Analysis I:
Calculus of One Real Variable

Peter Philip*

Lecture Notes

Created for the Class of Winter Semester 2015/2016 at LMU Munich

January 27, 2016

Contents

---

*E-Mail: philip@math.lmu.de

1 Foundations: Mathematical Logic and Set Theory

1.1 Introductory Remarks

The task of *mathematics* is to establish the truth or falsehood of (formalizable) statements using rigorous logic, and to provide methods for the solution of classes of (e.g. applied) problems, ideally including rigorous logical proofs verifying the validity of the methods (proofs that the method under consideration will, indeed, provide a correct solution).

The topic of this class is *calculus*, which is short for *infinitesimal calculus*, usually understood (as it is here) to mean differential and integral calculus of real and complex numbers (more generally, calculus may refer to any method or system of calculation guided by the symbolic manipulation of expressions, we will briefly touch on another example in Sec. 1.2 below). In that sense, calculus is the beginning part of the broader field of (mathematical) *analysis*, the section of mathematics concerned with the notion of a limit (for us, the most important examples will be limits of sequences (Def. 7.1 below) and limits of functions (Def. 8.17 below)).

Before we can properly define our first limit, however, it still needs some preparatory work. In modern mathematics, the objects under investigation are almost always so-called *sets*. So one aims at deriving (i.e. proving) true (and interesting and useful) statements about sets from other statements about sets known or assumed to be true. Such a derivation or proof means applying logical rules that guarantee the truth of the derived (i.e. proved) statement.

However, unfortunately, a proper definition of the notion of set is not easy, and neither is an appropriate treatment of logic and proof theory. Here, we will only be able to briefly touch on the bare necessities from logic and set theory needed to proceed to the core matter of this class. We begin with logic in Sec. 1.2, followed by set theory in Sec. 1.3, combining both in Sec. 1.4. The interested student can find an introductory presentation of axiomatic set theory in Appendix A and he/she should consider taking a separate class on set theory, logic, and proof theory at a later time.

1.2 Propositional Calculus

1.2.1 Statements

Mathematical logic is a large field in its own right and, as indicated above, a thorough introduction is beyond the scope of this class – the interested reader may refer to [EFT07], [Kun12], and references therein. Here, we will just introduce some basic concepts using common English (rather than formal symbolic languages – a concept touched on in Sec. A.2 of the Appendix and more thoroughly explained in books like [EFT07]).

As mentioned before, mathematics establishes the truth or falsehood of statements. By a *statement* or *proposition* we mean any sentence (any sequence of symbols) that can

reasonably be assigned a *truth value*, i.e. a value of either *true*, abbreviated T, or *false*, abbreviated F. The following example illustrates the difference between statements and sentences that are not statements:

Example 1.1. (a) Sentences that are statements:

> Every dog is an animal. (T)
> Every animal is a dog. (F)
> The number 4 is odd. (F)
> $2 + 3 = 5$. (T)
> $\sqrt{2} < 0$. (F)
> $x + 1 > 0$ holds for each natural number x. (T)

(b) Sentences that are *not* statements:

> Let's study calculus!
> Who are you?
> $3 \cdot 5 + 7$.
> $x + 1 > 0$.
> All natural numbers are green.

The fourth sentence in Ex. 1.1(b) is not a statement, as it can not be said to be either true or false without any further knowledge on x. The fifth sentence in Ex. 1.1(b) is not a statement as it lacks any meaning and can, hence, not be either true or false. It would become a statement if given a definition of what it means for a natural number to be green.

1.2.2   Logical Operators

The next step now is to *combine* statements into new statements using *logical operators*, where the truth value of the combined statements depends on the truth values of the original statements and on the type of logical operator facilitating the combination.

The simplest logical operator is *negation*, denoted $\neg$. It is actually a so-called *unary* operator, i.e. it does not combine statements, but is merely applied to one statement. For example, if A stands for the statement "Every dog is an animal.", then $\neg A$ stands for the statement "Not every dog is an animal."; and if B stands for the statement "The number 4 is odd.", then $\neg B$ stands for the statement "The number 4 is not odd.", which can also be expressed as "The number 4 is even."

To completely understand the action of a logical operator, one usually writes what is known as a *truth table*. For negation, the truth table is

$$
\begin{array}{c|c}
A & \neg A \\
\hline
\text{T} & \text{F} \\
\text{F} & \text{T}
\end{array}
\tag{1.1}
$$

that means if the input statement A is true, then the output statement $\neg A$ is false; if the input statement A is false, then the output statement $\neg A$ is true.

We now proceed to discuss *binary* logical operators, i.e. logical operators combining precisely two statements. The following four operators are essential for mathematical reasoning:

Conjunction: *A and B*, usually denoted $A \wedge B$.

Disjunction: *A or B*, usually denoted $A \vee B$.

Implication: *A implies B*, usually denoted $A \Rightarrow B$.

Equivalence: *A is equivalent to B*, usually denoted $A \Leftrightarrow B$.

Here is the corresponding truth table:

$$
\begin{array}{c|c||c|c|c|c}
A & B & A \wedge B & A \vee B & A \Rightarrow B & A \Leftrightarrow B \\
\hline
T & T & T & T & T & T \\
T & F & F & T & F & F \\
F & T & F & T & T & F \\
F & F & F & F & T & T
\end{array}
\tag{1.2}
$$

When first seen, some of the assignments of truth values in (1.2) might not be completely intuitive, due to the fact that logical operators are often used somewhat differently in common English. Let us consider each of the four logical operators of (1.2) in sequence:

For the use in subsequent examples, let $A_1, \ldots, A_6$ denote the six statements from Ex. 1.1(a).

Conjunction: Most likely the easiest of the four, basically identical to common language use: $A \wedge B$ is true if, and only if, both A and B are true. For example, using Ex. 1.1(a), $A_1 \wedge A_4$ is the statement "Every dog is an animal and $2 + 3 = 5$.", which is true since both A_1 and A_4 are true. On the other hand, $A_1 \wedge A_3$ is the statement "Every dog is an animal and the number 4 is odd.", which is false, since A_3 is false.

Disjunction: The disjunction $A \vee B$ is true if, and only if, at least one of the statements A, B is true. Here one already has to be a bit careful – $A \vee B$ defines the *inclusive* or, whereas "or" in common English is often understood to mean the exclusive or (which is false if both input statements are true). For example, using Ex. 1.1(a), $A_1 \vee A_4$ is the statement "Every dog is an animal or $2 + 3 = 5$.", which is true since both A_1 and A_4 are true. The statement $A_1 \vee A_3$, i.e. "Every dog is an animal or the number 4 is odd." is also true, since A_1 is true. However, the statement $A_2 \vee A_5$, i.e. "Every animal is a dog or $\sqrt{2} < 0$." is false, as both A_2 and A_5 are false.

As you will have noted in the above examples, logical operators can be applied to combine statements that have no obvious contents relation. While this might seem strange, introducing contents-related restrictions is unnecessary as well as undesirable, since it is often not clear which seemingly unrelated statements might suddenly appear in a common context in the future. The same occurs when considering implications and equivalences, where it might seem even more obscure at first.

Implication: Instead of A implies B, one also says *if A then B, B is a consequence of A, B is concluded or inferred from A, A is sufficient for B*, or *B is necessary for A*. The implication $A \Rightarrow B$ is always true, except if A is true and B is false. At first glance, it might be surprising that $A \Rightarrow B$ is defined to be true for A false and B true, however, there are many examples of incorrect statements implying correct statements. For instance, squaring the (false) equality of integers $-1 = 1$, implies the (true) equality of integers $1 = 1$. However, as with conjunction and disjunction, it is perfectly valid to combine statements without any obvious context relation: For example, using Ex. 1.1(a), the statement $A_1 \Rightarrow A_6$, i.e. "Every dog is an animal implies $x + 1 > 0$ holds for each natural number x." is true, since A_6 is true, whereas the statement $A_4 \Rightarrow A_2$, i.e. "$2 + 3 = 5$ implies every animal is a dog." is false, as A_4 is true and A_2 is false.

Of course, the implication $A \Rightarrow B$ is not really useful in situations, where the truth values of both A and B are already known. Rather, in a typical application, one tries to establish the truth of A to prove the truth of B (a strategy that will fail if A happens to be false).

Example 1.2. Suppose we know Sasha to be a member of a group of children. Then the statement A "Sasha is a girl." implies the statement B "There is at least one girl in the group." A priori, we might not know if Sasha is a girl or a boy, but if we can establish Sasha to be a girl, then we also know B to be true. If we find Sasha to be a boy, then we do not know, whether B is true or false.

—

Equivalence: $A \Leftrightarrow B$ means A is true if, and only if, B is true. Once again, using input statements from Ex. 1.1(a), we see that $A_1 \Leftrightarrow A_4$, i.e. "Every dog is an animal is equivalent to $2 + 3 = 5$.", is true as well as $A_2 \Leftrightarrow A_3$, i.e. "Every animal is a dog is equivalent to the number 4 is odd.". On the other hand, $A_4 \Leftrightarrow A_5$, i.e. "$2 + 3 = 5$ is equivalent to $\sqrt{2} < 0$, is false.

Analogous to the situation of implications, $A \Leftrightarrow B$ is not really useful if the truth values of both A and B are known a priori, but can be a powerful tool to prove B to be true or false by establishing the truth value of A. It is obviously more powerful than the implication as illustrated by the following example (compare with Ex. 1.2):

Example 1.3. Suppose we know Sasha is the tallest member of a group of children. Then the statement A "Sasha is a girl." is equivalent to the statement B "The tallest kid in the group is a girl." As in Ex. 1.2, if we can establish Sasha to be a girl, then we also know B to be true. However, in contrast to Ex. 1.2, if we find Sasha to be a boy, we know B to be false.

Remark 1.4. In computer science, the truth value T is often coded as 1 and the truth value F is often coded as 0.

1.2.3 Rules

Note that the expressions in the first row of the truth table (1.2) (e.g. $A \wedge B$) are *not* statements in the sense of Sec. 1.2.1, as they contain the *statement variables* (also known

as *propositional variables*) A or B. However, the expressions become statements if all statement variables are substituted with actual statements. We will call expressions of this form *propositional formulas*. Moreover, if a truth value is assigned to each statement variable of a propositional formula, then this uniquely determines the truth value of the formula. In other words, the truth value of the propositional formula can be *calculated* from the respective truth values of its statement variables – a first justification for the name *propositional calculus*.

Example 1.5. (a) Consider the propositional formula $(A \wedge B) \vee (\neg B)$. Suppose A is true and B is false. The truth value of the formula is obtained according to the following truth table:

$$
\begin{array}{c|c|c|c|c}
A & B & A \wedge B & \neg B & (A \wedge B) \vee (\neg B) \\
\hline
T & F & F & T & T
\end{array}
\tag{1.3}
$$

(b) The propositional formula $A \vee (\neg A)$, also known as the *law of the excluded middle*, has the remarkable property that its truth value is T for every possible choice of truth values for A:

$$
\begin{array}{c|c|c}
A & \neg A & A \vee (\neg A) \\
\hline
T & F & T \\
\hline
F & T & T
\end{array}
\tag{1.4}
$$

Formulas with this property are of particular importance.

Definition 1.6. A propositional formula is called a *tautology* or *universally true* if, and only if, its truth value is T for all possible assignments of truth values to all the statement variables it contains.

Notation 1.7. We write $\phi(A_1, \ldots, A_n)$ if, and only if, the propositional formula ϕ contains precisely the n statement variables $A_1, \ldots, A_n$.

Definition 1.8. The propositional formulas $\phi(A_1, \ldots, A_n)$ and $\psi(A_1, \ldots, A_n)$ are called *equivalent* if, and only if, $\phi(A_1, \ldots, A_n) \Leftrightarrow \psi(A_1, \ldots, A_n)$ is a tautology.

Lemma 1.9. *The propositional formulas $\phi(A_1, \ldots, A_n)$ and $\psi(A_1, \ldots, A_n)$ are equivalent if, and only if, they have the same truth value for all possible assignments of truth values to $A_1, \ldots, A_n$.*

Proof. If $\phi(A_1, \ldots, A_n)$ and $\psi(A_1, \ldots, A_n)$ are equivalent and A_i is assigned the truth value t_i, $i = 1, \ldots, n$, then $\phi(A_1, \ldots, A_n) \Leftrightarrow \psi(A_1, \ldots, A_n)$ being a tautology implies it has truth value T. From (1.2) we see that either $\phi(A_1, \ldots, A_n)$ and $\psi(A_1, \ldots, A_n)$ both have truth value T or they both have truth value F.

If, on the other hand, we know $\phi(A_1, \ldots, A_n)$ and $\psi(A_1, \ldots, A_n)$ have the same truth value for all possible assignments of truth values to $A_1, \ldots, A_n$, then, given such an assignment, either $\phi(A_1, \ldots, A_n)$ and $\psi(A_1, \ldots, A_n)$ both have truth value T or both have truth value F, i.e. $\phi(A_1, \ldots, A_n) \Leftrightarrow \psi(A_1, \ldots, A_n)$ has truth value T in each case, showing it is a tautology. ∎

For all logical purposes, two equivalent formulas are exactly the same – it does not matter if one uses one or the other. The following theorem provides some important equivalences of propositional formulas. As too many parentheses tend to make formulas less readable, we first introduce some precedence conventions for logical operators:

Convention 1.10. $\neg$ takes precedence over $\wedge, \vee$, which take precedence over $\Rightarrow, \Leftrightarrow$. So, for example,

$$(A \vee \neg B \Rightarrow \neg B \wedge \neg A) \Leftrightarrow \neg C \wedge (A \vee \neg D)$$

is the same as

$$\Big((A \vee (\neg B)) \Rightarrow ((\neg B) \wedge (\neg A))\Big) \Leftrightarrow \Big((\neg C) \wedge (A \vee (\neg D))\Big).$$

Theorem 1.11. (a) $(A \Rightarrow B) \Leftrightarrow \neg A \vee B$. *This means one can actually define implication via negation and disjunction.*

(b) $(A \Leftrightarrow B) \Leftrightarrow \big((A \Rightarrow B) \wedge (B \Rightarrow A)\big)$, *i.e. A and B are equivalent if, and only if, A is both necessary and sufficient for B. One also calls the implication $B \Rightarrow A$ the* converse *of the implication $A \Rightarrow B$. Thus, A and B are equivalent if, and only if, both $A \Rightarrow B$ and its converse hold true.*

(c) *Commutativity of Conjunction:* $A \wedge B \Leftrightarrow B \wedge A$.

(d) *Commutativity of Disjunction:* $A \vee B \Leftrightarrow B \vee A$.

(e) *Associativity of Conjunction:* $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$.

(f) *Associativity of Disjunction:* $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$.

(g) *Distributivity I:* $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$.

(h) *Distributivity II:* $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$.

(i) *De Morgan's Law I:* $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$.

(j) *De Morgan's Law II:* $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$.

(k) *Double Negative:* $\neg\neg A \Leftrightarrow A$.

(l) *Contraposition:* $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$.

Proof. Each equivalence is proved by providing a truth table and using Lem. 1.9.

(a):

A	B	$\neg A$	$A \Rightarrow B$	$\neg A \vee B$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

(b) – (h): Exercise.

(i):

A	B	$\neg A$	$\neg B$	$A \wedge B$	$\neg(A \wedge B)$	$\neg A \vee \neg B$
T	T	F	F	T	F	F
T	F	F	T	F	T	T
F	T	T	F	F	T	T
F	F	T	T	F	T	T

(j): Exercise.

(k):

A	$\neg A$	$\neg\neg A$
T	F	T
F	T	F

(l):

A	B	$\neg A$	$\neg B$	$A \Rightarrow B$	$\neg B \Rightarrow \neg A$
T	T	F	F	T	T
T	F	F	T	F	F
F	T	T	F	T	T
F	F	T	T	T	T

Having checked all the rules completes the proof of the theorem. ∎

The importance of the rules provided by Th. 1.11 lies in their providing *proof techniques*, i.e. methods for establishing the truth of statements from statements known or assumed to be true. The rules of Th. 1.11 will be used frequently in proofs throughout this class.

Remark 1.12. Another important proof technique is the so-called *proof by contradiction*, also called *indirect proof*. It is based on the observation, called the *principle of contradiction*, that $A \wedge \neg A$ is always false:

$$
\begin{array}{c|c||c}
A & \neg A & A \wedge \neg A \\
\hline
T & F & F \\
F & T & F
\end{array}
\tag{1.5}
$$

Thus, one possibility of proving a statement B to be true is to show $\neg B \Rightarrow A \wedge \neg A$ for some arbitrary statement A. Since the right-hand side of the implication is false, the left-hand side must also be false, proving B is true.

—

Two more rules we will use regularly in subsequent proofs are the so-called transitivity of implication and the transitivity of equivalence (we will encounter equivalence again in the context of relations in Sec. 1.3 below). In preparation for the transitivity rules, we generalize implication to propositional formulas:

Definition 1.13. In generalization of the implication operator defined in (1.2), we say the propositional formula $\phi(A_1, \ldots, A_n)$ *implies* the propositional formula $\psi(A_1, \ldots, A_n)$ (denoted $\phi(A_1, \ldots, A_n) \Rightarrow \psi(A_1, \ldots, A_n)$) if, and only if, each assignment of truth values to the $A_1, \ldots, A_n$ that makes $\phi(A_1, \ldots, A_n)$ true, makes $\psi(A_1, \ldots, A_n)$ true as well.

Theorem 1.14. **(a)** *Transitivity of Implication:* $(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)$.

(b) *Transitivity of Equivalence:* $(A \Leftrightarrow B) \wedge (B \Leftrightarrow C) \Rightarrow (A \Leftrightarrow C)$.

Proof. According to Def. 1.13, the rules can be verified by providing truth tables that show that, for all possible assignments of truth values to the propositional formulas on the left-hand side of the implications, either the left-hand side is false or both sides are true. (a):

A	B	C	$A \Rightarrow B$	$B \Rightarrow C$	$(A \Rightarrow B) \wedge (B \Rightarrow C)$	$A \Rightarrow C$
T	T	T	T	T	T	T
T	F	T	F	T	F	T
F	T	T	T	T	T	T
F	F	T	T	T	T	T
T	T	F	T	F	F	F
T	F	F	F	T	F	F
F	T	F	T	F	F	T
F	F	F	T	T	T	T

(b):

A	B	C	$A \Leftrightarrow B$	$B \Leftrightarrow C$	$(A \Leftrightarrow B) \wedge (B \Leftrightarrow C)$	$A \Leftrightarrow C$
T	T	T	T	T	T	T
T	F	T	F	F	F	T
F	T	T	F	T	F	F
F	F	T	T	F	F	F
T	T	F	T	F	F	F
T	F	F	F	T	F	F
F	T	F	F	F	F	T
F	F	F	T	T	T	T

Having checked both rules, the proof is complete. ∎

Definition and Remark 1.15. A *proof* of the statement B is a finite sequence of statements $A_1, A_2, \ldots, A_n$ such that A_1 is true; for $1 \leq i < n$, A_i implies A_{i+1}, and A_n implies B. If there exists a proof for B, then Th. 1.14(a) guarantees that B is true.

Remark 1.16. *Principle of Duality:* In Th. 1.11, there are several pairs of rules that have an analogous form: (c) and (d), (e) and (f), (g) and (h), (i) and (j). These analogies are due to the general law called the principle of duality: If $\phi(A_1, \ldots, A_n) \Rightarrow \psi(A_1, \ldots, A_n)$ and only the operators $\wedge, \vee, \neg$ occur in ϕ and ψ, then the reverse implication $\Phi(A_1, \ldots, A_n) \Leftarrow \Psi(A_1, \ldots, A_n)$ holds, where one obtains Φ from ϕ and Ψ from ψ by replacing each $\wedge$ with $\vee$ and each $\vee$ with $\wedge$. In particular, if, instead of an implication, we start with an equivalence (as in the examples from Th. 1.11), then we obtain another equivalence.

1.3 Set Theory

In the previous section, we have had a first glance at statements and corresponding truth values. In the present section, we will move our focus to the objects such statements are about. Reviewing Example 1.1(a), and recalling that this is a mathematics class rather than one in zoology, the first two statements of Example 1.1(a) are less relevant for us than statements 3–6. As in these examples, we will nearly always be interested in statements involving numbers or collections of numbers or collections of such collections etc.

In modern mathematics, the term one usually uses instead of "collection" is "set". In 1895, Georg Cantor defined a set as "any collection into a whole M of definite and separate objects m of our intuition or our thought". The objects m are called the *elements* of the set M. As explained in Appendix A, without restrictions and refinements, Cantor's set theory is not free of contradictions and, thus, not viable to be used in the foundation of mathematics. Axiomatic set theory provides these necessary restrictions and refinements and an introductory treatment can also be found in Appendix A. However, it is possible to follow and understand the rest of this class, without having studied Appendix A.

Notation 1.17. We write $m \in M$ for the statement "m is an element of the set M".

