 + z_1, \ldots, w_m + z_m) \quad \text{for all } w, z \in \mathbb{C}^m \ .$$

The addition of the $i$th components, i.e., $w_i + z_i$, $i = 1, \ldots, m$, is the usual addition in $\mathbb{C}$.

Multiplication $\cdot\colon \mathbb{C} \times \mathbb{C}^m \to \mathbb{C}^m$ is defined canonically as follows:
Let $\alpha \in \mathbb{C}$, and let $z \in \mathbb{C}^m$; then we set

$$\alpha \cdot z =_{\mathrm{df}} (\alpha z_1, \ldots, \alpha z_m) \ .$$

Note that the multiplication of the $i$th components, i.e., $\alpha z_i$, is the usual multiplication in $\mathbb{C}$.

It is easy to see that the properties stated in Proposition 1.4 directly translate to the complex case. But now, it becomes more complicated, since we have to define the complex analogue for a scalar product. Recalling that the scalar product for $\mathbb{R}^m$ has been used to induce the Euclidean norm we see that a new idea is needed. Theorem 1.19, Assertion (2), hints that one should define

$$\langle w, z\rangle =_{\mathrm{df}} \sum_{i=1}^{m} w_i \overline{z}_i \tag{1.38}$$

for all $w = (w_1, \ldots, w_m) \in \mathbb{C}^m$ and $z = (z_1, \ldots, z_m) \in \mathbb{C}^m$. Now, it is easy to see that $\langle\,\cdot\,,\,\cdot\,\rangle \colon \mathbb{C}^m \times \mathbb{C}^m \to \mathbb{C}$. The product $\langle\,\cdot\,,\,\cdot\,\rangle$ is called the *Hermitian form* in honor of Charles Hermite.

The following definition provides the remaining parts:

**Definition 1.19 ($m$-Dimensional Complex Linear Space).**

(1) We call $(\mathbb{C}^m, +, \cdot\,)$ the $m$-*dimensional complex linear space* or $m$-*dimensional complex vector space.*
(2) The mapping $\|\cdot\| : \mathbb{C}^m \to \mathbb{R}$ defined as $\|z\| =_{\mathrm{df}} \langle z, z\rangle^{1/2}$ for all $z \in \mathbb{C}^m$ is said to be the *complex Euclidean norm* on $\mathbb{C}^m$.
(3) The number $\|w - z\|$ is called the *complex Euclidean distance* of $w$ and $z$ in $\mathbb{C}^m$.
(5) We call $(\mathbb{C}^m, \|\cdot\|)$ the $m$-*dimensional complex Euclidean space.*

Further properties of the $m$-dimensional complex Euclidean space are given in the problem set for this chapter. There we shall also point out similarities and differences between the scalar product and the Hermitian form.

**Exercise 1.20.** *Prove the following identities:*

(1) $((1+i)/2))^4 = -1/4$ *;*
(2) $5/(1-2i) = 1+2i$ .

**Exercise 1.21.** *Show the following:*

(1) $i^{4n+1} = i,\;\; i^{4n+2} = -1,\;\; i^{4n+3} = -i,$ *and* $i^{4n+4} = 1$ *for all* $n \in \mathbb{N}_0$*;*
(2) $\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2, \quad \overline{z_1 \cdot z_2} = \overline{z}_1 \cdot \overline{z}_2$ , *and* $z = \overline{\overline{z}}$ *for all* $z, z_1, z_2 \in \mathbb{C}$*;*
(3) $\Re\,(z_1/z_2) = \Re(z_1 \cdot \overline{z}_2)/\,|z_2|^2$ *and* $\Im\,(z_1/z_2) = \Im(z_1 \cdot \overline{z}_2)/\,|z_2|^2$*;*
(4) $z \in \mathbb{R}$ *if and only if* $z = \overline{z}$.

**Exercise 1.22.** *Determine all complex numbers* $z$ *such that the condition* $\Im(2\overline{z} + z) = 1$ *is satisfied.*

## Problems for Chapter 1

**1.1.** Show that for all $a, b \in \mathbb{R}$ the following assertions hold:

(1) If $0 < a$ and $0 < b$ then $a/b + b/a \geqslant 2$;
(2) if $0 < a$ and $0 < b$ such that $ab > 1$ then $a + b \geqslant 2$.