Definition 1.18. The sets M and N are equal, denoted $M = N$, if, and only if, M and N have precisely the same elements.

—

Definition 1.18 means we know everything about a set M if, and only if, we know all its elements.

Definition 1.19. The set with no elements is called the *empty set*; it is denoted by the symbol $\emptyset$.

Example 1.20. For finite sets, we can simply write down all its elements, for example, $A := \{0\}$, $B := \{0, 17.5\}$, $C := \{5, 1, 5, 3\}$, $D := \{3, 5, 1\}$, $E := \{2, \sqrt{2}, -2\}$, where the symbolism ":=" is to be read as "is defined to be equal to".

Note $C = D$, since both sets contain precisely the same elements. In particular, the order in which the elements are written down plays no role and a set does not change if an element is written down more than once.

If a set has many elements, instead of writing down all its elements, one might use abbreviations such as $F := \{-4, -2, \ldots, 20, 22, 24\}$, where one has to make sure the meaning of the dots is clear from the context.

Definition 1.21. The set A is called a *subset* of the set B (denoted $A \subseteq B$ and also referred to as the *inclusion* of A in B) if, and only if, every element of A is also an element of B (one sometimes also calls B a *superset* of A and writes $B \supseteq A$). Please note that $A = B$ is allowed in the above definition of a subset. If $A \subseteq B$ and $A \neq B$, then A is called a *strict* subset of B, denoted $A \subsetneq B$.

If B is a set and $P(x)$ is a statement about an element x of B (i.e., for each $x \in B$, $P(x)$ is either true or false), then we can define a subset A of B by writing

$$A := \{x \in B : P(x)\}. \tag{1.7}$$

This notation is supposed to mean that the set A consists precisely of those elements of B such that $P(x)$ is true (has the truth value T in the language of Sec. 1.2).

Example 1.22. (a) For each set A, one has $A \subseteq A$ and $\emptyset \subseteq A$.

(b) If $A \subseteq B$, then $A = \{x \in B : x \in A\}$.

(c) We have $\{3\} \subseteq \{6.7, 3, 0\}$. Letting $A := \{-10, -8, \ldots, 8, 10\}$, we have $\{-2, 0, 2\} = \{x \in A : x^3 \in A\}$, $\emptyset = \{x \in A : x + 21 \in A\}$.

Remark 1.23. As a consequence of Def. 1.18, the sets A and B are equal if, and only if, one has both inclusions, namely $A \subseteq B$ and $B \subseteq A$. Thus, when proving the equality of sets, one often divides the proof into two parts, first proving one inclusion, then the other.

Definition 1.24. (a) The *intersection* of the sets A and B, denoted $A \cap B$, consists of all elements that are in A and in B. The sets A, B are said to be *disjoint* if, and only if, $A \cap B = \emptyset$.

(b) The *union* of the sets A and B, denoted $A \cup B$, consists of all elements that are in A or in B (as in the logical disjunction in (1.2), the or is meant nonexclusively). If A and B are disjoint, one sometimes writes $A \dot\cup B$ and speaks of the *disjoint union* of A and B.

(c) The *difference* of the sets A and B, denoted $A \setminus B$ (read "A minus B" or "A without B"), consists of all elements of A that are not elements of B, i.e. $A \setminus B := \{x \in A : x \notin B\}$. If B is a subset of a given set A (sometimes called the *universe* in this context), then $A \setminus B$ is also called the *complement* of B with respect to A. In that case, one also writes $B^c := A \setminus B$ (note that this notation suppresses the dependence on A).

Example 1.25. (a) Examples of Intersections:

$$\{1, 2, 3\} \cap \{3, 4, 5\} = \{3\}, \tag{1.8a}$$

$$\{\sqrt{2}\} \cap \{1, 2, \ldots, 10\} = \emptyset, \tag{1.8b}$$

$$\{-1, 2, -3, 4, 5\} \cap \{-10, -9, \ldots, -1\} \cap \{-1, 7, -3\} = \{-1, -3\}. \tag{1.8c}$$

(b) Examples of Unions:

$$\{1, 2, 3\} \cup \{3, 4, 5\} = \{1, 2, 3, 4, 5\}, \tag{1.9a}$$

$$\{1, 2, 3\} \dot\cup \{4, 5\} = \{1, 2, 3, 4, 5\}, \tag{1.9b}$$

$$\{-1, 2, -3, 4, 5\} \cup \{-99, -98, \ldots, -1\} \cup \{-1, 7, -3\}$$
$$= \{-99, -98, \ldots, -2, -1, 2, 4, 5, 7\}. \tag{1.9c}$$

(c) Examples of Differences:

$$\{1, 2, 3\} \setminus \{3, 4, 5\} = \{1, 2\}, \tag{1.10a}$$

$$\{1, 2, 3\} \setminus \{3, 2, 1, \sqrt{5}\} = \emptyset, \tag{1.10b}$$

$$\{-10, -9, \ldots, 9, 10\} \setminus \{0\} = \{-10, -9, \ldots, -1\} \cup \{1, 2, \ldots, 9, 10\}. \tag{1.10c}$$

With respect to the universe $\{1, 2, 3, 4, 5\}$, it is

$$\{1, 2, 3\}^{\mathrm{c}} = \{4, 5\}; \tag{1.10d}$$

with respect to the universe $\{0, 1, \ldots, 20\}$, it is

$$\{1, 2, 3\}^{\mathrm{c}} = \{0\} \cup \{4, 5, \ldots, 20\}. \tag{1.10e}$$

As mentioned earlier, it will often be unavoidable to consider sets of sets. Here are first examples: $\{\emptyset, \{0\}, \{0, 1\}\}$, $\{\{0, 1\}, \{1, 2\}\}$.

Definition 1.26. Given a set A, the set of all subsets of A is called the *power set* of A, denoted $\mathcal{P}(A)$ (for reasons explained later (cf. Prop. 2.18), the power set is sometimes also denoted as 2^A).

Example 1.27. Examples of Power Sets:

$$\mathcal{P}(\emptyset) = \{\emptyset\}, \tag{1.11a}$$

$$\mathcal{P}(\{0\}) = \{\emptyset, \{0\}\}, \tag{1.11b}$$

$$\mathcal{P}(\mathcal{P}(\{0\})) = \mathcal{P}(\{\emptyset, \{0\}\}) = \{\emptyset, \{\emptyset\}, \{\{0\}\}, \mathcal{P}(\{0\})\}. \tag{1.11c}$$

—

So far, we have restricted our set-theoretic examples to finite sets. However, not surprisingly, many sets of interest to us will be infinite (we will have to postpone a mathematically precise definition of finite and infinite to Sec. 3.2). We will now introduce the most simple infinite set.

Definition 1.28. The set $\mathbb{N} := \{1, 2, 3, \ldots\}$ is called the set of *natural numbers* (for a more rigorous construction of $\mathbb{N}$, based on the axioms of axiomatic set theory, see Sec. A.3.4 of the Appendix, where Th. A.46 shows $\mathbb{N}$ to be, indeed, infinite). Moreover, we define $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$.

—

The following theorem compiles important set-theoretic rules:

Theorem 1.29. *Let A, B, C, U be sets.*

(a) *Commutativity of Intersections: $A \cap B = B \cap A$.*

(b) *Commutativity of Unions: $A \cup B = B \cup A$.*

(c) *Associativity of Intersections:* $(A \cap B) \cap C = A \cap (B \cap C)$.

(d) *Associativity of Unions:* $(A \cup B) \cup C = A \cup (B \cup C)$.

(e) *Distributivity I:* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

(f) *Distributivity II:* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

(g) *De Morgan's Law I:* $U \setminus (A \cap B) = (U \setminus A) \cup (U \setminus B)$.

(h) *De Morgan's Law II:* $U \setminus (A \cup B) = (U \setminus A) \cap (U \setminus B)$.

(i) *Double Complement: If $A \subseteq U$, then $U \setminus (U \setminus A) = A$.*

Proof. In each case, the proof results from the corresponding rule of Th. 1.11:

(a):
$$x \in A \cap B \Leftrightarrow x \in A \wedge x \in B \overset{\text{Th. 1.11(c)}}{\Leftrightarrow} x \in B \wedge x \in A \Leftrightarrow x \in B \cap A.$$

(g): Under the general assumption of $x \in U$, we have the following equivalences:

$$x \in U \setminus (A \cap B) \Leftrightarrow \neg(x \in A \cap B) \Leftrightarrow \neg\big(x \in A \wedge x \in B\big) \overset{\text{Th. 1.11(i)}}{\Leftrightarrow} \neg(x \in A) \vee \neg(x \in B)$$
$$\Leftrightarrow x \in U \setminus A \vee x \in U \setminus B \Leftrightarrow x \in (U \setminus A) \cup (U \setminus B).$$

The proofs of the remaining rules are left as an exercise.   ∎

Remark 1.30. The correspondence between Th. 1.11 and Th. 1.29 is no coincidence. One can actually prove that, starting with an equivalence of propositional formulas $\phi(A_1, \ldots, A_n) \Leftrightarrow \psi(A_1, \ldots, A_n)$, where both formulas contain only the operators $\wedge, \vee, \neg$, one obtains a set-theoretic rule (stating an equality of sets) by reinterpreting all statement variables $A_1, \ldots, A_n$ as variables for sets, all subsets of a universe U, and replacing $\wedge$ by $\cap$, $\vee$ by $\cup$, and $\neg$ by $U\setminus$ (if there are no multiple negations, then we do not need the hypothesis that $A_1, \ldots, A_n$ are subsets of U). The procedure also works in the opposite direction – one can start with a set-theoretic formula for an equality of sets and translate it into two equivalent propositional formulas.

1.4 Predicate Calculus

Now that we have introduced sets in the previous section, we have to return to the subject of mathematical logic once more. As it turns out, propositional calculus, which we discussed in Sec. 1.2, does not quite suffice to develop the theory of calculus (nor most other mathematical theories). The reason is that we need to consider statements

such as

$$x + 1 > 0 \text{ holds for each natural number } x. \text{ (T)} \tag{1.12a}$$

$$\text{All real numbers are positive. (F)} \tag{1.12b}$$

$$\text{There exists a natural number bigger than 10. (T)} \tag{1.12c}$$

$$\text{There exists a real number } x \text{ such that } x^2 = -1. \text{ (F)} \tag{1.12d}$$

$$\text{For all natural numbers } n, \text{ there exists a natural number bigger than } n. \text{ (T)} \tag{1.12e}$$

That means we are interested in statements involving *universal quantification* via the quantifier "for all" (one also often uses "for each" or "for every" instead), *existential quantification* via the quantifier "there exists", or both. The quantifier of universal quantification is denoted by $\forall$ and the quantifier of existential quantification is denoted by $\exists$. Using these symbols as well as $\mathbb{N}$ and $\mathbb{R}$ to denote the sets of natural and real numbers, respectively, we can restate (1.12) as

$$\forall_{x \in \mathbb{N}} \; x + 1 > 0. \text{ (T)} \tag{1.13a}$$

$$\forall_{x \in \mathbb{R}} \; x > 0. \text{ (F)} \tag{1.13b}$$

$$\exists_{n \in \mathbb{N}} \; n > 10. \text{ (T)} \tag{1.13c}$$

$$\exists_{x \in \mathbb{R}} \; x^2 = -1. \text{ (F)} \tag{1.13d}$$

$$\forall_{n \in \mathbb{N}} \; \exists_{m \in \mathbb{N}} \; m > n. \text{ (T)} \tag{1.13e}$$

Definition 1.31. A *universal statement* has the form

$$\forall_{x \in A} \; P(x), \tag{1.14a}$$

whereas an *existential statement* has the form

$$\exists_{x \in A} \; P(x). \tag{1.14b}$$

In (1.14), A denotes a set and $P(x)$ is a sentence involving the variable x, a so-called *predicate* of x, that becomes a statement (i.e. becomes either true or false) if x is substituted with any concrete element of the set A (in particular, $P(x)$ is allowed to contain further quantifiers, but it must not contain any other quantifier involving x – one says x must be a *free* variable in $P(x)$, not bound by any quantifier in $P(x)$).

The universal statement (1.14a) has the truth value T if, and only if, $P(x)$ has the truth value T for *all* elements $x \in A$; the existential statement (1.14b) has the truth value T if, and only if, $P(x)$ has the truth value T for *at least one* element $x \in A$.

Remark 1.32. Some people prefer to write $\bigwedge_{x \in A}$ instead of $\forall_{x \in A}$ and $\bigvee_{x \in A}$ instead of $\exists_{x \in A}$. Even though this notation has the advantage of emphasizing that the universal statement can be interpreted as a big logical conjunction and the existential statement can be interpreted as a big logical disjunction, it is significantly less common. So we will stick to $\forall$ and $\exists$ in this class.

Remark 1.33. According to Def. 1.31, the existential statement (1.14b) is true if, and only if, $P(x)$ is true for at least one $x \in A$. So if there is precisely one such x, then (1.14b) is true; and if there are several different $x \in A$ such that $P(x)$ is true, then (1.14b) is still true. Uniqueness statements are often of particular importance, and one sometimes writes

$$\underset{x \in A}{\exists!} \; P(x) \tag{1.15}$$

for the statement "there exists a unique $x \in A$ such that $P(x)$ is true". This notation can be defined as an abbreviation for

$$\underset{x \in A}{\exists} \left(P(x) \wedge \underset{y \in A}{\forall} \big(P(y) \Rightarrow x = y \big) \right). \tag{1.16}$$

Example 1.34. Here are some examples of uniqueness statements:

$$\underset{n \in \mathbb{N}}{\exists!} \; n > 10. \; (\text{F}) \tag{1.17a}$$

$$\underset{n \in \mathbb{N}}{\exists!} \; 12 > n > 10. \; (\text{T}) \tag{1.17b}$$

$$\underset{n \in \mathbb{N}}{\exists!} \; 11 > n > 10. \; (\text{F}) \tag{1.17c}$$

$$\underset{x \in \mathbb{R}}{\exists!} \; x^2 = -1. \; (\text{F}) \tag{1.17d}$$

$$\underset{x \in \mathbb{R}}{\exists!} \; x^2 = 1. \; (\text{F}) \tag{1.17e}$$

$$\underset{x \in \mathbb{R}}{\exists!} \; x^2 = 0. \; (\text{T}) \tag{1.17f}$$

Remark 1.35. As for propositional calculus, we also have some important rules for predicate calculus:

(a) Consider the negation of a universal statement, $\neg \underset{x \in A}{\forall} \; P(x)$, which is true if, and only if, $P(x)$ does *not* hold for each $x \in A$, i.e. if, and only if, there exists at least one $x \in A$ such that $P(x)$ is false (such that $\neg P(x)$ is true). We have just proved the rule

$$\neg \underset{x \in A}{\forall} \; P(x) \; \Leftrightarrow \; \underset{x \in A}{\exists} \; \neg P(x). \tag{1.18a}$$

Similarly, consider the negation of an existential statement. We claim the corresponding rule is

$$\neg \underset{x \in A}{\exists} \; P(x) \; \Leftrightarrow \; \underset{x \in A}{\forall} \; \neg P(x). \tag{1.18b}$$

Indeed, we can prove (1.18b) from (1.18a):

$$\neg \underset{x \in A}{\exists} \; P(x) \overset{\text{Th. 1.11(k)}}{\Leftrightarrow} \neg \underset{x \in A}{\exists} \; \neg\neg P(x) \overset{(1.18a)}{\Leftrightarrow} \neg\neg \underset{x \in A}{\forall} \; \neg P(x) \overset{\text{Th. 1.11(k)}}{\Leftrightarrow} \underset{x \in A}{\forall} \; \neg P(x). \tag{1.19}$$

One can interpret (1.18) as a generalization of the De Morgan's laws Th. 1.11(i),(j).

One can actually generalize (1.18) even a bit more: If a statement starts with several quantifiers, then one negates the statement by replacing each $\forall$ with $\exists$ and vice versa plus negating the predicate after the quantifiers (see the example in (1.22e) below).

(b) If A, B are sets and $P(x, y)$ denotes a predicate of both x and y, then $\underset{x \in A}{\forall} \underset{y \in B}{\forall} P(x, y)$ and $\underset{y \in B}{\forall} \underset{x \in A}{\forall} P(x, y)$ both hold true if, and only if, $P(x, y)$ holds true for each $x \in A$ and each $y \in B$, i.e. the order of two consecutive universal quantifiers does not matter:

$$\underset{x \in A}{\forall} \underset{y \in B}{\forall} P(x, y) \Leftrightarrow \underset{y \in B}{\forall} \underset{x \in A}{\forall} P(x, y) \tag{1.20a}$$

In the same way, we obtain the following rule:

$$\underset{x \in A}{\exists} \underset{y \in B}{\exists} P(x, y) \Leftrightarrow \underset{y \in B}{\exists} \underset{x \in A}{\exists} P(x, y). \tag{1.20b}$$

If $A = B$, one also uses abbreviations of the form

$$\underset{x,y \in A}{\forall} P(x, y) \quad \text{for} \quad \underset{x \in A}{\forall} \underset{y \in A}{\forall} P(x, y), \tag{1.21a}$$

$$\underset{x,y \in A}{\exists} P(x, y) \quad \text{for} \quad \underset{x \in A}{\exists} \underset{y \in A}{\exists} P(x, y). \tag{1.21b}$$

Generalizing rules (1.20), we can always commute *identical* quantifiers. Caveat: Quantifiers that are not identical must not be commuted (see Ex. 1.36(d) below).

Example 1.36. (a) Negation of universal and existential statements:

$$\text{Negation of (1.13a)}: \quad \underset{x \in \mathbb{N}}{\exists} \overbrace{x + 1 \leq 0}^{\neg(x+1>0)}. \ (\text{F}) \tag{1.22a}$$

$$\text{Negation of (1.13b)}: \quad \underset{x \in \mathbb{R}}{\exists} \overbrace{x \leq 0}^{\neg(x>0)}. \ (\text{T}) \tag{1.22b}$$

$$\text{Negation of (1.13c)}: \quad \underset{n \in \mathbb{N}}{\forall} \overbrace{n \leq 10}^{\neg(n>10)}. \ (\text{F}) \tag{1.22c}$$

$$\text{Negation of (1.13d)}: \quad \underset{x \in \mathbb{R}}{\forall} \overbrace{x^2 \neq -1}^{\neg(x^2=-1)}. \ (\text{T}) \tag{1.22d}$$

$$\text{Negation of (1.13e)}: \quad \underset{n \in \mathbb{N}}{\exists} \underset{m \in \mathbb{N}}{\forall} \overbrace{m \leq n}^{\neg(m>n)}. \ (\text{F}) \tag{1.22e}$$

(b) As a more complicated example, consider the negation of the uniqueness statement (1.15), i.e. of (1.16):

$$\neg \underset{x \in A}{\exists !} P(x) \quad \Leftrightarrow \quad \neg \underset{x \in A}{\exists} \left(P(x) \wedge \underset{y \in A}{\forall} \big(P(y) \Rightarrow x = y \big) \right)$$

$$\overset{\text{(1.18b), Th. 1.11(a)}}{\Leftrightarrow} \quad \underset{x \in A}{\forall} \neg \left(P(x) \wedge \underset{y \in A}{\forall} \big(\neg P(y) \vee x = y \big) \right)$$

$$\overset{\text{Th. 1.11(i)}}{\Leftrightarrow} \quad \underset{x \in A}{\forall} \left(\neg P(x) \vee \neg \underset{y \in A}{\forall} \big(\neg P(y) \vee x = y \big) \right)$$

$$\overset{\text{(1.18a)}}{\Leftrightarrow} \quad \underset{x \in A}{\forall} \left(\neg P(x) \vee \underset{y \in A}{\exists} \neg \big(\neg P(y) \vee x = y \big) \right)$$

$$\overset{\text{Th. 1.11(j),(k)}}{\Leftrightarrow} \quad \underset{x \in A}{\forall} \left(\neg P(x) \vee \underset{y \in A}{\exists} \big(P(y) \wedge x \neq y \big) \right)$$

$$\overset{\text{Th. 1.11(a)}}{\Leftrightarrow} \quad \underset{x \in A}{\forall} \left(P(x) \Rightarrow \underset{y \in A}{\exists} \big(P(y) \wedge x \neq y \big) \right). \tag{1.23}$$

So how to decode the expression, we have obtained at the end? It states that if $P(x)$ holds for some $x \in A$, then there must be at least a second, different, element $y \in A$ such that $P(y)$ is true. This is, indeed, precisely the negation of $\underset{x \in A}{\exists!} P(x)$.

(c) Identical quantifiers commute:

$$\underset{x \in \mathbb{R}}{\forall} \ \underset{n \in \mathbb{N}}{\forall} \ x^{2n} \geq 0 \ \Leftrightarrow \ \underset{n \in \mathbb{N}}{\forall} \ \underset{x \in \mathbb{R}}{\forall} \ x^{2n} \geq 0, \tag{1.24a}$$

$$\underset{x \in \mathbb{R}}{\forall} \ \underset{y \in \mathbb{R}}{\exists} \ \underset{n \in \mathbb{N}}{\exists} \ ny > x^2 \ \Leftrightarrow \ \underset{x \in \mathbb{R}}{\forall} \ \underset{n \in \mathbb{N}}{\exists} \ \underset{y \in \mathbb{R}}{\exists} \ ny > x^2. \tag{1.24b}$$

(d) The following example shows that different quantifiers do, in general, not commute (i.e. do not yield equivalent statements when commuted): While the statement

$$\underset{x \in \mathbb{R}}{\forall} \ \underset{y \in \mathbb{R}}{\exists} \ y > x \tag{1.25a}$$

is true (for each real number x, there is a bigger real number y, e.g. $y := x + 1$ will do the job), the statement

$$\underset{y \in \mathbb{R}}{\exists} \ \underset{x \in \mathbb{R}}{\forall} \ y > x \tag{1.25b}$$

is false (for example, since $y > y$ is false). In particular, (1.25a) and (1.25b) are not equivalent.

(e) Even though (1.15) provides useful notation, it is better not to think of $\exists!$ as a quantifier. It is really just an abbreviation for (1.16), and it behaves very differently from $\exists$ and $\forall$: The following examples show that, in general, $\exists!$ commutes neither with $\exists$, nor with itself:

$$\underset{n \in \mathbb{N}}{\exists} \ \underset{m \in \mathbb{N}}{\exists!} \ m < n \ \not\Leftrightarrow \ \underset{m \in \mathbb{N}}{\exists!} \ \underset{n \in \mathbb{N}}{\exists} \ m < n$$

(the statement on the left is true, as one can choose $n = 2$, but the statement on the right is false, as $\underset{n \in \mathbb{N}}{\exists} \ m < n$ holds for every $m \in \mathbb{N}$). Similarly,

$$\underset{n \in \mathbb{N}}{\exists!} \ \underset{m \in \mathbb{N}}{\exists!} \ m < n \ \not\Leftrightarrow \ \underset{m \in \mathbb{N}}{\exists!} \ \underset{n \in \mathbb{N}}{\exists!} \ m < n$$

(the statement on the left is still true and the statement on the right is still false (there is no $m \in \mathbb{N}$ such that $\underset{n \in \mathbb{N}}{\exists!} \ m < n$)).

Remark 1.37. One can make the following observations regarding the strategy for proving universal and existential statements:

(a) To prove that $\underset{x \in A}{\forall} P(x)$ is true, one must check the truth of $P(x)$ for every element $x \in A$ – examples are *not* enough!

(b) To prove that $\underset{x \in A}{\forall} P(x)$ is false, it suffices to find *one* $x \in A$ such that $P(x)$ is false – such an x is then called a *counterexample* and *one* counterexample is always enough to prove $\underset{x \in A}{\forall} P(x)$ is false!

(c) To prove that $\underset{x \in A}{\exists} P(x)$ is true, it suffices to find *one* $x \in A$ such that $P(x)$ is true – such an x is then called an *example* and *one* example is always enough to prove $\underset{x \in A}{\exists} P(x)$ is true!

———

The subfield of mathematical logic dealing with quantified statements is called *predicate calculus*. In general, one does not restrict the quantified variables to range only over elements of sets (as we have done above). Again, we refer to [EFT07] for a deeper treatment of the subject.

As an application of quantified statements, let us generalize the notion of union and intersection:

Definition 1.38. Let $I \neq \emptyset$ be a nonempty set, usually called an *index set* in the present context. For each $i \in I$, let A_i denote a set (some or all of the A_i can be identical).

(a) The *intersection*

$$\bigcap_{i \in I} A_i := \left\{ x : \underset{i \in I}{\forall} \ x \in A_i \right\} \tag{1.26a}$$

consists of all elements x that belong to every A_i.

(b) The *union*

$$\bigcup_{i \in I} A_i := \left\{ x : \underset{i \in I}{\exists} \ x \in A_i \right\} \tag{1.26b}$$

consists of all elements x that belong to at least one A_i. The union is called *disjoint* if, and only if, for each $i, j \in I$, $i \neq j$ implies $A_i \cap A_j = \emptyset$.

Proposition 1.39. *Let $I \neq \emptyset$ be an index set, let M denote a set, and, for each $i \in I$, let A_i denote a set. The following set-theoretic rules hold:*

(a) $\left(\bigcap_{i \in I} A_i \right) \cap M = \bigcap_{i \in I} (A_i \cap M).$

(b) $\left(\bigcup_{i \in I} A_i \right) \cup M = \bigcup_{i \in I} (A_i \cup M).$

(c) $\left(\bigcap_{i \in I} A_i \right) \cup M = \bigcap_{i \in I} (A_i \cup M).$

(d) $\left(\bigcup_{i \in I} A_i \right) \cap M = \bigcup_{i \in I} (A_i \cap M).$

(e) $M \setminus \bigcap_{i \in I} A_i = \bigcup_{i \in I} (M \setminus A_i).$

(f) $M \setminus \bigcup_{i \in I} A_i = \bigcap_{i \in I} (M \setminus A_i).$

Proof. We prove (c) and (e) and leave the remaining proofs as an exercise.

(c):

$$x \in \left(\bigcap_{i \in I} A_i \right) \cup M \; \Leftrightarrow \; x \in M \vee \underset{i \in I}{\forall} \, x \in A_i \; \overset{(*)}{\Leftrightarrow} \; \underset{i \in I}{\forall} \, \left(x \in A_i \vee x \in M \right)$$

$$\Leftrightarrow \; x \in \bigcap_{i \in I}(A_i \cup M).$$

To justify the equivalence at $(*)$, we make use of Th. 1.11(b) and verify $\Rightarrow$ and $\Leftarrow$. For $\Rightarrow$ note that the truth of $x \in M$ implies $x \in A_i \vee x \in M$ is true for each $i \in I$. If $x \in A_i$ is true for each $i \in I$, then $x \in A_i \vee x \in M$ is still true for each $i \in I$. To verify $\Leftarrow$, note that the existence of $i \in I$ such that $x \in M$ implies the truth of $x \in M \vee \underset{i \in I}{\forall} \, x \in A_i$. If $x \in M$ is false for each $i \in I$, then $x \in A_i$ must be true for each $i \in I$, showing $x \in M \vee \underset{i \in I}{\forall} \, x \in A_i$ is true also in this case.

(e):

$$x \in M \setminus \bigcap_{i \in I} A_i \; \Leftrightarrow \; x \in M \wedge \neg \underset{i \in I}{\forall} \, x \in A_i \; \Leftrightarrow \; x \in M \wedge \underset{i \in I}{\exists} \, x \notin A_i$$

$$\Leftrightarrow \; \underset{i \in I}{\exists} \, x \in M \setminus A_i \; \Leftrightarrow \; x \in \bigcup_{i \in I}(M \setminus A_i),$$

completing the proof. ∎

Example 1.40. We have the following identities of sets:

$$\bigcap_{x \in \mathbb{R}} \mathbb{N} = \mathbb{N}, \tag{1.27a}$$

$$\bigcap_{n \in \mathbb{N}} \{1, 2, \ldots, n\} = \{1\}, \tag{1.27b}$$

$$\bigcup_{x \in \mathbb{R}} \mathbb{N} = \mathbb{N}, \tag{1.27c}$$

$$\bigcup_{n \in \mathbb{N}} \{1, 2, \ldots, n\} = \mathbb{N}, \tag{1.27d}$$

$$\mathbb{N} \setminus \bigcup_{n \in \mathbb{N}} \{2n\} = \{1, 3, 5, \ldots\} = \bigcap_{n \in \mathbb{N}} \left(\mathbb{N} \setminus \{2n\} \right). \tag{1.27e}$$

2 Functions and Relations

2.1 Functions

Definition 2.1. Let A, B be sets. Given $x \in A$, $y \in B$, the set

$$(x, y) := \Big\{ \{x\}, \{x, y\} \Big\} \tag{2.1}$$

is called the *ordered pair* (often shortened to just *pair*) consisting of x and y. The set of all such pairs is called the Cartesian product $A \times B$, i.e.

$$A \times B := \{(x, y) : x \in A \wedge y \in B\}. \tag{2.2}$$

Example 2.2. Let A be a set.

$$A \times \emptyset = \emptyset \times A = \emptyset, \tag{2.3a}$$
$$\{1, 2\} \times \{1, 2, 3\} = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3)\} \tag{2.3b}$$
$$\neq \{1, 2, 3\} \times \{1, 2\} = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 1), (3, 2)\}. \tag{2.3c}$$

Also note that, for $x \neq y$,

$$(x, y) = \big\{\{x\}, \{x, y\}\big\} \neq \big\{\{y\}, \{x, y\}\big\} = (y, x). \tag{2.4}$$

Definition 2.3. Given sets A, B, a *function* or *map* f is an assignment rule that assigns to each $x \in A$ a unique $y \in B$. One then also writes $f(x)$ for the element y. The set A is called the *domain* of f, denoted $\mathcal{D}(f)$, and B is called the *range* of f, denoted $\mathcal{R}(f)$. The information about a map f can be concisely summarized by the notation

$$f : A \longrightarrow B, \quad x \mapsto f(x), \tag{2.5}$$

where $x \mapsto f(x)$ is called the *assignment rule* for f, $f(x)$ is called the *image* of x, and x is called a *preimage* of $f(x)$ (the image must be unique, but there might be several preimages). The set

$$\mathrm{graph}(f) := \big\{(x, y) \in A \times B : y = f(x)\big\} \tag{2.6}$$

is called the *graph* of f (not to be confused with pictures visualizing the function f, which are also called graph of f). If one wants to be completely precise, then one identifies the function f with the ordered triple $(A, B, \mathrm{graph}(f))$.

The set of all functions with domain A and range B is denoted by $\mathcal{F}(A, B)$ or B^A, i.e.

$$\mathcal{F}(A, B) := B^A := \big\{(f : A \longrightarrow B) : A = \mathcal{D}(f) \wedge B = \mathcal{R}(f)\big\}. \tag{2.7}$$

Caveat: Some authors reserve the word *map* for continuous functions, but we use function and map synonymously.

Definition 2.4. Let A, B be sets and $f : A \longrightarrow B$ a function.

(a) If T is a subset of A, then

$$f(T) := \{f(x) \in B : x \in T\} \tag{2.8}$$

is called the *image* of T under f.

(b) If U is a subset of B, then

$$f^{-1}(U) := \{x \in A : f(x) \in U\} \tag{2.9}$$

is called the *preimage* or *inverse image* of U under f.

(c) f is called *injective* or *one-to-one* if, and only if, every $y \in B$ has at most one preimage, i.e. if, and only if, the preimage of $\{y\}$ has at most one element:

$$f \text{ injective} \quad \Leftrightarrow \quad \mathop{\forall}_{y \in B} \left(f^{-1}\{y\} = \emptyset \vee \mathop{\exists!}_{x \in A} f(x) = y \right)$$

$$\Leftrightarrow \quad \mathop{\forall}_{x_1, x_2 \in A} \left(x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2) \right). \tag{2.10}$$

(d) f is called *surjective* or *onto* if, and only if, every element of the range of f has a preimage:

$$f \text{ surjective} \quad \Leftrightarrow \quad \mathop{\forall}_{y \in B} \mathop{\exists}_{x \in A} y = f(x) \quad \Leftrightarrow \quad \mathop{\forall}_{y \in B} f^{-1}\{y\} \neq \emptyset. \tag{2.11}$$

(e) f is called *bijective* if, and only if, f is injective and surjective.

Example 2.5. Examples of Functions:

$$f : \{1, 2, 3, 4, 5\} \longrightarrow \{1, 2, 3, 4, 5\}, \qquad f(x) := -x + 6, \tag{2.12a}$$
$$g : \mathbb{N} \longrightarrow \mathbb{N}, \qquad g(n) := 2n, \tag{2.12b}$$
$$h : \mathbb{N} \longrightarrow \{2, 4, 6, \dots\}, \qquad h(n) := 2n, \tag{2.12c}$$

$$\tilde{h} : \mathbb{N} \longrightarrow \{2, 4, 6, \dots\}, \qquad \tilde{h}(n) := \begin{cases} n & \text{for } n \text{ even,} \\ n+1 & \text{for } n \text{ odd,} \end{cases} \tag{2.12d}$$

$$G : \mathbb{N} \longrightarrow \mathbb{R}, \qquad G(n) := n/(n+1), \tag{2.12e}$$
$$F : \mathcal{P}(\mathbb{N}) \longrightarrow \mathcal{P}\big(\mathcal{P}(\mathbb{N})\big), \qquad F(A) := \mathcal{P}(A). \tag{2.12f}$$

Instead of $f(x) := -x + 6$ in (2.12a), one can also write $x \mapsto -x + 6$ and analogously in the other cases. Also note that, in the strict sense, functions g and h are different, since their ranges are different (however, using the following Def. 2.4(a), they have the same *image* in the sense that $g(\mathbb{N}) = h(\mathbb{N})$). Furthermore,

$$f(\{1, 2\}) = \{5, 4\} = f^{-1}(\{1, 2\}), \quad \tilde{h}^{-1}(\{2, 4, 6\}) = \{1, 2, 3, 4, 5, 6\}, \tag{2.13}$$

f is bijective; g is injective, but not surjective; h is bijective; $\tilde{h}$ is surjective, but not injective. Can you figure out if G and F are injective and/or surjective?

Example 2.6. (a) For each nonempty set A, the map $\text{Id} : A \longrightarrow A$, $\text{Id}(x) := x$, is called the *identity* on A. If one needs to emphasize that Id operates on A, then one also writes Id_A instead of Id. The identity is clearly bijective.

(b) Let A, B be nonempty sets. A map $f : A \longrightarrow B$ is called *constant* if, and only if, there exists $c \in B$ such that $f(x) = c$ for each $x \in A$. In that case, one also writes $f \equiv c$, which can be read as "f is identically equal to c". If $f \equiv c$, $\emptyset \neq T \subseteq A$, and $U \subseteq B$, then

$$f(T) = \{c\}, \quad f^{-1}(U) = \begin{cases} A & \text{for } c \in U, \\ \emptyset & \text{for } c \notin U. \end{cases} \tag{2.14}$$

f is injective if, and only if, $A = \{x\}$; f is surjective if, and only if, $B = \{c\}$.

(c) Given $A \subseteq X$, the map
$$\iota : A \longrightarrow X, \quad \iota(x) := x, \qquad (2.15)$$

is called *inclusion* (also *embedding* or *imbedding*). An inclusion is always injective; it is surjective if, and only if $A = X$, i.e. if, and only if, it is the identity on A.

(d) Given $A \subseteq X$ and a map $f : X \longrightarrow B$, the map $g : A \longrightarrow B$, $g(x) = f(x)$, is called the *restriction* of f to A; f is called the *extension* of g to X. In this situation, one also uses the notation $f{\restriction}_A$ for g (some authors prefer the notation $f|_A$ or $f|A$).

Theorem 2.7. *Let $f : A \to B$ be a map, let $\emptyset \neq I$ be an index set, and assume S, T, S_i, $i \in I$, are subsets of A, whereas U, V, U_i, $i \in I$, are subsets of B. Then we have the following rules concerning functions and set-theoretic operations:*

$$f(S \cap T) \;\; \subseteq \;\; f(S) \cap f(T), \qquad (2.16\text{a})$$

$$f\left(\bigcap_{i \in I} S_i\right) \;\; \subseteq \;\; \bigcap_{i \in I} f(S_i), \qquad (2.16\text{b})$$

$$f(S \cup T) \;\; = \;\; f(S) \cup f(T), \qquad (2.16\text{c})$$

$$f\left(\bigcup_{i \in I} S_i\right) \;\; = \;\; \bigcup_{i \in I} f(S_i), \qquad (2.16\text{d})$$

$$f^{-1}(U \cap V) \;\; = \;\; f^{-1}(U) \cap f^{-1}(V), \qquad (2.16\text{e})$$

$$f^{-1}\left(\bigcap_{i \in I} U_i\right) \;\; = \;\; \bigcap_{i \in I} f^{-1}(U_i), \qquad (2.16\text{f})$$

$$f^{-1}(U \cup V) \;\; = \;\; f^{-1}(U) \cup f^{-1}(V), \qquad (2.16\text{g})$$

$$f^{-1}\left(\bigcup_{i \in I} U_i\right) \;\; = \;\; \bigcup_{i \in I} f^{-1}(U_i), \qquad (2.16\text{h})$$

$$f(f^{-1}(U)) \;\; \subseteq \;\; U, \quad f^{-1}(f(S)) \supseteq S, \qquad (2.16\text{i})$$

$$f^{-1}(U \setminus V) \;\; = \;\; f^{-1}(U) \setminus f^{-1}(V). \qquad (2.16\text{j})$$

Proof. We prove (2.16b) (which includes (2.16a) as a special case) and the second part of (2.16i), and leave the remaining cases as exercises.

For (2.16b), one argues

$$y \in f\left(\bigcap_{i \in I} S_i\right) \;\Leftrightarrow\; \underset{x \in A}{\exists} \; \underset{i \in I}{\forall} \left(x \in S_i \wedge y = f(x)\right) \;\Rightarrow\; \underset{i \in I}{\forall} \; y \in f(S_i) \;\Leftrightarrow\; y \in \bigcap_{i \in I} f(S_i).$$

The observation

$$x \in S \;\Rightarrow\; f(x) \in f(S) \;\Leftrightarrow\; x \in f^{-1}(f(S)).$$

establishes the second part of (2.16i). $\blacksquare$

It is an exercise to find counterexamples that show one can not, in general, replace the four subset symbols in (2.16) by equalities (it is possible to find examples with sets that have at most 2 elements).

Definition 2.8. The *composition* of maps f and g with $f : A \longrightarrow B$, $g : C \longrightarrow D$, and $f(A) \subseteq C$ is defined to be the map

$$g \circ f : A \longrightarrow D, \quad (g \circ f)(x) := g\big(f(x)\big). \tag{2.17}$$

The expression $g \circ f$ is read as "g after f" or "g composed with f".

Example 2.9. Consider the maps

$$f : \mathbb{N} \longrightarrow \mathbb{R}, \qquad\qquad n \mapsto n^2, \tag{2.18a}$$
$$g : \mathbb{N} \longrightarrow \mathbb{R}, \qquad\qquad n \mapsto 2n. \tag{2.18b}$$

We obtain $f(\mathbb{N}) = \{1, 4, 9, \dots\} \subseteq \mathcal{D}(g)$, $g(\mathbb{N}) = \{2, 4, 6, \dots\} \subseteq \mathcal{D}(f)$, and the compositions

$$(g \circ f) : \mathbb{N} \longrightarrow \mathbb{R}, \qquad (g \circ f)(n) = g(n^2) = 2n^2, \tag{2.19a}$$
$$(f \circ g) : \mathbb{N} \longrightarrow \mathbb{R}, \qquad (f \circ g)(n) = f(2n) = 4n^2, \tag{2.19b}$$

showing that composing functions is, in general, not commutative, even if the involved functions have the same domain and the same range.

Proposition 2.10. *Consider maps* $f : A \longrightarrow B$, $g : C \longrightarrow D$, $h : E \longrightarrow F$, *satisfying* $f(A) \subseteq C$ *and* $g(C) \subseteq E$.

(a) *Associativity of Compositions:*

$$h \circ (g \circ f) = (h \circ g) \circ f. \tag{2.20}$$

(b) *One has the following law for forming preimages:*

$$\underset{W \in \mathcal{P}(D)}{\forall} (g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W)). \tag{2.21}$$

Proof. (a): Both $h \circ (g \circ f)$ and $(h \circ g) \circ f$ map A into F. So it just remains to prove $\big(h \circ (g \circ f)\big)(x) = \big((h \circ g) \circ f\big)(x)$ for each $x \in A$. One computes, for each $x \in A$,

$$\big(h \circ (g \circ f)\big)(x) = h\big((g \circ f)(x)\big) = h\big(g(f(x))\big) = (h \circ g)(f(x))$$
$$= \big((h \circ g) \circ f\big)(x), \tag{2.22}$$

establishing the case.

(b): Exercise. ∎

Definition 2.11. A function $g : B \longrightarrow A$ is called a *right inverse* (resp. *left inverse*) of a function $f : A \longrightarrow B$ if, and only if, $f \circ g = \mathrm{Id}_B$ (resp. $g \circ f = \mathrm{Id}_A$). Moreover, g is called an *inverse* of f if, and only if, it is both a right and a left inverse. If g is an inverse of f, then one also writes f^{-1} instead of g. The map f is called (*right*, *left*) *invertible* if, and only if, there exists a (right, left) inverse for f.