**1.2.** Show that for all $a, b, c \in \mathbb{R}$, $a, b, c > 0$ the inequalities

$$\sqrt{ab} \leqslant \frac{a+b}{2} \quad \text{and}$$
$$\sqrt[3]{abc} \leqslant \frac{a+b+c}{3} \ .$$

are satisfied.

Prove or disprove that the following generalization holds: Let $n \in \mathbb{N}$ and let $a_i \in \mathbb{R}$, $a_i \geqslant 0$, $i = 1, \ldots, n$, then we have

$$\sqrt[n]{\prod_{i=1}^{n} a_i} \leqslant \frac{1}{n} \cdot \sum_{i=1}^{n} a_i \ . \tag{1.39}$$

If the answer is affirmative then determine under what conditions equality holds. Note that the left-hand side of Inequality (1.39) is called the *geometric mean* and the right-hand side is called the *arithmetic mean.*

**1.3.** Let $n \in \mathbb{N}$, and let $a_i \in \mathbb{R}$, $a_i > 0$ for all $i = 1, \ldots, n$. Prove or disprove the following inequality:

$$\frac{n}{\sum_{i=1}^{n}(1/a_i)} \leqslant \sqrt[n]{\prod_{i=1}^{n} a_i} \ . \tag{1.40}$$

Note that the left-hand side of Inequality (1.40) is called the *harmonic mean.*

**1.4.** Show that $\binom{2n}{n} \geqslant 2^n$ for all $n \in \mathbb{N}_0$.

**1.5.** Show that $\prod\limits_{k=1}^{n} (2k-1)/(2k) \leqslant 1/\sqrt{3n+1}$ for all $n \in \mathbb{N}$.

**1.6.** Prove or disprove that

$$\frac{1}{\sqrt{n}} < \sqrt{n+1} - \sqrt{n-1} \quad \text{for all } n \in \mathbb{N} \ .$$

**1.7.** Let $\mathbb{N}^*$ be the set of all finite tuples of natural numbers, i.e., define $\mathbb{N}^* =_{df} \bigcup_{n \in \mathbb{N}} \mathbb{N}^n$. Prove or disprove that $\mathbb{N}^*$ is countable.

**1.8.** Prove or disprove that $\{f \mid f \colon \mathbb{N}_0 \to \{0, 1\}\}$ is countable.

**1.9.** Let $n \in \mathbb{N}$ be arbitrarily fixed, and let $a_k \in \mathbb{R}$ for all $k \in \{1, \ldots, n\}$. Prove or disprove that

$$\left(\sum_{k=1}^{n} \frac{a_k}{k}\right)^2 \leqslant \left(\sum_{k=1}^{n} k^3 a_k^2\right) \left(\sum_{k=1}^{n} k^{-5}\right) \ .$$

**1.10.** Show that for all $x, y, z \in \mathbb{R}^m$ and all $\alpha \in \mathbb{R}$ the following properties are satisfied:

(i) $\langle x, y\rangle = \langle y, x\rangle$ (symmetry);
(ii) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$;
(iii) $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$;
(iv) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle = \langle x, \alpha y\rangle$;
(v) $\langle x, x\rangle \geqslant 0$ and $\langle x, x\rangle = 0$ iff $x = 0$ (positive-definiteness).

Note that Properties (ii) through (iv) establish the bilinearity of the scalar product.

**1.11.** Show that for the Hermitian form the following properties are satisfied for $w, \widetilde{w}, z \in \mathbb{C}^m$ and all $\alpha \in \mathbb{C}$:

(i) $\langle w, z\rangle = \overline{\langle z, w\rangle}$ (Hermitian symmetry);
(ii) $\langle w + \widetilde{w}, z\rangle = \langle w, z\rangle + \langle \widetilde{w}, z\rangle$;
(iii) $\langle w, \widetilde{w} + z\rangle = \langle w, \widetilde{w}\rangle + \langle w, z\rangle$;
(iv) $\langle \alpha w, z\rangle = \alpha\langle w, z\rangle$ and $\langle w, \alpha z\rangle = \overline{\alpha}\langle w, z\rangle$;
(v) $\langle z, z\rangle \geqslant 0$ and $\langle z, z\rangle = 0$ iff $z = 0$ (positive-definiteness).

Note that Properties (ii) through (iv) establish the sesquilinearity of the Hermitian form.