Example 2.12. (a) Consider the map

$$f : \mathbb{N} \longrightarrow \mathbb{N}, \quad f(n) := 2n. \tag{2.23a}$$

The maps

$$g_1 : \mathbb{N} \longrightarrow \mathbb{N}, \quad g_1(n) := \begin{cases} n/2 & \text{if } n \text{ even,} \\ 1 & \text{if } n \text{ odd,} \end{cases} \tag{2.23b}$$

$$g_2 : \mathbb{N} \longrightarrow \mathbb{N}, \quad g_2(n) := \begin{cases} n/2 & \text{if } n \text{ even,} \\ 2 & \text{if } n \text{ odd,} \end{cases} \tag{2.23c}$$

both constitute left inverses of f. It follows from Th. 2.13(c) below that f does not have a right inverse.

(b) Consider the map

$$f : \mathbb{N} \longrightarrow \mathbb{N}, \quad f(n) := \begin{cases} n/2 & \text{for } n \text{ even,} \\ (n+1)/2 & \text{for } n \text{ odd.} \end{cases} \tag{2.24a}$$

The maps

$$g_1 : \mathbb{N} \longrightarrow \mathbb{N}, \quad g_1(n) := 2n, \tag{2.24b}$$
$$g_2 : \mathbb{N} \longrightarrow \mathbb{N}, \quad g_2(n) := 2n - 1, \tag{2.24c}$$

both constitute right inverses of f. It follows from Th. 2.13(c) below that f does not have a left inverse.

(c) The map

$$f : \mathbb{N} \longrightarrow \mathbb{N}, \quad f(n) := \begin{cases} n - 1 & \text{for } n \text{ even,} \\ n + 1 & \text{for } n \text{ odd,} \end{cases} \tag{2.25a}$$

is its own inverse, i.e. $f^{-1} = f$. For the map

$$g : \mathbb{N} \longrightarrow \mathbb{N}, \quad g(n) := \begin{cases} 2 & \text{for } n = 1, \\ 3 & \text{for } n = 2, \\ 1 & \text{for } n = 3, \\ n & \text{for } n \notin \{1, 2, 3\}, \end{cases} \tag{2.25b}$$

the inverse is

$$g^{-1} : \mathbb{N} \longrightarrow \mathbb{N}, \quad g^{-1}(n) := \begin{cases} 3 & \text{for } n = 1, \\ 1 & \text{for } n = 2, \\ 2 & \text{for } n = 3, \\ n & \text{for } n \notin \{1, 2, 3\}. \end{cases} \tag{2.25c}$$

While Examples 2.12(a),(b) show that left and right inverses are usually not unique, they *are* unique provided f is bijective (see Th. 2.13(c)).

Theorem 2.13. *Let A, B be nonempty sets.*

(a) $f : A \longrightarrow B$ *is right invertible if, and only if, f is surjective (where the implication "⟸" makes use of the axiom of choice (AC), see Appendix A.4).*

(b) $f : A \longrightarrow B$ *is left invertible if, and only if, f is injective.*

(c) $f : A \longrightarrow B$ *is invertible if, and only if, f is bijective. In this case, the right inverse and the left inverse are unique and both identical to the inverse.*

Proof. (a): If f is surjective, then, for each $y \in B$, there exists $x_y \in f^{-1}\{y\}$ such that $f(x_y) = y$. By AC, we can define the choice function

$$g : B \longrightarrow A, \quad g(y) := x_y. \tag{2.26}$$

Then, for each $y \in B$, $f(g(y)) = y$, showing g is a right inverse of f. Conversely, if $g : B \longrightarrow A$ is a right inverse of f, then, for each $y \in B$, it is $y = f(g(y))$, showing that $g(y) \in A$ is a preimage of y, i.e. f is surjective.

(b): Fix $a \in A$. If f is injective, then, for each $y \in B$ with $f^{-1}\{y\} \neq \emptyset$, let x_y denote the unique element in A satisfying $f(x_y) = y$. Define

$$g : B \longrightarrow A, \quad g(y) := \begin{cases} x_y & \text{for } f^{-1}\{y\} \neq \emptyset, \\ a & \text{otherwise.} \end{cases} \tag{2.27}$$

Then, for each $x \in A$, $g(f(x)) = x$, showing g is a left inverse of f. Conversely, if $g : B \longrightarrow A$ is a left inverse of f and $x_1, x_2 \in A$ with $f(x_1) = f(x_2) = y$, then $x_1 = (g \circ f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \circ f)(x_2) = x_2$, showing y has precisely one preimage and f is injective.

The first part of (c) follows immediately by combining (a) and (b) (and, actually, without using AC, since, if f is both injective and surjective, then, for each $y \in B$, the element $x_y \in f^{-1}\{y\}$ is unique, and (2.26) can be defined without AC). It merely remains to verify the uniqueness of right and left inverse for bijective maps. So let g be a left inverse of f, let h be a right inverse of f, and let f^{-1} be an inverse of f. Then, for each $y \in B$,

$$g(y) = \big(g \circ (f \circ f^{-1})\big)(y) = \big((g \circ f) \circ f^{-1}\big)(y) = f^{-1}(y), \tag{2.28a}$$

$$h(y) = \big((f^{-1} \circ f) \circ h\big)(y) = \big(f^{-1} \circ (f \circ h)\big)(y) = f^{-1}(y), \tag{2.28b}$$

thereby proving the uniqueness of left and right inverse for bijective maps. ∎

Theorem 2.14. *Consider maps $f : A \longrightarrow B$, $g : B \longrightarrow C$. If f and g are both injective (resp. both surjective, both bijective), then so is $g \circ f$. Moreover, in the bijective case, one has*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}. \tag{2.29}$$

Proof. Exercise.   ■

Definition 2.15. (a) Given an index set I and a set A, a map $f : I \longrightarrow A$ is sometimes called a *family* (of elements in A), and is denoted in the form $f = (a_i)_{i \in I}$ with $a_i := f(i)$. When using this representation, one often does not even specify f and A, especially if the a_i are themselves sets.

(b) A *sequence* in a set A is a family of elements in A, where the index set is the set of natural numbers $\mathbb{N}$. In this case, one writes $(a_n)_{n \in \mathbb{N}}$ or $(a_1, a_2, \dots)$. More generally, a family is called a *sequence*, given a bijective map between the index set I and a subset of $\mathbb{N}$.

(c) Given a family of sets $(A_i)_{i \in I}$, we define the *Cartesian product* of the A_i to be the set of functions

$$\prod_{i \in I} A_i := \left\{ \left(f : I \longrightarrow \bigcup_{j \in I} A_j \right) : \underset{i \in I}{\forall} \, f(i) \in A_i \right\}. \tag{2.30}$$

If I has precisely n elements with $n \in \mathbb{N}$, then the elements of the Cartesian product $\prod_{i \in I} A_i$ are called (ordered) *n-tuples*, (ordered) *triples* for $n = 3$.

Example 2.16. (a) Using the notion of family, we can now say that the intersection $\bigcap_{i \in I} A_i$ and union $\bigcup_{i \in I} A_i$ as defined in Def. 1.38 are the intersection and union of the family of sets $(A_i)_{i \in I}$, respectively. As a concrete example, let us revisit (1.27b), where we have

$$(A_n)_{n \in \mathbb{N}}, \quad A_n := \{1, 2, \dots, n\}, \quad \bigcap_{n \in \mathbb{N}} A_n = \{1\}. \tag{2.31}$$

(b) Examples of Sequences:

Sequence in $\{0,1\}$:	$(1, 0, 1, 0, 1, 0, \dots)$,	(2.32a)
Sequence in $\mathbb{N}$:	$(n^2)_{n \in \mathbb{N}} = (1, 4, 9, 16, 25, \dots)$,	(2.32b)
Sequence in $\mathbb{R}$:	$((-1)^n \sqrt{n})_{n \in \mathbb{N}} = \left(-1, \sqrt{2}, -\sqrt{3}, \dots\right)$,	(2.32c)
Sequence in $\mathbb{R}$:	$(1/n)_{n \in \mathbb{N}} = \left(1, \frac{1}{2}, \frac{1}{3}, \dots\right)$,	(2.32d)
Finite Sequence in $\mathcal{P}(\mathbb{N})$:	$(\{3, 2, 1\}, \{2, 1\}, \{1\}, \emptyset)$.	(2.32e)

(c) The Cartesian product $\prod_{i \in I} A$, where all sets $A_i = A$, is the same as A^I, the set of all functions from I into A. So, for example, $\prod_{n \in \mathbb{N}} \mathbb{R} = \mathbb{R}^{\mathbb{N}}$ is the set of all sequences in $\mathbb{R}$. If $I = \{1, 2, \dots, n\}$ with $n \in \mathbb{N}$, then

$$\prod_{i \in I} A = A^{\{1, 2 \dots, n\}} =: \prod_{i=1}^{n} A =: A^n \tag{2.33}$$

is the set of all n-tuples with entries from A.

—

In the following, we explain the common notation 2^A for the power set $\mathcal{P}(A)$ of a set A. It is related to a natural identification between subsets and their corresponding characteristic function.

Definition 2.17. Let A be a set and let $B \subseteq A$ be a subset of A. Then

$$\chi_B : A \longrightarrow \{0, 1\}, \quad \chi_B(x) := \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{if } x \notin B, \end{cases} \tag{2.34}$$

is called the *characteristic function* of the set B (with respect to the universe A). One also finds the notations 1_B and $\mathbb{1}_B$ instead of χ_B (note that all the notations supress the dependence of the characteristic function on the universe A).

Proposition 2.18. *Let A be a set. Then the map*

$$\chi : \mathcal{P}(A) \longrightarrow \{0, 1\}^A, \quad \chi(B) := \chi_B, \tag{2.35}$$

is bijective (recall that $\mathcal{P}(A)$ denotes the power set of A and $\{0, 1\}^A$ denotes the set of all functions from A into $\{0, 1\}$).

Proof. χ is injective: Let $B, C \in \mathcal{P}(A)$ with $B \neq C$. By possibly switching the names of B and C, we may assume there exists $x \in B$ such that $x \notin C$. Then $\chi_B(x) = 1$, whereas $\chi_C(x) = 0$, showing $\chi(B) \neq \chi(C)$, proving χ is injective.

χ is surjective: Let $f : A \longrightarrow \{0, 1\}$ be an arbitrary function and define $B := \{x \in A : f(x) = 1\}$. Then $\chi(B) = \chi_B = f$, proving χ is surjective. ∎

Proposition 2.18 allows one to identify the sets $\mathcal{P}(A)$ and $\{0, 1\}^A$ via the bijective map χ. This fact together with the common practise of set theory to identify the number 2 with the set $\{0, 1\}$ (cf. the first paragraph of Sec. D.1 in the Appendix) explains the notation 2^A for $\mathcal{P}(A)$.

2.2 Relations

Definition 2.19. Given sets A and B, a *relation* is a subset R of $A \times B$ (if one wants to be completely precise, a relation is an ordered triple (A, B, R), where $R \subseteq A \times B$). If $A = B$, then we call R a relation on A. One says that $a \in A$ and $b \in B$ are *related* according to the relation R if, and only if, $(a, b) \in R$. In this context, one usually writes $a \, R \, b$ instead of $(a, b) \in R$.

Example 2.20. (a) The relations we are probably most familiar with are $=$ and $\leq$. The relation R of equality, usually denoted $=$, makes sense on every nonempty set A:

$$R := \Delta(A) := \{(x, x) \in A \times A : x \in A\}. \tag{2.36}$$

The set $\Delta(A)$ is called the *diagonal* of the Cartesian product, i.e., as a subset of $A \times A$, the relation of equality is identical to the diagonal:

$$x = y \Leftrightarrow x \, R \, y \Leftrightarrow (x, y) \in R = \Delta(A). \tag{2.37}$$

Similarly, the relation $\leq$ on $\mathbb{R}$ is identical to the set

$$R_{\leq} := \{(x, y) \in \mathbb{R}^2 : x \leq y\}. \tag{2.38}$$

(b) Every function $f : A \longrightarrow B$ is a relation, namely the relation

$$R_f = \{(x, y) \in A \times B : y = f(x)\} = \mathrm{graph}(f). \tag{2.39}$$

Conversely, if $B \neq \emptyset$, then every relation $R \subseteq A \times B$ uniquely corresponds to the function

$$f_R : A \longrightarrow \mathcal{P}(B), \quad f_R(x) = \{y \in B : x\,R\,y\}. \tag{2.40}$$

Definition 2.21. Let R be a relation on the set A.

(a) R is called *reflexive* if, and only if,

$$\underset{x \in A}{\forall}\ x\,R\,x, \tag{2.41}$$

i.e. if, and only if, every element is related to itself.

(b) R is called *symmetric* if, and only if,

$$\underset{x,y \in A}{\forall}\ \big(x\,R\,y\ \Rightarrow\ y\,R\,x\big), \tag{2.42}$$

i.e. if, and only if, each x is related to y if, and only if, y is related to x.

(c) R is called *antisymmetric* if, and only if,

$$\underset{x,y \in A}{\forall}\ \big((x\,R\,y\ \wedge\ y\,R\,x)\ \Rightarrow\ x = y\big), \tag{2.43}$$

i.e. if, and only if, the only possibility for x to be related to y at the same time that y is related to x is in the case $x = y$.

(d) R is called *transitive* if, and only if,

$$\underset{x,y,z \in A}{\forall}\ \big((x\,R\,y\ \wedge\ y\,R\,z)\ \Rightarrow\ x\,R\,z\big), \tag{2.44}$$

i.e. if, and only if, the relatedness of x and y together with the relatedness of y and z implies the relatedness of x and z.

Example 2.22. The relations $=$ and $\leq$ on $\mathbb{R}$ (or $\mathbb{N}$) are reflexive, antisymmetric, and transitive; $=$ is also symmetric, whereas $\leq$ is not; $<$ is antisymmetric (since $x < y \wedge y < x$ is always false) and transitive, but neither reflexive nor symmetric. The relation

$$R := \big\{(x, y) \in \mathbb{N}^2 : (x, y \text{ are both even}) \vee (x, y \text{ are both odd})\big\} \tag{2.45}$$

on $\mathbb{N}$ is not antisymmetric, but reflexive, symmetric, and transitive. The relation

$$S := \{(x, y) \in \mathbb{N}^2 : y = x^2\} \tag{2.46}$$

is not transitive (for example, $2\,S\,4$ and $4\,S\,16$, but not $2\,S\,16$), not reflexive, not symmetric; it is only antisymmetric.

Definition 2.23. A relation R on a set A is called an *equivalence relation* if, and only if, R is reflexive, symmetric, and transitive. If R is an equivalence relations, then one often writes $x \sim y$ instead of $x\,R\,y$.

Example 2.24. (a) The equality relation $=$ is an equivalence relation on each $A \neq \emptyset$.

(b) The relation R defined in (2.45) is an equivalence relation on $\mathbb{N}$.

(c) Given a disjoint union $A = \dot{\bigcup}_{i \in I} A_i$ with every $A_i \neq \emptyset$ (which is sometimes called a *decomposition* of A), an equivalence relation on A is defined by

$$x \sim y \;\Leftrightarrow\; \underset{i \in I}{\exists} \left(x \in A_i \wedge y \in A_i \right). \tag{2.47}$$

Conversely, given an equivalence relation $\sim$ on a nonempty set A, we can construct a decomposition $A = \dot{\bigcup}_{i \in I} A_i$ such that (2.47) holds: For each $x \in A$, define

$$[x] := \{ y \in A : x \sim y \}, \tag{2.48}$$

called the *equivalence class* of x; each $y \in [x]$ is called a *representative* of $[x]$. One verifies that the properties of $\sim$ guarantee

$$\left([x] = [y] \Leftrightarrow x \sim y \right) \quad \wedge \quad \left([x] \cap [y] = \emptyset \Leftrightarrow \neg(x \sim y) \right). \tag{2.49}$$

The set of all equivalence classes $I := A/\!\sim\; := \{ [x] : x \in A \}$ is called the *quotient set* of A by $\sim$, and $A = \dot{\bigcup}_{i \in I} A_i$ with $A_i := i$ for each $i \in I$ is the desired decomposition of A.

Definition 2.25. A relation R on a set A is called a *partial order* if, and only if, R is reflexive, antisymmetric, and transitive. If R is a partial order, then one usually writes $x \leq y$ instead of $x\,R\,y$. A partial order $\leq$ is called a *total* or *linear order* if, and only if, for each $x, y \in A$, one has $x \leq y$ or $y \leq x$.

Notation 2.26. Given a (partial or total) order $\leq$ on $A \neq \emptyset$, we write $x < y$ if, and only if, $x \leq y$ and $x \neq y$, calling $<$ the *strict* order corresponding to $\leq$ (note that the strict order is never a partial order).

Definition 2.27. Let $\leq$ be a partial order on $A \neq \emptyset$, $\emptyset \neq B \subseteq A$.

(a) $x \in A$ is called *lower* (resp. *upper*) *bound* for B if, and only if, $x \leq b$ (resp. $b \leq x$) for each $b \in B$. Moreover, B is called *bounded from below* (resp. *from above*) if, and only if, there exists a lower (resp. upper) bound for B; B is called *bounded* if, and only if, it is bounded from above and from below.

(b) $x \in B$ is called *minimum* or just *min* (resp. *maximum* or *max*) of B if, and only if, x is a lower (resp. upper) bound for B. One writes $x = \min B$ if x is minimum and $x = \max B$ if x is maximum.

(c) A maximum of the set of lower bounds of B (i.e. a largest lower bound) is called *infimum* of B, denoted $\inf B$; a minimum of the set of upper bounds of B (i.e. a smallest upper bound) is called *supremum* of B, denoted $\sup B$.

Example 2.28. (a) For each $A \subseteq \mathbb{R}$, the usual relation $\leq$ defines a total order on A. For $A = \mathbb{R}$, we see that $\mathbb{N}$ has 0 and 1 as lower bound with $1 = \min \mathbb{N} = \inf \mathbb{N}$. On the other hand, $\mathbb{N}$ is unbounded from above. The set $M := \{1, 2, 3\}$ is bounded with $\min M = 1$, $\max M = 3$. The positive real numbers $\mathbb{R}^+ := \{x \in \mathbb{R} : x > 0\}$ have $\inf \mathbb{R}^+ = 0$, but they do not have a minimum (if $x > 0$, then $0 < x/2 < x$).

(b) Consider $A := \mathbb{N} \times \mathbb{N}$. Then

$$(m_1, m_2) \leq (n_1, n_2) \iff m_1 \leq n_1 \ \wedge \ m_2 \leq n_2, \tag{2.50}$$

defines a partial order on A that is not a total order (for example, neither $(1, 2) \leq (2, 1)$ nor $(2, 1) \leq (1, 2)$). For the set

$$B := \big\{(1, 1), (2, 1), (1, 2)\big\}, \tag{2.51}$$

we have $\inf B = \min B = (1, 1)$, B does not have a max, but $\sup B = (2, 2)$ (if $(m, n) \in A$ is an upper bound for B, then $(2, 1) \leq (m, n)$ implies $2 \leq m$ and $(1, 2) \leq (m, n)$ implies $2 \leq n$, i.e. $(2, 2) \leq (m, n)$; since $(2, 2)$ is clearly an upper bound for B, we have proved $\sup B = (2, 2)$).

A different order on A is the so-called *lexicographic order* defined by

$$(m_1, m_2) \leq (n_1, n_2) \iff m_1 < n_1 \ \vee \ (m_1 = n_1 \ \wedge \ m_2 \leq n_2). \tag{2.52}$$

In contrast to the order from (2.50), the lexicographic order does define a total order on A.

Lemma 2.29. *Let $\leq$ be a partial order on $A \neq \emptyset$, $\emptyset \neq B \subseteq A$. Then the relation $\geq$, defined by*

$$x \geq y \iff y \leq x, \tag{2.53}$$

is also a partial order on A. Moreover, using obvious notation, we have, for each $x \in A$,

$x \leq$-*lower bound for B*	$\iff$	$x \geq$-*upper bound for B,*	(2.54a)
$x \leq$-*upper bound for B*	$\iff$	$x \geq$-*lower bound for B,*	(2.54b)
$x = \min_{\leq} B$	$\iff$	$x = \max_{\geq} B,$	(2.54c)
$x = \max_{\leq} B$	$\iff$	$x = \min_{\geq} B,$	(2.54d)
$x = \inf_{\leq} B$	$\iff$	$x = \sup_{\geq} B,$	(2.54e)
$x = \sup_{\leq} B$	$\iff$	$x = \inf_{\geq} B.$	(2.54f)

Proof. Reflexivity, antisymmetry, and transitivity of $\leq$ clearly imply the same properties for $\geq$, respectively. Moreover

$$x \leq\text{-lower bound for } B \iff \underset{b \in B}{\forall} \ x \leq b \iff \underset{b \in B}{\forall} \ b \geq x \iff x \geq\text{-upper bound for } B,$$

proving (2.54a). Analogously, we obtain (2.54b). Next, (2.54c) and (2.54d) are implied by (2.54a) and (2.54b), respectively. Finally, (2.54e) is proved by

$$x = \inf_{\leq} B \iff x = \max_{\leq} \{y \in A : y \leq\text{-lower bound for } B\}$$
$$\iff x = \min_{\geq} \{y \in A : y \geq\text{-upper bound for } B\} \iff x = \sup_{\geq} B,$$

and (2.54f) follows analogously. ∎

Proposition 2.30. *Let $\leq$ be a partial order on $A \neq \emptyset$, $\emptyset \neq B \subseteq A$. The elements $\max B$, $\min B$, $\sup B$, $\inf B$ are all unique, provided they exist.*

Proof. Exercise. ∎

Definition 2.31. Let A, B be nonempty sets with partial orders, both denoted by $\leq$ (even though they might be different). A function $f : A \longrightarrow B$, is called (*strictly*) *isotone*, *order-preserving*, or *increasing* if, and only if,

$$\underset{x,y \in A}{\forall} \left(x < y \implies f(x) \leq f(y) \;\; (\text{resp. } f(x) < f(y)) \right); \tag{2.55a}$$

f is called (*strictly*) *antitone*, *order-reversing*, or *decreasing* if, and only if,

$$\underset{x,y \in A}{\forall} \left(x < y \implies f(x) \geq f(y) \;\; (\text{resp. } f(x) > f(y)) \right). \tag{2.55b}$$

Functions that are (strictly) isotone or antitone are called (strictly) *monotone*.

Proposition 2.32. *Let A, B be nonempty sets with partial orders, both denoted by $\leq$.*

(a) *A (strictly) isotone function $f : A \longrightarrow B$ becomes a (strictly) antitone function and vice versa if precisely one of the relations $\leq$ is replaced by $\geq$.*

(b) *If the order $\leq$ on A is total and $f : A \longrightarrow B$ is strictly isotone or strictly antitone, then f is one-to-one.*

(c) *If the order $\leq$ on A is total and $f : A \longrightarrow B$ is invertible and strictly isotone (resp. antitone), then f^{-1} is also strictly isotone (resp. antitone).*

Proof. (a) is immediate from (2.55).

(b): Due to (a), it suffices to consider the case that f is strictly isotone. If f is strictly isotone and $x \neq y$, then $x < y$ or $y < x$ since the order on A is total. Thus, $f(x) < f(y)$ or $f(y) < f(x)$, i.e. $f(x) \neq f(y)$ in every case, showing f is one-to-one.

(c): Again, due to (a), it suffices to consider the isotone case. If $u, v \in B$ such that $u < v$, then $u = f(f^{-1}(u))$, $v = f(f^{-1}(v))$, and the isotonicity of f imply $f^{-1}(u) < f^{-1}(v)$ (we are using that the order on A is total – otherwise, $f^{-1}(u)$ and $f^{-1}(v)$ need not be comparable). ∎

Example 2.33. **(a)** $f : \mathbb{N} \longrightarrow \mathbb{N}$, $f(n) := 2n$, is strictly increasing, every constant map on $\mathbb{N}$ is both increasing and decreasing, but not strictly increasing or decreasing. All maps occurring in (2.25) are neither increasing nor decreasing.

(b) The map $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) := -2x$, is invertible and strictly decreasing, and so is $f^{-1} : \mathbb{R} \longrightarrow \mathbb{R}$, $f^{-1}(x) := -x/2$.

(c) The following counterexamples show that the assertions of Prop. 2.32(b),(c) are no longer correct if one does not assume the order on A is total. Let A be the set from (2.51) (where it had been called B) with the (nontotal) order from (2.50). The map

$$f : A \longrightarrow \mathbb{N}, \quad \begin{cases} f(1,1) := 1, \\ f(1,2) := 2, \\ f(2,1) := 2, \end{cases} \tag{2.56}$$

is strictly isotone, but not one-to-one. The map

$$f : A \longrightarrow \{1,2,3\}, \quad \begin{cases} f(1,1) := 1, \\ f(1,2) := 2, \\ f(2,1) := 3, \end{cases} \tag{2.57}$$

is strictly isotone and invertible, however f^{-1} is not isotone (since $2 < 3$, but $f^{-1}(2) = (1,2)$ and $f^{-1}(3) = (2,1)$ are not comparable, i.e. $f^{-1}(2) \leq f^{-1}(3)$ is *not* true).

3 Natural Numbers, Induction, and the Size of Sets

3.1 Induction and Recursion

One of the most useful proof techniques is the method of induction – it is used in situations, where one needs to verify the truth of statements $\phi(n)$ for each $n \in \mathbb{N}$, i.e. the truth of the statement

$$\underset{n \in \mathbb{N}}{\forall} \phi(n). \tag{3.1}$$

Induction is based on the fact that $\mathbb{N}$ satisfies the so-called *Peano axioms*:

P1: $\mathbb{N}$ contains a special element called *one*, denoted 1.

P2: There exists an injective map $S : \mathbb{N} \longrightarrow \mathbb{N} \setminus \{1\}$, called the *successor function* (for each $n \in \mathbb{N}$, $S(n)$ is called the *successor* of n).

P3: If a subset A of $\mathbb{N}$ has the property that $1 \in A$ and $S(n) \in A$ for each $n \in A$, then A is equal to $\mathbb{N}$. Written as a formula, the third axiom is:

$$\underset{A \in \mathcal{P}(\mathbb{N})}{\forall} \big(1 \in A \,\wedge\, S(A) \subseteq A \,\Rightarrow\, A = \mathbb{N}\big).$$

Remark 3.1. In Def. 1.28, we had introduced the natural numbers $\mathbb{N} := \{1,2,3,\dots\}$. The successor function is $S(n) = n + 1$. In axiomatic set theory, one starts with the Peano axioms and shows that the axioms of set theory allow the construction of a set $\mathbb{N}$ which satisfies the Peano axioms. One then *defines* $2 := S(1)$, $3 := S(2)$, ..., $n + 1 := S(n)$. The interested reader can find more details in Appendix D.1.

Theorem 3.2 (Principle of Induction). *Suppose, for each $n \in \mathbb{N}$, $\phi(n)$ is a statement (i.e. a predicate of n in the language of Def. 1.31). If* (a) *and* (b) *both hold, where*

(a) $\phi(1)$ *is true,*

(b) $\underset{n \in \mathbb{N}}{\forall} \big(\phi(n) \Rightarrow \phi(n+1)\big),$

then (3.1) *is true, i.e. $\phi(n)$ is true for every $n \in \mathbb{N}$.*

Proof. Let $A := \{n \in \mathbb{N} : \phi(n)\}$. We have to show $A = \mathbb{N}$. Since $1 \in A$ by (a), and

$$n \in A \;\Rightarrow\; \phi(n) \overset{\text{(b)}}{\Rightarrow} \phi(n+1) \Rightarrow S(n) = n+1 \in A, \tag{3.2}$$

i.e. $S(A) \subseteq A$, the Peano axiom P3 implies $A = \mathbb{N}$. ∎

Remark 3.3. To prove some $\phi(n)$ for each $n \in \mathbb{N}$ by induction according to Th. 3.2 consists of the following two steps:

(a) Prove $\phi(1)$, the so-called *base case*.

(b) Perform the *inductive step*, i.e. prove that $\phi(n)$ (the *induction hypothesis*) implies $\phi(n+1)$.

Example 3.4. We use induction to prove the statement

$$\underset{n \in \mathbb{N}}{\forall} \underbrace{\left(1 + 2 + \cdots + n = \frac{n(n+1)}{2}\right)}_{\phi(n)} : \tag{3.3}$$

Base Case ($n = 1$): $1 = \frac{1 \cdot 2}{2}$, i.e. $\phi(1)$ is true.

Induction Hypothesis: Assume $\phi(n)$, i.e. $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ holds.

Induction Step: One computes

$$1 + 2 + \cdots + n + (n+1) \overset{\left(\phi(n)\right)}{=} \frac{n(n+1)}{2} + n + 1 = \frac{n(n+1) + 2n + 2}{2}$$

$$= \frac{n^2 + 3n + 2}{2} = \frac{(n+1)(n+2)}{2}, \tag{3.4}$$

i.e. $\phi(n+1)$ holds and the induction is complete.

Corollary 3.5. *Theorem 3.2 remains true if* (b) *is replaced by*

$$\underset{n \in \mathbb{N}}{\forall} \left(\left(\underset{1 \leq m \leq n}{\forall} \phi(m)\right) \Rightarrow \phi(n+1)\right). \tag{3.5}$$

Proof. If, for each $n \in \mathbb{N}$, we use $\psi(n)$ to denote $\bigvee_{1 \leq m \leq n} \phi(m)$, then (3.5) is equivalent to $\bigvee_{n \in \mathbb{N}} \big(\psi(n) \Rightarrow \psi(n+1)\big)$, i.e. to Th. 3.2(b) with ϕ replaced by ψ. Thus, Th. 3.2 implies $\psi(n)$ holds true for each $n \in \mathbb{N}$, i.e. $\phi(n)$ holds true for each $n \in \mathbb{N}$. ∎

Corollary 3.6. *Let I be an index set. Suppose, for each $i \in I$, $\phi(i)$ is a statement. If there is a bijective map $f : \mathbb{N} \longrightarrow I$ and (a) and (b) both hold, where*

(a) $\phi\big(f(1)\big)$ *is true,*

(b) $\bigvee_{n \in \mathbb{N}} \Big(\phi\big(f(n)\big) \Rightarrow \phi\big(f(n+1)\big)\Big)$,

then $\phi(i)$ is true for every $i \in I$.

Finite Induction: The above assertion remains true if $f : \{1, \ldots, m\} \longrightarrow I$ is bijective for some $m \in \mathbb{N}$ and $\mathbb{N}$ in (b) is replaced by $\{1, \ldots, m-1\}$.

Proof. If, for each $n \in \mathbb{N}$, we use $\psi(n)$ to denote $\phi\big(f(n)\big)$, then Th. 3.2 shows $\psi(n)$ is true for every $n \in \mathbb{N}$. Given $i \in I$, we have $n := f^{-1}(i) \in \mathbb{N}$ with $f(n) = i$, showing that $\phi(i) = \phi\big(f(n)\big) = \psi(n)$ is true.

For the finite induction, let $\psi(n)$ denote $\big(n \leq m \wedge \phi(f(n))\big) \vee n > m$. Then, for $1 \leq n < m$, we have $\psi(n) \Rightarrow \psi(n+1)$ due to (b). For $n \geq m$, we also have $\psi(n) \Rightarrow \psi(n+1)$ due to $n \geq m \Rightarrow n+1 > m$. Thus, Th. 3.2 shows $\psi(n)$ is true for every $n \in \mathbb{N}$. Given $i \in I$, it is $n := f^{-1}(i) \in \{1, \ldots, m\}$ with $f(n) = i$. Since $n \leq m \wedge \psi(n) \Rightarrow \phi\big(f(n)\big)$, we obtain that $\phi(i)$ is true. ∎

Apart from providing a widely employable proof technique, the most important application of Th. 3.2 is the possibility to define sequences inductively, using so-called recursion:

Theorem 3.7 (Recursion Theorem). *Let A be a nonempty set and $x \in A$. Given a sequence of functions $(f_n)_{n \in \mathbb{N}}$, where $f_n : A^n \longrightarrow A$, there exists a unique sequence $(x_n)_{n \in \mathbb{N}}$ in A satisfying the following two conditions:*

(i) $x_1 = x$.

(ii) $\bigvee_{n \in \mathbb{N}} x_{n+1} = f_n(x_1, \ldots, x_n)$.

The same holds if $\mathbb{N}$ is replaced by an index set I as in Cor. 3.6.

Proof. To prove uniqueness, let $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ be sequences in A, both satisfying (i) and (ii), i.e.

$$x_1 = y_1 = x \quad \text{and} \tag{3.6a}$$

$$\bigvee_{n \in \mathbb{N}} \big(x_{n+1} = f_n(x_1, \ldots, x_n) \wedge y_{n+1} = f_n(y_1, \ldots, y_n)\big). \tag{3.6b}$$

We prove by induction (in the form of Cor. 3.5) that $(x_n)_{n \in \mathbb{N}} = (y_n)_{n \in \mathbb{N}}$, i.e.

$$\underset{n \in \mathbb{N}}{\forall} \; \underbrace{x_n = y_n}_{\phi(n)} : \tag{3.7}$$

Base Case ($n = 1$): $\phi(1)$ is true according to (3.6a).

Induction Hypothesis: Assume $\phi(m)$ for each $m \in \{1, \ldots, n\}$, i.e. $x_m = y_m$ holds for each $m \in \{1, \ldots, n\}$.

Induction Step: One computes

$$x_{n+1} \overset{(3.6b)}{=} f_n(x_1, \ldots, x_n) \overset{\left(\phi(1), \ldots, \phi(n)\right)}{=} f_n(y_1, \ldots, y_n) \overset{(3.6b)}{=} y_{n+1}, \tag{3.8}$$

i.e. $\phi(n+1)$ holds and the induction is complete.

To prove existence, we have to show that there is a function $F : \mathbb{N} \longrightarrow A$ such that the following two conditions hold:

$$F(1) = x, \tag{3.9a}$$

$$\underset{n \in \mathbb{N}}{\forall} \; F(n+1) = f_n\big(F(1), \ldots, F(n)\big). \tag{3.9b}$$

To this end, let

$$\mathcal{F} := \left\{ B \subseteq \mathbb{N} \times A : (1, x) \in B \wedge \underset{\substack{n \in \mathbb{N}, \\ (1, a_1), \ldots, (n, a_n) \in B}}{\forall} \big(n+1, f_n(a_1, \ldots, a_n)\big) \in B \right\} \tag{3.10}$$

and

$$G := \bigcap_{B \in \mathcal{F}} B. \tag{3.11}$$

Note that G is well-defined, as $\mathbb{N} \times A \in \mathcal{F}$. Also, clearly, $G \in \mathcal{F}$. We would like to define F such that $G = \operatorname{graph}(F)$. For this to be possible, we will show, by induction,

$$\underset{n \in \mathbb{N}}{\forall} \; \underbrace{\underset{x_n \in A}{\exists !} \; (n, x_n) \in G}_{\phi(n)}. \tag{3.12}$$

Base Case ($n = 1$): From the definition of G, we know $(1, x) \in G$. If $(1, a) \in G$ with $a \neq x$, then $H := G \setminus \{(1, a)\} \in \mathcal{F}$, implying $G \subseteq H$ in contradiction to $(1, a) \notin H$. This shows $a = x$ and proves $\phi(1)$.

Induction Hypothesis: Assume $\phi(m)$ for each $m \in \{1, \ldots, n\}$.

Induction Step: From the induction hypothesis, we know

$$\underset{(x_1, \ldots, x_n) \in A^n}{\exists !} \; (1, x_1), \ldots, (n, x_n) \in G.$$

Thus, if we let $x_{n+1} := f_n(x_1, \ldots, x_n)$, then $(n+1, x_{n+1}) \in G$ by the definition of G. If $(n+1, a) \in G$ with $a \neq x_{n+1}$, then $H := G \setminus \{(n+1, a)\} \in \mathcal{F}$ (using the uniqueness of

the $(1, x_1), \ldots, (n, x_n) \in G)$, implying $G \subseteq H$ in contradiction to $(n + 1, a) \notin H$. This shows $a = x_{n+1}$, proves $\phi(n + 1)$, and completes the induction.

Due to (3.12), we can now define $F : \mathbb{N} \longrightarrow A$, $F(n) := x_n$, and the definition of G then guarantees the validity of (3.9). ∎

Example 3.8. In many applications of Th. 3.7, one has functions $g_n : A \longrightarrow A$ and uses

$$\underset{n \in \mathbb{N}}{\forall} \left(f_n : A^n \longrightarrow A, \quad f_n(a_1, \ldots, a_n) := g_n(a_n) \right). \tag{3.13}$$

Here are some important concrete examples:

(a) The *factorial function* $F : \mathbb{N}_0 \longrightarrow \mathbb{N}$, $n \mapsto n!$, is defined recursively by

$$0! := 1, \quad 1! := 1, \quad \underset{n \in \mathbb{N}}{\forall} (n + 1)! := (n + 1) \cdot n!, \tag{3.14a}$$

i.e. we have $A = \mathbb{N}$ and $g_n(x) := (n + 1) \cdot x$. So we obtain

$$(n!)_{n \in \mathbb{N}_0} = (1, 1, 2, 6, 24, 120, \ldots). \tag{3.14b}$$

(b) For each $a \in \mathbb{R}$ and each $d \in \mathbb{R}$, we define the following *arithmetic progression* (also called *arithmetic sequence*) recursively by

$$a_1 := a, \quad \underset{n \in \mathbb{N}}{\forall} a_{n+1} := a_n + d, \tag{3.15a}$$

i.e. we have $A = \mathbb{R}$ and $g_n = g$ with $g(x) := x + d$. For example, for $a = 2$ and $d = -0.5$, we obtain

$$(a_n)_{n \in \mathbb{N}} = (2, \ 1.5, \ 1, \ 0.5, \ 0, \ -0.5, \ -1, \ -1.5, \ \ldots). \tag{3.15b}$$

(c) For each $a \in \mathbb{R}$ and each $q \in \mathbb{R} \setminus \{0\}$, we define the following *geometric progression* (also called *geometric sequence*) recursively by

$$x_1 := a, \quad \underset{n \in \mathbb{N}}{\forall} x_{n+1} := x_n \cdot q, \tag{3.16a}$$

i.e. we have $A = \mathbb{R}$ and $g_n = g$ with $g(x) := x \cdot q$. For example, for $a = 3$ and $q = -2$, we obtain

$$(x_n)_{n \in \mathbb{N}} = (3, \ -6, \ 12, \ -24, \ 48, \ \ldots). \tag{3.16b}$$

For the time being, we will continue to always specify A and the g_n or f_n in subsequent recursive definitions, but in the literature, most of the time, the g_n or f_n are not provided explicitly.

Example 3.9. (a) The *Fibonacci sequence* consists of the *Fibonacci numbers*, defined recursively by

$$F_0 := 0, \quad F_1 := 1, \quad \underset{n \in \mathbb{N}}{\forall} F_{n+1} := F_n + F_{n-1}, \tag{3.17a}$$

i.e. we have $A = \mathbb{N}_0$ and

$$f_n : A^n \longrightarrow A, \quad f_n(a_1, \ldots, a_n) := \begin{cases} 1 & \text{for } n = 1, \\ a_n + a_{n-1} & \text{for } n \geq 2. \end{cases} \tag{3.17b}$$

So we obtain

$$(F_n)_{n \in \mathbb{N}_0} = (0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots). \tag{3.17c}$$

(b) For $A := \mathbb{N}$, $x := 1$, and

$$f_n : A^n \longrightarrow A, \quad f_n(a_1, \ldots, a_n) := a_1 + \cdots + a_n, \tag{3.18a}$$

one obtains

$$x_1 = 1, \quad x_2 = f_1(1) = 1, \quad x_3 = f_2(1, 1) = 2, \quad x_4 = f_3(1, 1, 2) = 4,$$
$$x_5 = f_4(1, 1, 2, 4) = 8, \quad x_6 = f_5(1, 1, 2, 4, 8) = 16, \quad \ldots \tag{3.18b}$$

Definition 3.10. (a) *Summation Symbol:* On $A = \mathbb{R}$ (or, more generally, on every set where an addition $+ : A \times A \longrightarrow A$ is defined), define recursively, for each given (possibly finite) sequence $(a_1, a_2, \ldots)$ in A:

$$\sum_{i=1}^{1} a_i := a_1, \quad \sum_{i=1}^{n+1} a_i := a_{n+1} + \sum_{i=1}^{n} a_i \text{ for } n \geq 1, \tag{3.19a}$$

i.e.

$$f_n : A^n \longrightarrow A, \quad f_n(x_1, \ldots, x_n) := x_n + a_{n+1}. \tag{3.19b}$$

In (3.19a), one can also use other symbols for i, except a and n; for a finite sequence, n needs to be less than the maximal index of the finite sequence.

More generally, if I is an index set and $\phi : \{1, \ldots, n\} \longrightarrow I$ a bijective map, then define

$$\sum_{i \in I} a_i := \sum_{i=1}^{n} a_{\phi(i)}. \tag{3.19c}$$

The commutativity of addition implies that the definition in (3.19c) is actually independent of the chosen bijective map ϕ (cf. Th. B.2). Also define

$$\sum_{i \in \emptyset} a_i := 0 \tag{3.19d}$$

(for a general A, 0 is meant to be an element such that $a + 0 = 0 + a = a$ for each $a \in A$ and we can even define this if $0 \notin A$).

(b) *Product Symbol:* On $A = \mathbb{R}$ (or, more generally, on every set where a multiplication $\cdot : A \times A \longrightarrow A$ is defined), define recursively, for each given (possibly finite) sequence $(a_1, a_2, \ldots)$ in A:

$$\prod_{i=1}^{1} a_i := a_1, \quad \prod_{i=1}^{n+1} a_i := a_{n+1} \cdot \prod_{i=1}^{n} a_i \text{ for } n \geq 1, \tag{3.20a}$$

i.e.

$$f_n : A^n \longrightarrow A, \quad f_n(x_1, \ldots, x_n) := x_n \cdot a_{n+1}. \tag{3.20b}$$

In (3.20a), one can also use other symbols for i, except a and n; for a finite sequence, n needs to be less than the maximal index of the finite sequence.