**1.12.** Show the Cauchy–Schwarz inequality for the complex case, i.e., for all $w, z \in \mathbb{C}^m$ we have $|\langle w, z\rangle| \leqslant \|w\|\,\|z\|$. Determine under what conditions equality holds.

**1.13.** Prove that for all $z, z_1, z_2 \in \mathbb{C} \setminus \{0\}$, and $m, n \in \mathbb{Z}$ the following assertions are satisfied:

(i) $z^{m+n} = z^n \cdot z^n$;
(ii) $(z^m)^n = z^{mn}$;
(iii) $(z_1 \cdot z_2)^n = z_1^n \cdot z_2^n$.

**1.14.** Prove or disprove that

$$\frac{1}{\sqrt{2}}\left(|x| + |y|\right) \leqslant |z| \leqslant |x| + |y|$$

for all $z \in \mathbb{C}$, where $z = x + yi$. Determine under what conditions equality holds.

**1.15.** Determine the set of all complex numbers $z$ for which $z \cdot (1 + z^2)^{-1} \in \mathbb{R}$.

**1.16.** Prove the binomial theorem for complex numbers.

**1.17.** Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be defined as $f(x) =_{df} x^2 + 2x$ and $g(x) =_{df} x + 1$ for all $x \in \mathbb{R}$. Prove that $f \circ g \neq g \circ f$.

**1.18.** Let $M \neq \emptyset$ be any set, and let $S(M) =_{df} \{f \mid f\colon M \to M \text{ is bijective}\}$. Furthermore, let $\circ\colon S(M) \to S(M)$ be the composition (cf. Definition 1.14). Prove or disprove that $(S(M), \circ)$ is a group.

**1.19.** Provide a function $f\colon \mathbb{R} \to \mathbb{R}$ such that

(i) the function $f$ is neither injective nor surjective;
(ii) the function $f$ is injective but not surjective;
(iii) the function $f$ is not injective but surjective;
(iv) the function $f$ is injective and surjective.

**1.20.** Let $M, N$, and $K$ be arbitrary sets. Prove or disprove the following distributive laws:

(i) $M \times (N \cup K) = (M \times N) \cup (M \times K)$;
(ii) $M \times (N \cap K) = (M \times N) \cap (M \times K)$.

**1.21.** Let $M$ be any set, and let $f\colon \wp(M) \to \wp(M)$ be any mapping such that for all $A$, $B \subseteq M$ the condition if $A \subseteq B$ then $f(A) \subseteq f(B)$ is satisfied. Prove that there must exist a set $C$ such that $C = f(C)$.

**1.22.** Prove or disprove the following: For every countable set $X$ the set $\wp(X)$ is uncountable.

**1.23.** Let $A \neq \emptyset$ be any set, and let $R, L \subseteq A \times A$ be any binary relations over $A$. We define

$$L \circ R =_{df} \{(a, c) \mid \text{ there is a } b \in A \text{ such that } (a, b) \in L \text{ and } (b, c) \in R\} .$$

We set $R^0 =_{df} \{(a, a) \mid a \in A\}$ and define inductively $R^{n+1} =_{df} R^n \circ R$ for all $n \in \mathbb{N}_0$. Furthermore, we set $\langle R \rangle =_{df} \bigcup_{n \in \mathbb{N}_0} R^n$. Prove or disprove the following:

(i) $\langle R \rangle$ is a binary relation over $A$;
(ii) $\langle R \rangle$ is reflexive;
(iii) $\langle R \rangle$ is transitive;
(iv) $\langle R \rangle = \langle \langle R \rangle \rangle$.

**1.24.** Determine the set of all $x \in \mathbb{R}$ such that the inequality $\sqrt[4]{x} \leqslant 3/8 + 2x$ is satisfied.

**1.25.** Let $M \neq \emptyset$ be any set. Prove or disprove that $(\wp(M), \subseteq)$ is an ordered set.

**1.26.** Consider the set $S =_{df} \{z \mid z \in \mathbb{C},\ |z| = 1\}$. Prove that $S$ is a group with respect to multiplication.

Springer

http://www.springer.com/978-3-319-42753-9

Mathematical Analysis and the Mathematics of Computation
Römisch, W.; Zeugmann, Th.
2016, XXIII, 703 p. 52 illus., 32 illus. in color., Hardcover
ISBN: 978-3-319-42753-9