More generally, if I is an index set and $\phi : \{1, \ldots, n\} \longrightarrow I$ a bijective map, then define

$$\prod_{i \in I} a_i := \prod_{i=1}^{n} a_{\phi(i)}. \tag{3.20c}$$

The commutativity of multiplication implies that the definition in (3.20c) is actually independent of the chosen bijective map ϕ (cf. Th. B.2). Also define

$$\prod_{i \in \emptyset} a_i := 1 \tag{3.20d}$$

(for a general A, 1 is meant to be an element such that $a \cdot 1 = 1 \cdot a = a$ for each $a \in A$ and we can even define this if $1 \notin A$).

Example 3.11. (a) Given $a, d \in \mathbb{R}$, let $(a_n)_{n \in \mathbb{N}}$ be the arithmetic sequence as defined in (3.15a). It is an exercise to prove by induction that

$$\underset{n \in \mathbb{N}}{\forall} \quad a_n = a + (n-1)d, \tag{3.21a}$$

$$\underset{n \in \mathbb{N}}{\forall} \quad S_n := \sum_{i=1}^{n} a_i = \frac{n}{2}(a_1 + a_n) = \frac{n}{2}\big(2a + (n-1)d\big), \tag{3.21b}$$

where the S_n are called *arithmetic sums*.

(b) Given $a \in \mathbb{R}$ and $q \in \mathbb{R} \setminus \{0\}$, let $(x_n)_{n \in \mathbb{N}}$ be the geometric sequence as defined in (3.16a). We will prove by induction that

$$\underset{n \in \mathbb{N}}{\forall} \quad x_n = a\, q^{n-1}, \tag{3.22a}$$

$$\underset{n \in \mathbb{N}}{\forall} \quad S_n := \sum_{i=1}^{n} x_i = \sum_{i=1}^{n}(a\, q^{i-1}) = a \sum_{i=0}^{n-1} q^i = \begin{cases} n\,a & \text{for } q = 1, \\ \frac{a\,(1-q^n)}{1-q} & \text{for } q \neq 1, \end{cases} \tag{3.22b}$$

where the S_n are called *geometric sums*.

For the induction proof of (3.22a), $\phi(n)$ is $x_n = a\, q^{n-1}$. The base case, $\phi(1)$, is the statement $x_1 = a\, q^0 = a$, which is true. For the induction step, we assume $\phi(n)$ and compute

$$x_{n+1} = x_n \cdot q \overset{\left(\phi(n)\right)}{=} a\, q^{n-1} \cdot q = a\, q^n, \tag{3.23}$$

showing $\phi(n) \Rightarrow \phi(n+1)$ and completing the proof.

For $q = 1$, the sum S_n is actually arithmetic with $d = 0$, i.e. $S_n = na$ can be obtained from (3.21b). For the induction proof of (3.22b) with $q \neq 1$, $\phi(n)$ is

$S_n = \frac{a(1-q^n)}{1-q}$. The base case, $\phi(1)$, is the statement $S_1 = \frac{a(1-q)}{1-q} = a$, which is true. For the induction step, we assume $\phi(n)$ and compute

$$S_{n+1} = S_n + x_{n+1} \overset{(\phi(n))}{=} \frac{a(1-q^n)}{1-q} + aq^n = \frac{a(1-q^n) + aq^n(1-q)}{1-q} = \frac{a(1-q^{n+1})}{1-q}, \tag{3.24}$$

showing $\phi(n) \Rightarrow \phi(n+1)$ and completing the proof.

3.2   Cardinality: The Size of Sets

Cardinality measures the size of sets. For a finite set A, it is precisely the number of elements in A. For an infinite set, it classifies the set's degree or level of infinity (it turns out that not all infinite sets have the same size).

Definition 3.12. (a) The sets A, B are defined to have the same *cardinality* or the same *size* if, and only if, there exists a bijective map $\varphi : A \longrightarrow B$. One can show that this defines an equivalence relation on every set of sets (see Th. A.53 of the Appendix).

(b) The *cardinality* of a set A is $n \in \mathbb{N}$ (denoted $\#A = n$) if, and only if, there exists a bijective map $\varphi : A \longrightarrow \{1, \ldots, n\}$. The cardinality of $\emptyset$ is defined as 0, i.e. $\#\emptyset := 0$. A set A is called *finite* if, and only if, there exists $n \in \mathbb{N}_0$ such that $\#A = n$; A is called *infinite* if, and only if, A is not finite, denoted $\#A = \infty$ (in the strict sense, this is an abuse of notation, since ∞ is *not* a cardinality – for example $\#\mathbb{N} = \infty$ and $\#\mathcal{P}(\mathbb{N}) = \infty$, but $\mathbb{N}$ and $\mathcal{P}(\mathbb{N})$ do *not* have the same cardinality, since the power set $\mathcal{P}(A)$ is always strictly bigger than A (see Th. A.69 of the Appendix) – $\#A = \infty$ is merely an abbreviation for the statement "A is infinite"). The interested student finds additional material regarding the uniqueness of finite cardinality in Th. A.61 and Cor. A.62, and regarding characterizations of infinite sets in Th. A.54 of the Appendix.

(c) The set A is called *countable* if, and only if, A is finite or A has the same cardinality as $\mathbb{N}$. Otherwise, A is called *uncountable*.

———

In the rest of the section, we present a number of important results regarding the natural numbers and countability.

Theorem 3.13. (a) *Every nonempty finite subset of a totally ordered set has a minimum and a maximum.*

(b) *Every nonempty subset of $\mathbb{N}$ has a minimum.*

Proof. (a): Let A be a set and let $\leq$ denote a total order on A. Moreover, let $\emptyset \neq B \subseteq A$. We show by induction

$$\underset{n \in \mathbb{N}}{\forall} \underbrace{\left(\#B = n \; \Rightarrow \; B \text{ has a min}\right)}_{\phi(n)}.$$

Base Case ($n = 1$): For $n = 1$, B contains a unique element b, i.e. $b = \min B$, proving $\phi(1)$.

Induction Step: Suppose $\phi(n)$ holds and consider B with $\#B = n + 1$. Let b be one element from B. Then $C := B \setminus \{b\}$ has cardinality n and, according to the induction hypothesis, there exists $c \in C$ satisfying $c = \min C$. If $c \leq b$, then $c \leq x$ for each $x \in B$, proving $c = \min B$. If $b \leq c$, then $b \leq x$ for each $x \in B$, proving $b = \min B$. In each case, B has a min, proving $\phi(n + 1)$ and completing the induction.

(b): Let $\emptyset \neq A \subseteq \mathbb{N}$. We have to show A has a min. If A is finite, then A has a min by (a). If A is infinite, let n be an element from A. Then the finite set $B := \{k \in A : k \leq n\}$ must have a min m by (a). Since $m \leq x$ for each $x \in B$ and $m \leq n < x$ for each $x \in A \setminus B$, we have $m = \min A$. ∎

Proposition 3.14. *Every subset A of $\mathbb{N}$ is countable.*

Proof. Since $\emptyset$ is countable, we may assume $A \neq \emptyset$. From Th. 3.13(b), we know that every nonempty subset of $\mathbb{N}$ has a min. We recursively define a sequence in A by

$$a_1 := \min A, \quad a_{n+1} := \begin{cases} \min A_n & \text{if } A_n := A \setminus \{a_i : 1 \leq i \leq n\} \neq \emptyset, \\ a_n & \text{if } A_n = \emptyset. \end{cases}$$

This sequence is the same as the function $f : \mathbb{N} \longrightarrow A$, $f(n) = a_n$. An easy induction shows that, for each $n \in \mathbb{N}$, $a_n \neq a_{n+1}$ implies the restriction $f \restriction_{\{1,\ldots,n+1\}}$ is injective. Thus, if there exists $n \in \mathbb{N}$ such that $a_n = a_{n+1}$, then $f \restriction_{\{1,\ldots,k\}}: \{1,\ldots,k\} \longrightarrow A$ is bijective, where $k := \min\{n \in \mathbb{N} : a_n = a_{n+1}\}$, showing A is finite, i.e. countable. If there does not exist $n \in \mathbb{N}$ with $a_n = a_{n+1}$, then f is injective. Another easy induction shows that, for each $n \in \mathbb{N}$, $f(\{1,\ldots,n\}) \supseteq \{k \in A : k \leq n\}$, showing f is also surjective, proving A is countable. ∎

Proposition 3.15. *For each set $A \neq \emptyset$, the following three statements are equivalent:*

(i) *A is countable.*

(ii) *There exists an injective map $f : A \longrightarrow \mathbb{N}$.*

(iii) *There exists a surjective map $g : \mathbb{N} \longrightarrow A$.*

Proof. Directly from the definition of countable in Def. 3.12(c), one obtains (i)⇒(ii) and (i)⇒(iii). To prove (ii)⇒(i), let $f : A \longrightarrow \mathbb{N}$ be injective. Then $f : A \longrightarrow f(A)$ is bijective, and, since $f(A) \subseteq \mathbb{N}$, $f(A)$ is countable by Prop. 3.14, proving A is countable as well. To prove (iii)⇒(i), let $g : \mathbb{N} \longrightarrow A$ be surjective. Then g has a right inverse $f : A \longrightarrow \mathbb{N}$. One can obtain this from Th. 2.13(a), but, here, we can actually construct f without the axiom of choice: For $a \in A$, let $f(a) := \min g^{-1}(\{a\})$ (recall Th. 3.13(b)). Then, clearly, $g \circ f = \text{Id}_A$. But this means g is a left inverse for f, showing f is injective according to Th. 2.13(b). Then A is countable by an application of (ii). ∎

Theorem 3.16. *If $(A_1, \ldots, A_n)$, $n \in \mathbb{N}$, is a finite family of countable sets, then $\prod_{i=1}^{n} A_i$ is countable.*

Proof. We first consider the special case $n = 2$ with $A_1 = A_2 = \mathbb{N}$ and show the map

$$\varphi : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}, \quad \varphi(m, n) := 2^m \cdot 3^n,$$

is injective: If $\varphi(m, n) = \varphi(p, q)$, then $2^m \cdot 3^n = 2^p \cdot 3^q$. Moreover $m \leq p$ or $p \leq m$. If $m \leq p$, then $3^n = 2^{p-m} \cdot 3^q$. Since 3^n is odd, $2^{p-m} \cdot 3^q$ must also be odd, implying $p - m = 0$, i.e. $m = p$. Moreover, we now have $3^n = 3^q$, implying $n = q$, showing $(m, n) = (p, q)$, i.e. φ is injective.

We now come back to the general case stated in the theorem. If at least one of the A_i is empty, then A is empty. So it remains to consider the case, where all A_i are nonempty. The proof is conducted by induction by showing

$$\underset{n \in \mathbb{N}}{\forall} \underbrace{\prod_{i=1}^{n} A_i \text{ is countable}}_{\phi(n)}.$$

Base Case ($n = 1$): $\phi(1)$ is merely the hypothesis that A_1 is countable.

Induction Step: Assuming $\phi(n)$, Prop. 3.15(ii) provides injective maps $f_1 : \prod_{i=1}^{n} A_i \longrightarrow \mathbb{N}$ and $f_2 : A_{n+1} \longrightarrow \mathbb{N}$. To prove $\phi(n+1)$, we provide an injective map $h : \prod_{i=1}^{n+1} A_i \longrightarrow \mathbb{N}$: Define

$$h : \prod_{i=1}^{n+1} A_i \longrightarrow \mathbb{N}, \quad h(a_1, \ldots, a_n, a_{n+1}) := \varphi\big(f_1(a_1, \ldots, a_n), f_2(a_{n+1})\big).$$

The injectivity of f_1, f_2, and φ clearly implies the injectivity of h, thereby proving $\phi(n + 1)$ and completing the induction. ∎

Theorem 3.17. *If $(A_i)_{i \in I}$ is a countable family of countable sets (i.e. $\emptyset \neq I$ is countable and each A_i, $i \in I$, is countable), then the union $A := \bigcup_{i \in I} A_i$ is also countable (this result makes use of AC, cf. Rem. 3.18 below).*

Proof. It suffices to consider the case that all A_i are nonempty. Moreover, according to Prop. 3.15(iii), it suffices to construct a surjective map $\varphi : \mathbb{N} \longrightarrow A$. Also according to Prop. 3.15(iii), the countability of I and the A_i provides us with surjective maps $f : \mathbb{N} \longrightarrow I$ and $g_i : \mathbb{N} \longrightarrow A_i$ (here AC is used to select each g_i from the set of all surjective maps from $\mathbb{N}$ onto A_i). Define

$$F : \mathbb{N} \times \mathbb{N} \longrightarrow A, \quad F(m, n) := g_{f(m)}(n).$$

Then F is surjective: Given $x \in A$, there exists $i \in I$ such that $x \in A_i$. Since f is surjective, there is $m \in \mathbb{N}$ satisfying $f(m) = i$. Moreover, since g_i is surjective, there exists $n \in \mathbb{N}$ with $g_i(n) = x$. Then $F(m, n) = g_i(n) = x$, verifying that F is surjective. As $\mathbb{N} \times \mathbb{N}$ is countable by Th. 3.16, there exists a surjective map $h : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$. Thus, $F \circ h$ is the desired surjective map from $\mathbb{N}$ onto A. ∎

Remark 3.18. The axiom of choice is, indeed, essential for the proof of Th. 3.17. It is shown in [Jec73, Th. 10.6] that it is consistent with the axioms of ZF (i.e. with the axioms of Sec. A.3 of the Appendix) that, e.g., the uncountable sets $\mathcal{P}(\mathbb{N})$ and $\mathbb{R}$ (cf. Th. F.2 of the Appendix) are countable unions of countable sets.

4 Real Numbers

4.1 The Real Numbers as a Complete Totally Ordered Field

The set of real numbers, denoted $\mathbb{R}$, is a set with special properties, namely a so-called *complete totally ordered field*, which, after some preliminaries, will be defined in Def. 4.3 below.

Definition 4.1. A total order $\leq$ on a nonempty set A is called *complete* if, and only if, every nonempty subset B of A that is bounded from above has a supremum, i.e.

$$\underset{B \in \mathcal{P}(A) \setminus \{\emptyset\}}{\forall} \left(\left(\underset{x \in A}{\exists} \ \underset{b \in B}{\forall} \ b \leq x \right) \ \Rightarrow \ \underset{s \in A}{\exists} \ s = \sup B \right). \tag{4.1}$$

Lemma 4.2. *A total order $\leq$ on a nonempty set A is complete if, and only if, every nonempty subset B of A that is bounded from below has an infimum.*

Proof. According to Lem. 2.29, it suffices to prove one implication. We show that (4.1) implies that every nonempty B bounded from below has an infimum: Define

$$C := \{x \in A : x \text{ is lower bound for } B\}. \tag{4.2}$$

Then every $b \in B$ is an upper bound for C and (4.1) implies there exists $s = \sup C \in A$. To verify $s = \inf B$, it remains to show $s \in C$, i.e. that s is a lower bound for B. However, every $b \in B$ is an upper bound for C and $s = \sup C$ is the min of all upper bounds for C, i.e. $s \leq b$ for each $b \in B$, showing $s \in C$. $\blacksquare$

Definition 4.3. Let $(A, +, \cdot)$ be a field and let $\leq$ be a total order on A. Then A (or, more precisely, $(A, +, \cdot, \leq)$) is called a *totally ordered field* if, and only if, the order is compatible with addition and multiplication, i.e. if, and only if,

$$\underset{x,y,z \in A}{\forall} \left(x \leq y \ \Rightarrow \ x + z \leq y + z \right), \tag{4.3a}$$

$$\underset{x,y \in A}{\forall} \left(0 \leq x \wedge 0 \leq y \ \Rightarrow \ 0 \leq xy \right). \tag{4.3b}$$

Finally, A is called a *complete totally ordered field* if, and only if, A is a totally ordered field that is complete in the sense of Def. 4.1.

Theorem 4.4. *There exists a complete totally ordered field $\mathbb{R}$ (it is called the set of real numbers). Moreover, $\mathbb{R}$ is unique up to isomorphism, i.e. if A is a complete totally ordered field, then there exists an isomorphism $\phi : A \longrightarrow \mathbb{R}$, i.e. a bijective map $\phi : A \longrightarrow \mathbb{R}$, satisfying*

$$\underset{x,y \in A}{\forall} \phi(x + y) = \phi(x) + \phi(y), \tag{4.4a}$$

$$\underset{x,y \in A}{\forall} \phi(xy) = \phi(x)\phi(y), \tag{4.4b}$$

$$\underset{x,y \in A}{\forall} \left(x < y \ \Rightarrow \ \phi(x) < \phi(y) \right). \tag{4.4c}$$

It also turns out that the isomorphism is unique.

Proof. To really prove the existence of the real numbers by providing a construction is tedious and not easy. One possible construction is provided in Appendix D (the existence proof is completed in Th. D.39, the results regarding the isomorphism can be found in Th. D.43). ∎

Theorem 4.5. *The following statements and rules are valid in the set of real numbers $\mathbb{R}$ (and, more generally, in every totally ordered field):*

(a) $x \leq y \Rightarrow -x \geq -y$.

(b) $x \leq y \wedge z \geq 0 \Rightarrow xz \leq yz$ *holds as well as* $x \leq y \wedge z \leq 0 \Rightarrow xz \geq yz$.

(c) $x \neq 0 \Rightarrow x^2 := x \cdot x > 0$. *In particular* $1 > 0$.

(d) $x > 0 \Rightarrow 1/x > 0$, *whereas* $x < 0 \Rightarrow 1/x < 0$.

(e) *If* $0 < x < y$, *then* $x/y < 1$, $y/x > 1$, *and* $1/x > 1/y$.

(f) $x < y \wedge u < v \Rightarrow x + u < y + v$.

(g) $0 < x < y \wedge 0 < u < v \Rightarrow xu < yv$.

(h) $x < y \wedge 0 < \lambda < 1 \Rightarrow x < \lambda x + (1 - \lambda)y < y$. *In particular* $x < \frac{x+y}{2} < y$.

Proof. (a): Using (4.3a): $x \leq y \Rightarrow 0 \leq y - x \Rightarrow -y \leq -x$.

(b): One argues, for $z \geq 0$,

$$x \leq y \Rightarrow 0 \leq y - x \overset{(4.3b)}{\Rightarrow} 0 \leq (y - x)z = yz - xz \Rightarrow xz \leq yz,$$

and, for $z \leq 0$,

$$x \leq y \Rightarrow 0 \leq y - x \overset{(4.3b)}{\Rightarrow} 0 \leq (y - x)(-z) = xz - yz \Rightarrow xz \geq yz.$$

(c): From (b), one obtains $x^2 \geq 0$. From Th. C.10(i), one then gets $x^2 > 0$.

(d): If $x > 0$, then $x^{-1} < 0$ implies the false statement $1 = xx^{-1} < 0$, i.e. $x^{-1} > 0$. The case $x < 0$ is treated analogously.

(e): Using (d), we obtain from $0 < x < y$ that $x/y = xy^{-1} < yy^{-1} = 1$ and $1 = xx^{-1} < yx^{-1} = y/x$.

(f): $x < y \Rightarrow x + u < y + u$ and $u < v \Rightarrow y + u < y + v$; both combined yield $x + u < y + v$.

(g): $0 < x < y \wedge 0 < u < v \Rightarrow xu < yu \wedge yu < yv \Rightarrow xu < yv$.

(h): Since $0 < \lambda$ and $1 - \lambda > 0$, $x < y$ implies

$$\lambda x < \lambda y \quad \wedge \quad (1 - \lambda)x < (1 - \lambda)y.$$

Using (4.3a), we obtain

$$x = \lambda x + (1 - \lambda)x < \lambda x + (1 - \lambda)y < \lambda y + (1 - \lambda)y = y,$$

completing the proof of the theorem. ∎

Theorem 4.6. *Let $\emptyset \neq A, B \subseteq \mathbb{R}$, $\lambda \in \mathbb{R}$, and define*

$$A + B := \{a + b : a \in A \wedge b \in B\}, \tag{4.5a}$$
$$\lambda A := \{\lambda a : a \in A\}. \tag{4.5b}$$

If A and B are bounded, then

$$\sup(A + B) = \sup A + \sup B, \tag{4.6a}$$
$$\inf(A + B) = \inf A + \inf B, \tag{4.6b}$$
$$\sup(\lambda A) = \begin{cases} \lambda \cdot \sup A & \text{for } \lambda \geq 0, \\ \lambda \cdot \inf A & \text{for } \lambda < 0, \end{cases} \tag{4.6c}$$
$$\inf(\lambda A) = \begin{cases} \lambda \cdot \inf A & \text{for } \lambda \geq 0, \\ \lambda \cdot \sup A & \text{for } \lambda < 0. \end{cases} \tag{4.6d}$$

Proof. Exercise. ∎

4.2 Important Subsets

Remark 4.7. We would like to recover the natural numbers $\mathbb{N}$ as a subset of $\mathbb{R}$. Indeed, if we start with 1 as the neutral element of multiplication and define $2 := 1+1$, $3 := 2+1$, ..., then $\mathbb{N} := \{1, 2, \dots\}$ is a subset of $\mathbb{R}$, satisfying the Peano axioms P1, P2, P3 of Sec. 3.1. However, if one does actually construct $\mathbb{R}$ according to the axioms of axiomatic set theory, then one starts by constructing $\mathbb{N}$ first, constructing $\mathbb{R}$ from $\mathbb{N}$ in several steps (cf. Appendix D). Depending on the construction used, the original set of natural numbers will typically not be the same set as the natural numbers as a subset of $\mathbb{R}$. However, both sets will satisfy the Peano axioms and you will have a canonical bijection between the two sets. Which one you consider the "genuine" set of natural numbers depends on your personal taste and philosophy and is completely irrelevant. Any two models of $\mathbb{N}$ will always produce equivalent results, since they must both satisfy the three Peano axioms.

—

We now introduce a zoo of important subsets of $\mathbb{R}$ together with corresponding notation:

$$\mathbb{N} := \{1, 2, 3, \dots\} \qquad \text{(natural numbers)}, \tag{4.7a}$$
$$\mathbb{N}_0 := \mathbb{N} \cup \{0\}, \tag{4.7b}$$
$$\mathbb{Z}^- := \{-n : n \in \mathbb{N}\} \qquad \text{(negative integers)}, \tag{4.7c}$$
$$\mathbb{Z} := \mathbb{Z}^- \cup \mathbb{N}_0 \qquad \text{(integers)}, \tag{4.7d}$$
$$\mathbb{Q}^+ := \{m/n : m, n \in \mathbb{N}\} \qquad \text{(positive rational numbers)}, \tag{4.7e}$$
$$\mathbb{Q}_0^+ := \mathbb{Q}^+ \cup \{0\} \qquad \text{(nonnegative rational numbers)}, \tag{4.7f}$$
$$\mathbb{Q}^- := \{-q : q \in \mathbb{Q}^+\} \qquad \text{(negative rational numbers)}, \tag{4.7g}$$

$$\mathbb{Q}_0^- := \mathbb{Q}^- \cup \{0\} \qquad \text{(nonpositive rational numbers),} \qquad \text{(4.7h)}$$
$$\mathbb{Q} := \mathbb{Q}_0^+ \cup \mathbb{Q}^- \qquad \text{(rational numbers),} \qquad \text{(4.7i)}$$
$$\mathbb{R}^+ := \{x \in \mathbb{R} : \ x > 0\} \qquad \text{(positive real numbers),} \qquad \text{(4.7j)}$$
$$\mathbb{R}_0^+ := \{x \in \mathbb{R} : \ x \geq 0\} \qquad \text{(nonnegative real numbers),} \qquad \text{(4.7k)}$$
$$\mathbb{R}^- := \{x \in \mathbb{R} : \ x < 0\} \qquad \text{(negative real numbers),} \qquad \text{(4.7l)}$$
$$\mathbb{R}_0^- := \{x \in \mathbb{R} : \ x \leq 0\} \qquad \text{(nonpositive real numbers).} \qquad \text{(4.7m)}$$

For $a, b \in \mathbb{R}$ with $a \leq b$, one also defines the following *intervals*:

$$[a, b] := \{x \in \mathbb{R} : \ a \leq x \leq b\} \qquad \text{(bounded closed interval),} \qquad \text{(4.8a)}$$
$$]a, b[:= \{x \in \mathbb{R} : \ a < x < b\} \qquad \text{(bounded open interval),} \qquad \text{(4.8b)}$$
$$]a, b] := \{x \in \mathbb{R} : \ a < x \leq b\} \qquad \text{(bounded half-open interval),} \qquad \text{(4.8c)}$$
$$[a, b[:= \{x \in \mathbb{R} : \ a \leq x < b\} \qquad \text{(bounded half-open interval),} \qquad \text{(4.8d)}$$
$$]-\infty, b] := \{x \in \mathbb{R} : \ x \leq b\} \qquad \text{(unbounded closed interval),} \qquad \text{(4.8e)}$$
$$]-\infty, b[:= \{x \in \mathbb{R} : \ x < b\} \qquad \text{(unbounded open interval),} \qquad \text{(4.8f)}$$
$$[a, \infty[:= \{x \in \mathbb{R} : \ a \leq x\} \qquad \text{(unbounded closed interval),} \qquad \text{(4.8g)}$$
$$]a, \infty[:= \{x \in \mathbb{R} : \ a < x\} \qquad \text{(unbounded open interval).} \qquad \text{(4.8h)}$$

For $a = b$, one says that the intervals defined by (4.8a) – (4.8d) are *degenerate* or *trivial*, where $[a, a] = \{a\}$, $]a, a[=]a, a] = [a, a[= \emptyset$ – it is sometimes convenient to have included the degenerate cases in the definition. It is sometimes also useful to abandon the restriction $a \leq b$, to let $c := \min\{a, b\}$, $d := \max\{a, b\}$, and to define

$$[a, b] := [c, d], \quad]a, b[:=]c, d[, \quad]a, b] :=]c, d], \quad [a, b[:= [c, d[. \qquad \text{(4.8i)}$$

Theorem 4.8 (Archimedean Property). *Let ϵ, x be real numbers. If $\epsilon > 0$ and $x > 0$, then there exists $n \in \mathbb{N}$ such that $n \epsilon > x$.*

Proof. We conduct the proof by contradiction: Suppose x is an upper bound for the set $A := \{n \epsilon : \ n \in \mathbb{N}\}$. Since the order $\leq$ on $\mathbb{R}$ is complete, according to (4.1), there exists $s \in \mathbb{R}$ such that $s = \sup A$. In particular, $s - \epsilon$ is not an upper bound for A, i.e. there exists $n \in \mathbb{N}$ satisfying $n \epsilon > s - \epsilon$. But then $(n + 1) \epsilon > s$ in contradiction to $s = \sup A$. This shows x is not an upper bound for A, thereby establishing the case. ∎

5   Complex Numbers

5.1   Definition and Basic Arithmetic

According to Th. 4.5(c), $x^2 \geq 0$ holds for every real number $x \in \mathbb{R}$, i.e. the equation $x^2 + 1 = 0$ has no solution in $\mathbb{R}$. This deficiency of the real numbers motivates the effort to try to extend the field of real numbers to a larger field $\mathbb{C}$, the so-called *complex numbers*. The two requirements that $\mathbb{C}$ is to be a field containing $\mathbb{R}$ and that there is to

be some complex number $i \in \mathbb{C}$ satisfying $i^2 = -1$ already dictates the following laws of addition and multiplication for complex numbers $z = x + iy$ and $w = u + iv$ with $x, y, u, v \in \mathbb{R}$:

$$z + w = x + iy + u + iv = x + u + i(y + v), \tag{5.1a}$$
$$zw = (x + iy)(u + iv) = xu - yv + i(xv + yu). \tag{5.1b}$$

Moreover, if $x + iy = u + iv$, then $(x - u)^2 = -(v - y)^2$, i.e. $x - u = 0 = v - y$, implying $x = u$ and $y = v$. This suggests to try defining complex numbers as pairs of real numbers. Indeed, this works:

Definition 5.1. We define the set of *complex numbers* $\mathbb{C} := \mathbb{R} \times \mathbb{R}$, where, keeping in mind (5.1), *addition* on $\mathbb{C}$ is defined by

$$+ : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}, \quad \big((x, y), (u, v)\big) \mapsto (x, y) + (u, v) := (x + u, y + v), \tag{5.2}$$

and *multiplication* on $\mathbb{C}$ is defined by

$$\cdot : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}, \quad \big((x, y), (u, v)\big) \mapsto (x, y) \cdot (u, v) := (xu - yv, xv + yu). \tag{5.3}$$

Theorem 5.2. (a) *The set of complex numbers $\mathbb{C}$ with addition and multiplication as defined in Def. 5.1 forms a field, where $(0, 0)$ and $(1, 0)$ are the neutral elements with respect to addition and multiplication, respectively,*

$$-z := (-x, -y) \tag{5.4a}$$

is the additive inverse to $z = (x, y)$, whereas

$$z^{-1} := \frac{1}{z} := \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) \tag{5.4b}$$

is the multiplicative inverse to $z = (x, y) \neq (0, 0)$.

(b) *Defining subtraction and division in the usual way, for each $z, w \in \mathbb{C}$, by $w - z := w + (-z)$, and $w/z := wz^{-1}$ for $z \neq (0, 0)$, respectively, all the rules stated in Th. C.10 are valid in $\mathbb{C}$.*

(c) *The map*

$$\iota : \mathbb{R} \longrightarrow \mathbb{C}, \quad \iota(x) := (x, 0), \tag{5.5}$$

is a monomorphism, i.e. it is injective and satisfies

$$\underset{x,y \in \mathbb{R}}{\forall} \quad \iota(x + y) = \iota(x) + \iota(y), \tag{5.6a}$$

$$\underset{x,y \in \mathbb{R}}{\forall} \quad \iota(xy) = \iota(x) \cdot \iota(y). \tag{5.6b}$$

It is customary to identify $\mathbb{R}$ with $\iota(\mathbb{R})$, as it usually does not cause any confusion. One then just writes x instead of $(x, 0)$.

Proof. All computations required for (a) and (c) are straightforward and are left as an exercise; (b) is a consequence of (a), since Th. C.10 is valid for every field. ∎

Notation 5.3. The number $i := (0,1)$ is called the *imaginary unit* (note that, indeed, $i^2 = i \cdot i = (0,1) \cdot (0,1) = (0 \cdot 0 - 1 \cdot 1, \, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -1$). Using i, one obtains the commonly used representation of a complex number $z = (x,y) \in \mathbb{C}$:

$$z = (x,y) = x \cdot (1,0) + y \cdot (0,1) = x + iy, \tag{5.7}$$

where one calls $\operatorname{Re} z := x$ the *real part* of z and $\operatorname{Im} z := y$ the *imaginary part* of z. Moreover, z is called *purely imaginary* if, and only if, $\operatorname{Re} z = 0$.

Remark 5.4. There does not exist a total order $\leq$ on $\mathbb{C}$ that makes $\mathbb{C}$ into a totally ordered field (i.e. no total order on $\mathbb{C}$ can be compatible with addition and multiplication in the sense of (4.3)): Indeed, if there were such a total order $\leq$ on $\mathbb{C}$, then all the rules of Th. 4.5 had to be valid with respect to that total order $\leq$. In particular, $0 < 1^2 = 1$ and $0 < i^2 = -1$ had to be valid by Th. 4.5(c), and, then, $0 < 1 + (-1) = 0$ had to be valid by Th. 4.5(f). However, $0 < 0$ is false, showing that there is no total order on $\mathbb{C}$ that satisfies (4.3). Caveat: Of course, there *do* exist total orders on $\mathbb{C}$, just none compatible with addition and multiplication – for example, the lexicographic order on $\mathbb{R} \times \mathbb{R}$ (defined as it was in (2.52) for $\mathbb{N} \times \mathbb{N}$) constitutes a total order on $\mathbb{C}$.

Definition and Remark 5.5. Conjugation: For each complex number $z = x + iy$, we define its *complex conjugate* or just *conjugate* to be the complex number $\bar{z} := x - iy$. We then have the following rules that hold for each $z = x + iy, w = u + iv \in \mathbb{C}$:

(a) $\overline{z + w} = x + u - iy - iv = \bar{z} + \bar{w}$ and $\overline{zw} = xu - yv - (xv + yu)i = (x - iy)(u - iv) = \bar{z}\bar{w}$.

(b) $z + \bar{z} = 2x = 2\operatorname{Re} z$ and $z - \bar{z} = 2yi = 2i\operatorname{Im} z$.

(c) $z = \bar{z} \iff x + iy = x - iy \iff y = 0 \iff z \in \mathbb{R}$.

(d) $z\bar{z} = (x + iy)(x - iy) = x^2 + y^2 \in \mathbb{R}_0^+$.

5.2   Sign and Absolute Value (Modulus)

We face a certain conundrum regarding the handling of square roots. The problem is that we will need the notion of a continuous function to prove the existence of a unique square root $\sqrt{x}$ for every nonnegative real number x and, in consequence, we will have to wait until Section 7.2.5 below to carry out this proof. On the other hand, it is extremely desirable to present the theory of convergence simultaneously for real and for complex numbers, which requires the notion of the *absolute value* or *modulus* of a complex number, to be defined in Def. 5.7(b) below as the square root of a nonnegative real number.

Faced with this difficulty, we will introduce the notion of square root now, *assuming* the existence, until we can add the proof in Section 7.2.5. Some students might be worried

that this might lead to a circular argument, where our later proof of the existence of square roots would somehow make use of our previous assumption of that existence. Of course, we will be careful not to make such a circular (and, thereby, logically invalid) argument. The point is that for *real numbers* the notion of absolute value does in no way depend on the notion of a square root (see Lem. 5.8 below).

Definition and Remark 5.6. We define a nonnegative real number $y \in \mathbb{R}_0^+$ to be the *square root* of the nonnegative real number $x \in \mathbb{R}_0^+$ if, and only if, $y^2 = x$. If y is the square root of x, then one uses the notation $\sqrt{x} := y$. We will see in Rem. and Def. 7.61 that every $x \in \mathbb{R}_0^+$ has a unique square root and that the function $f : \mathbb{R}_0^+ \longrightarrow \mathbb{R}_0^+$, $f(x) := \sqrt{x}$, is strictly increasing (in particular, injective).

Definition 5.7. (a) The *sign* function is defined by

$$\mathrm{sgn} : \mathbb{R} \longrightarrow \mathbb{R}, \quad \mathrm{sgn}(x) := \begin{cases} 1 & \text{for } x > 0, \\ 0 & \text{for } x = 0, \\ -1 & \text{for } x < 0. \end{cases} \tag{5.8}$$

It is emphasized that the sign function is only defined for *real* numbers (cf. Rem. 5.4)!

(b) The *absolute value* or *modulus* function is defined by

$$\mathrm{abs} : \mathbb{C} \longrightarrow \mathbb{R}_0^+, \quad z = x + iy \mapsto |z| := \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}, \tag{5.9}$$

where the term *absolute value* is often preferred for real numbers $z \in \mathbb{R}$ and the term *modulus* is often preferred if one also considers complex numbers $z \notin \mathbb{R}$.

Lemma 5.8. *For each $x \in \mathbb{R}$, one has*

$$|x| = x \cdot \mathrm{sgn}(x) = \begin{cases} x & \text{for } x \geq 0, \\ -x & \text{for } x < 0. \end{cases} \tag{5.10}$$

Proof. One has

$$|x| = \sqrt{x^2} = \begin{cases} x & \text{for } x \geq 0, \\ -x & \text{for } x < 0, \end{cases} \tag{5.11}$$

as claimed. ∎

Theorem 5.9. *The following rules hold for each $z, w \in \mathbb{C}$:*

(a) $z \neq 0 \Rightarrow |z| > 0$.

(b) $||z|| = |z|$.

(c) $|z| = |\bar{z}|$.

(d) $\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leq |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|$.

(e) $|zw| = |z||w|$.

(f) *For $w \neq 0$, one has $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$.*

(g) Triangle Inequality:
$$|z + w| \leq |z| + |w|. \tag{5.12}$$

(h) Inverse Triangle Inequality:
$$\big||z| - |w|\big| \leq |z - w|. \tag{5.13}$$

Proof. We carry out the proofs for $z, w \in \mathbb{C}$. However, for $z, w \in \mathbb{R}$, everything can easily be shown directly from (5.10), without making use of square roots.

Let $z = x + iy$ with $x, y \in \mathbb{R}$.

(a): If $z \neq 0$, then $x \neq 0$ or $y \neq 0$, i.e. $x^2 > 0$ or $y^2 > 0$ by Th. 4.5(c), implying $x^2 + y^2 > 0$ by Th. 4.5(f), i.e. $|z| = \sqrt{x^2 + y^2} > 0$.

(b): Since $a := |z| \in \mathbb{R}_0^+$, we have $|a| = \sqrt{a^2} = a = |z|$.

(c): Since $\bar{z} = x - iy$, we have $|\bar{z}| = \sqrt{x^2 + (-y)^2} = \sqrt{x^2 + y^2} = |z|$.

(d): It is $x = \operatorname{Re} z$, $y = \operatorname{Im} z$. Let $a := \max\{|x|, |y|\}$. As remarked in Def. and Rem. 5.6, the square root function is increasing and, thus, taking square roots in the chain of inequalities $a^2 \leq x^2 + y^2 \leq (|x| + |y|)^2$ implies $a \leq |z| \leq |x| + |y|$ as claimed.

(e): As remarked in Def. and Rem. 5.6, the square root function is injective, and, thus, (e) follows from
$$|zw|^2 = zw\,\overline{zw} \overset{\text{Def. and Rem. 5.5(a)}}{=} zw\bar{z}\bar{w} = z\bar{z}\,w\bar{w} = |z|^2\,|w|^2.$$

(f): Let $w = u + iv$ with $u, v \in \mathbb{R}$. We first consider the special case $z = 1$. Applying the formula (5.4b) for the inverse to w, one obtains
$$|w^{-1}|^2 = \frac{u^2}{(u^2 + v^2)^2} + \frac{v^2}{(u^2 + v^2)^2} = \frac{1}{u^2 + v^2} = \left(|w|^{-1}\right)^2,$$

i.e. $|w^{-1}| = |w|^{-1}$. Now (f) follows from (e): $\left|\frac{z}{w}\right| = |zw^{-1}| = |z||w^{-1}| = |z||w|^{-1} = \frac{|z|}{|w|}$.

(g) follows from
$$
\begin{aligned}
|z + w|^2 \quad &= \quad (z + w)(\bar{z} + \bar{w}) = z\bar{z} + w\bar{z} + z\bar{w} + w\bar{w} \\
&\overset{\text{Def. and Rem. 5.5(b)}}{=} \quad |z|^2 + 2\operatorname{Re}(z\bar{w}) + |w|^2 \\
&\overset{\text{(d)}}{\leq} \quad |z|^2 + 2|z\bar{w}| + |w|^2 = \left(|z| + |w|\right)^2,
\end{aligned}
$$

once again using that the square root function is increasing.

(h): Using (g), we obtain
$$
\begin{aligned}
|z| = |z - w + w| \leq |z - w| + |w| \quad &\Rightarrow \quad |z| - |w| \leq |z - w|, \\
|w| = |w - z + z| \leq |z - w| + |z| \quad &\Rightarrow \quad -(|z| - |w|) \leq |z - w|,
\end{aligned}
$$

implying $\big||z| - |w|\big| \leq |z - w|$ by (5.10) (notice $|z| - |w| \in \mathbb{R}$). ■

Remark 5.10. Each complex number $(x, y) = x + iy$ can be visualized as a point in the so-called *complex plane*, where the horizontal x-axis represents real numbers and the veritcal y-axis represents purely imaginary numbers. Then the addition of complex numbers is precisely the vector addition of 2-dimensional vectors in the complex plane, and conjugation is represented by reflection through the x-axis. Moreover, the modulus $|z|$ of a complex number is precisely its distance from the origin $(0, 0)$, and $|z - w|$ is the distance between the points $z = (x, y)$ and $w = (u, v)$ in the plane. Complex multiplication can also be interpreted geometrically in the plane: If ϕ denotes the angle that the vector representing $z = (x, y)$ forms with the x-axis, and, likewise, ψ denotes the angle that the vector representing $w = (u, v)$ forms with the x-axis, then zw is the vector of length $|zw|$ that forms the angle $\phi + \psi$ with the x-axis (we will better understand this geometrical interpretation of complex multiplication later (see Def. and Rem. 8.29), when writing complex numbers in the polar form $z = x + iy = |z| \exp(i\phi)$, making use of the exponential function exp).

5.3   Sums and Products

Here we compile some important rules involving sums and products. We are mostly interested in applying them to real and complex numbers. However, most of the rules, without extra difficulty, can be proved to hold in more general structures. We will provide the more general statements, but the reader will not lose much by merely thinking of $\mathbb{C}$ rather than a general ring in Th. 5.11, and of $\mathbb{R}$ rather than a general totally ordered field in Th. 5.12. Some rules involve exponentiation as defined in (C.10a) of the Appendix (in particular, it is used that $z^0 = 1$ for each $z \in R$, where R is ring with unity), and some proofs make use of the corresponding exponentiation rules of Th. C.6.

Theorem 5.11. *Let $(R, +, \cdot)$ be a ring (cf. Def. C.7 of the Appendix).*

(a) *For each $n \in \mathbb{N}$ and each $\lambda, \mu, z_j, w_j \in R$, $j \in \{1, \dots, n\}$:*

$$\sum_{j=1}^{n} (\lambda z_j + \mu w_j) = \lambda \sum_{j=1}^{n} z_j + \mu \sum_{j=1}^{n} w_j.$$

(b) *If R is a ring with unity, then, for each $n \in \mathbb{N}_0$ and each $z \in R$:*

$$(1 - z)(1 + z + z^2 + \cdots + z^n) = (1 - z) \sum_{j=0}^{n} z^j = 1 - z^{n+1}.$$

(c) *If R is a commutative ring with unity, then, for each $n \in \mathbb{N}_0$ and each $z, w \in R$:*

$$w^{n+1} - z^{n+1} = (w - z) \sum_{j=0}^{n} z^j w^{n-j} = (w - z)(w^n + zw^{n-1} + \cdots + z^{n-1}w + z^n).$$

Proof. In each case, the proof can be conducted by an easy induction. We carry out (c) and leave the other cases as exercises. For (c), the base case $(n = 0)$ is provided by the true statement $w^{0+1} - z^{0+1} = w - z = (w - z)z^0 \, w^{0-0}$. For the induction step, one computes

$$
(w - z) \sum_{j=0}^{n+1} z^j \, w^{n+1-j} \quad = \quad (w - z) \left(z^{n+1} w^0 + \sum_{j=0}^{n} z^j \, w^{n+1-j} \right)
$$

$$
= \quad (w - z)z^{n+1} + (w - z)\, w \sum_{j=0}^{n} z^j \, w^{n-j}
$$

$$
\overset{\text{ind. hyp.}}{=} \quad (w - z)\, z^{n+1} + w(w^{n+1} - z^{n+1}) = w^{n+2} - z^{n+2},
$$

completing the induction. ∎

Theorem 5.12. *Let* $(F, +, \cdot)$ *be a totally ordered field (cf. Def. 4.3).*

(a) *For each* $n \in \mathbb{N}$ *and each* $x_j, y_j \in F$, $j \in \{1, \ldots, n\}$:

$$
\left(\underset{j \in \{1,\ldots,n\}}{\forall} x_j \leq y_j \right) \quad \Rightarrow \quad \sum_{j=1}^{n} x_j \leq \sum_{j=1}^{n} y_j,
$$

where equality can only hold if $x_j = y_j$ *for each* $j \in \{1, \ldots, n\}$.

(b) *For each* $n \in \mathbb{N}$ *and each* $x_j, y_j \in F$, $j \in \{1, \ldots, n\}$:

$$
\left(\underset{j \in \{1,\ldots,n\}}{\forall} 0 < x_j \leq y_j \right) \quad \Rightarrow \quad \prod_{j=1}^{n} x_j \leq \prod_{j=1}^{n} y_j,
$$

where equality can only hold if $x_j = y_j$ *for each* $j \in \{1, \ldots, n\}$.

Proof. Both cases are proved by simple inductions, where Th. 4.5(f) is used in (a) and Th. 4.5(g) is used in (b). ∎

Theorem 5.13. Triangle Inequality: *For each* $n \in \mathbb{N}$ *and each* $z_j \in \mathbb{C}$, $j \in \{1, \ldots, n\}$:

$$
\left| \sum_{j=1}^{n} z_j \right| \leq \sum_{j=1}^{n} |z_j|.
$$

Proof. Another very simple induction that is left to the reader. ∎

5.4 Binomial Coefficients and Binomial Theorem

The goal in this section is to expand $(z + w)^n$ into a sum. This sum involves the so-called *binomial coefficients* $\binom{n}{k}$, which are also useful in other contexts. To obtain an idea for

what to expect, let us compute the cases $n = 0, 1, 2, 3$: $(z + w)^0 = 1$, $(z + w)^1 = z + w$, $(z + w)^2 = z^2 + 2zw + w^2$, $(z + w)^3 = z^3 + 3z^2w + 3zw^2 + w^3$. One finds that the coefficients form what is known as *Pascal's triangle*, which we write for $n = 0, \ldots, 5$:

$$
\begin{array}{lccccccccccc}
n = 0: & & & & & & 1 & & & & & \\
n = 1: & & & & & 1 & & 1 & & & & \\
n = 2: & & & & 1 & & 2 & & 1 & & & \\
n = 3: & & & 1 & & 3 & & 3 & & 1 & & \\
n = 4: & & 1 & & 4 & & 6 & & 4 & & 1 & \\
n = 5: & 1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
\tag{5.14}
$$

The entries of the nth row of Pascal's triangle are denoted by $\binom{n}{0}, \ldots, \binom{n}{n}$. One also observes that one obtains each entry of the $(n+1)$st row, except the first and last entry, by adding the corresponding entries in row n to the left and to the right of the considered entry in row $n + 1$. The first and last entry of each row are always set to 1. This can be summarized as

$$
\underset{n \in \mathbb{N}_0}{\forall} \left(\binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k} \text{ for } k \in \{1, \ldots, n\} \right). \tag{5.15}
$$

The following Def. 5.14 provides a different and more general definition of binomial coefficients. We will then prove in Prop. 5.15 that the binomial coefficients as defined in Def. 5.14 do, indeed, satisfy (5.15).

Definition 5.14. For each $\alpha \in \mathbb{C}$ and each $k \in \mathbb{N}_0$, we define the *binomial coefficient*

$$
\binom{\alpha}{0} := 1, \quad \binom{\alpha}{k} := \prod_{j=1}^{k} \frac{\alpha + 1 - j}{j} = \frac{\alpha(\alpha - 1) \cdots (\alpha - k + 1)}{1 \cdot 2 \cdots k} \quad \text{for } k \in \mathbb{N}. \tag{5.16}
$$

Proposition 5.15. (a) *For each $\alpha \in \mathbb{C}$ and each $k \in \mathbb{N}$:*

$$
\binom{\alpha}{0} = 1, \quad \binom{\alpha + 1}{k} = \binom{\alpha}{k-1} + \binom{\alpha}{k}. \tag{5.17}
$$

(b) *For each $n \in \mathbb{N}_0$:*

$$
\binom{n}{n} = 1. \tag{5.18}
$$

The above statements include (5.15) as a special case.

Proof. (a): The first identity is part of the definition in (5.16). For the second identity, we first observe, for each $k \in \mathbb{N}$,

$$
\binom{\alpha}{k} = \prod_{j=1}^{k} \frac{\alpha + 1 - j}{j} = \frac{\alpha + 1 - k}{k} \prod_{j=1}^{k-1} \frac{\alpha + 1 - j}{j} = \binom{\alpha}{k-1} \frac{\alpha + 1 - k}{k}, \tag{5.19}
$$

which implies

$$\binom{\alpha}{k-1} + \binom{\alpha}{k} = \binom{\alpha}{k-1}\left(1 + \frac{\alpha+1-k}{k}\right) = \binom{\alpha}{k-1}\frac{\alpha+1}{k}$$

$$= \frac{\alpha+1}{k}\prod_{j=1}^{k-1}\frac{\alpha+1-j}{j} = \prod_{j=1}^{k}\frac{\alpha+2-j}{j} = \binom{\alpha+1}{k}. \tag{5.20}$$

(b): $\binom{0}{0} = 1$ according to (5.16). For $n \in \mathbb{N}$, (5.18) is proved by induction. The base case $(n = 1)$ is provided by the true statement $\binom{1}{1} = \frac{1+1-1}{1} = 1$. For the induction step, one computes

$$\binom{n+1}{n+1} = \prod_{j=1}^{n+1}\frac{n+1+1-j}{j} = \frac{n+1}{n+1}\prod_{j=1}^{n}\frac{n+1-j}{j} = \binom{n}{n} \stackrel{\text{ind. hyp.}}{=} 1, \tag{5.21}$$

which completes the induction. ∎

Theorem 5.16 (Binomial Theorem). *Let R be a commutative ring with unity (cf. Def. C.7 of the Appendix – for us, $R = \mathbb{C}$ is the most important example). For each $z, w \in R$ and each $n \in \mathbb{N}_0$, the following formula holds:*

$$(z+w)^n = \sum_{k=0}^{n}\binom{n}{k}z^{n-k}w^k = z^n + \binom{n}{1}z^{n-1}w + \cdots + \binom{n}{n-1}zw^{n-1} + w^n. \tag{5.22}$$

Proof. The proof is conducted via induction on n. The base case $(n = 0)$ is provided by the correct statement $(z+w)^0 = 1 = \binom{0}{0}z^{0-0}w^0$. For the induction step, we first observe

$$(z+w)^{n+1} = (z+w)(z+w)^n = z(z+w)^n + w(z+w)^n. \tag{5.23}$$

Using the induction hypothesis, we now further manipulate the two terms on the right-hand side of (5.23):

$$z(z+w)^n \stackrel{\text{ind. hyp.}}{=} z\sum_{k=0}^{n}\binom{n}{k}z^{n-k}w^k = \sum_{k=0}^{n}\binom{n}{k}z^{n+1-k}w^k$$

$$\stackrel{\binom{n}{n+1}=0}{=} \sum_{k=0}^{n+1}\binom{n}{k}z^{n+1-k}w^k, \tag{5.24}$$

$$w(z+w)^n \stackrel{\text{ind. hyp.}}{=} w\sum_{k=0}^{n}\binom{n}{k}z^{n-k}w^k = \sum_{k=0}^{n}\binom{n}{k}z^{n-k}w^{k+1}$$

$$= \sum_{k=1}^{n+1}\binom{n}{k-1}z^{n+1-k}w^k. \tag{5.25}$$

Plugging (5.24) and (5.25) into (5.23) yields

$$
\begin{aligned}
(z+w)^{n+1} \quad &= \quad \binom{n}{0} z^{n+1} w^0 + \sum_{k=1}^{n+1} \left(\binom{n}{k} + \binom{n}{k-1} \right) z^{n+1-k} w^k \\
&\overset{\text{Prop. 5.15}}{=} \binom{n+1}{0} z^{n+1} w^0 + \sum_{k=1}^{n+1} \binom{n+1}{k} z^{n+1-k} w^k \\
&= \quad \sum_{k=0}^{n+1} \binom{n+1}{k} z^{n+1-k} w^k, \qquad\qquad (5.26)
\end{aligned}
$$

completing the induction. ∎

The binomial theorem can now be used to infer a few more rules that hold for the binomial coefficients:

Corollary 5.17. *One has the following identities:*

$$
\forall_{n \in \mathbb{N}_0} \qquad \sum_{k=0}^{n} \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n, \qquad\qquad (5.27\text{a})
$$

$$
\forall_{n \in \mathbb{N}} \qquad \sum_{k=0}^{n} \binom{n}{k} (-1)^k = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - + \cdots + (-1)^n \binom{n}{n} = 0. \qquad (5.27\text{b})
$$

Proof. (5.27a) is just (5.22) with $z = w = 1$; (5.27b) is just (5.22) with $z = 1$ and $w = -1$. ∎

The formulas provided by the following proposition are also sometimes useful.

Proposition 5.18. (a) *For each $\alpha \in \mathbb{C}$ and each $k \in \mathbb{N}_0$:*

$$
\sum_{j=0}^{k} \binom{\alpha+j}{j} = \binom{\alpha}{0} + \binom{\alpha+1}{1} + \cdots + \binom{\alpha+k}{k} = \binom{\alpha+k+1}{k}. \qquad\qquad (5.28)
$$

(b) *For each $n, k \in \mathbb{N}_0$ with $k \leq n$:*

$$
\binom{n}{k} = \frac{n!}{k!(n-k)!}. \qquad\qquad (5.29)
$$

Moreover, for $n \geq 1$, one has $\binom{n}{k} = \#\mathcal{P}_k(\{1, \ldots, n\})$, where

$$
\mathcal{P}_k(A) := \big\{ B \in \mathcal{P}(A) : \#B = k \big\} \qquad\qquad (5.30)
$$

denotes the set of all subsets of a set A that have precisely k elements.

(c) *For each $n, k \in \mathbb{N}_0$:*

$$
\sum_{j=0}^{k} \binom{n+j}{n} = \binom{n}{n} + \binom{n+1}{n} + \cdots + \binom{n+k}{n} = \binom{n+k+1}{n+1}. \qquad\qquad (5.31)
$$

Proof. The induction proofs of (a) and (b) are left as exercises. For (c), one computes

$$\sum_{j=0}^{k}\binom{n+j}{n} \overset{(5.29)}{=} \sum_{j=0}^{k}\frac{(n+j)!}{n!(n+j-n)!} \overset{(5.29)}{=} \sum_{j=0}^{k}\binom{n+j}{j}$$

$$\overset{(5.28)}{=} \binom{n+k+1}{k} \overset{(5.29)}{=} \frac{(n+k+1)!}{k!(n+1)!} = \binom{n+k+1}{n+1},$$

thereby establishing the case. ∎

6   Polynomials

6.1   Arithmetic of $\mathbb{K}$-Valued Functions

Notation 6.1. We will write $\mathbb{K}$ in situations, where we allow $\mathbb{K}$ to be $\mathbb{R}$ or $\mathbb{C}$.

Notation 6.2. If A is any nonempty set, then one can add and multiply arbitrary functions $f, g : A \longrightarrow \mathbb{K}$, and one can define several further operations to create new functions from f and g:

$$\begin{aligned}
(f + g) : A \longrightarrow \mathbb{K}, &\qquad (f + g)(x) := f(x) + g(x), &\qquad \text{(6.1a)}\\
(\lambda f) : A \longrightarrow \mathbb{K}, &\qquad (\lambda f)(x) := \lambda f(x) \quad \text{for each } \lambda \in \mathbb{K}, &\qquad \text{(6.1b)}\\
(fg) : A \longrightarrow \mathbb{K}, &\qquad (fg)(x) := f(x)g(x), &\qquad \text{(6.1c)}\\
(f/g) : A \longrightarrow \mathbb{K}, &\qquad (f/g)(x) := f(x)/g(x) \quad \text{(assuming } g(x) \neq 0), &\qquad \text{(6.1d)}\\
\operatorname{Re} f : A \longrightarrow \mathbb{R}, &\qquad (\operatorname{Re} f)(x) := \operatorname{Re}(f(x)), &\qquad \text{(6.1e)}\\
\operatorname{Im} f : A \longrightarrow \mathbb{R}, &\qquad (\operatorname{Im} f)(x) := \operatorname{Im}(f(x)). &\qquad \text{(6.1f)}
\end{aligned}$$

For $\mathbb{K} = \mathbb{R}$, we further define

$$\begin{aligned}
\max(f, g) : A \longrightarrow \mathbb{R}, &\qquad \max(f, g)(x) := \max\big\{f(x), g(x)\big\}, &\qquad \text{(6.1g)}\\
\min(f, g) : A \longrightarrow \mathbb{R}, &\qquad \min(f, g)(x) := \min\big\{f(x), g(x)\big\}, &\qquad \text{(6.1h)}\\
f^{+} : A \longrightarrow \mathbb{R}, &\qquad f^{+} := \max(f, 0), &\qquad \text{(6.1i)}\\
f^{-} : A \longrightarrow \mathbb{R}, &\qquad f^{-} := \max(-f, 0). &\qquad \text{(6.1j)}
\end{aligned}$$

Finally, once again also allowing $\mathbb{K} = \mathbb{C}$,

$$|f| : A \longrightarrow \mathbb{R}, \qquad |f|(x) := |f(x)|. \qquad \text{(6.1k)}$$

One calls f^{+} and f^{-} the *positive part* and the *negative part* of f, respectively. For $\mathbb{R}$-valued functions f, we have

$$|f| = f^{+} + f^{-}. \qquad \text{(6.1l)}$$

6.2   Polynomials

Definition 6.3. Let $n \in \mathbb{N}$. Each function from $\mathbb{K}$ into $\mathbb{K}$, $x \mapsto x^n$, is called a *monomial*. A function P from $\mathbb{K}$ into $\mathbb{K}$ is called a *polynomial* if, and only if, it is a linear combination of monomials, i.e. if, and only if, P has the form

$$P : \mathbb{K} \longrightarrow \mathbb{K}, \quad P(x) = \sum_{j=0}^{n} a_j x^j = a_0 + a_1 x + \cdots + a_n x^n, \quad a_j \in \mathbb{K}. \qquad (6.2)$$

The a_j are called the *coefficients* of P. The largest number $d \leq n$ such that $a_d \neq 0$ is called the *degree* of P, denoted $\deg(P)$. If all coefficients are 0, then P is called the *zero polynomial*; the degree of the zero polynomial is defined as -1 (in Th. 6.6(b) below, we will see that each polynomial of degree $n \in \mathbb{N}_0$ is uniquely determined by its coefficients $a_0, \ldots, a_n$ and vice versa).

Polynomials of degree ≤ 0 are *constant*. Polynomials of degree ≤ 1 have the form $P(x) = a + bx$ and are called *affine* functions (often they are also called *linear* functions, even though this is not really correct for $a \neq 0$, since every function P that is linear (in the sense of linear algebra) must satisfy $P(0) = 0$). Polynomials of degree ≤ 2 have the form $P(x) = a + bx + cx^2$ and are called *quadratic* functions.

Each $\xi \in \mathbb{K}$ such that $P(\xi) = 0$ is called a *zero* or a *root* of P.

A *rational function* is a quotient P/Q of two polynomials P and Q.

Remark 6.4. Let $\lambda \in \mathbb{K}$ and let P, Q be polynomials. Then λP, $P+Q$, and PQ defined according to Not. 6.2 are polynomials as well. More precisely, if $\lambda = 0$ or $P \equiv 0$, then $\lambda P = 0$; if $P \equiv 0$, then $P + Q = Q$; if $Q \equiv 0$, then $P + Q = P$; if $P \equiv 0$ or $Q \equiv 0$, then $PQ = 0$. If $\lambda \neq 0$ and

$$P(x) = \sum_{j=0}^{n} a_j x^j, \quad Q(x) = \sum_{j=0}^{m} b_j x^j, \qquad (6.3)$$

$$\text{with } \deg(P) = n \geq 0, \quad \deg(Q) = m \geq 0, \quad n \geq m \geq 0,$$

then, defining $b_j := 0$ for each $j \in \{m+1, \ldots, n\}$ in case $n > m$,

$$(\lambda P)(x) = \sum_{j=0}^{n} (\lambda \, a_j) \, x^j, \qquad \deg(\lambda P) = n, \qquad (6.4a)$$

$$(P + Q)(x) = \sum_{j=0}^{n} (a_j + b_j) \, x^j, \qquad \deg(P + Q) \leq n = \max\{m, n\}, \qquad (6.4b)$$

$$(PQ)(x) = \sum_{j=0}^{m+n} c_j \, x^j, \qquad \deg(PQ) = m + n, \qquad (6.4c)$$

where, setting $a_k := 0$ for each $k \in \{n+1, \ldots, m+n\}$ and $b_k := 0$ for each $k \in \{m+1, \ldots, m+n\}$,

$$\underset{j \in \{0, \ldots, n+m\}}{\forall} c_j = \sum_{k=0}^{j} a_k b_{j-k}. \qquad (6.4d)$$

Formula (6.4c) can be proved by induction on $m = \deg(Q) \in \mathbb{N}_0$ as follows: For $m = 0$, we compute

$$(PQ)(x) = b_0 \sum_{j=0}^{n} a_j \, x^j = \sum_{j=0}^{n+0} b_0 a_j \, x^j,$$

i.e. $c_j = b_0 a_j = \sum_{k=0}^{j} a_k b_{j-k}$, which establishes the base case, remembering $b_{j-k} = 0$ for $j > k$. For the induction step, we compute, for $\deg(Q) = m + 1$,

$$
\begin{aligned}
(PQ)(x) \quad &= \quad \sum_{j=0}^{n} a_j \, x^j \sum_{\alpha=0}^{m+1} b_\alpha \, x^\alpha = \sum_{j=0}^{n} a_j \, x^j \left(b_{m+1} x^{m+1} + \sum_{\alpha=0}^{m} b_\alpha \, x^\alpha \right) \\
&\overset{\text{ind. hyp.}}{=} \quad \sum_{j=0}^{n} a_j b_{m+1} \, x^{m+1+j} + \sum_{j=0}^{m+n} \left(\sum_{k=0}^{j} a_k b_{j-k} \right) x^j \\
&= \quad \sum_{j=m+1}^{m+n+1} a_{j-m-1} b_{m+1} \, x^j + \sum_{j=0}^{m+n} \left(\sum_{k=0}^{j} a_k b_{j-k} \right) x^j \\
&= \quad \sum_{j=0}^{m+n+1} \left(\sum_{k=0}^{j} a_k b_{j-k} \right) x^j,
\end{aligned}
$$

which completes the induction step. There is a notational issue in the second and third line in of the above computation, since, in both lines, the b_{m+1} in the first sum is the actual b_{m+1} from Q, but $b_{m+1} = 0$ in the second sum in both lines, which is due to the induction hypothesis being applied for $m < m+1$. This is actually used when combining both sums in the last step, computing, for $m+1 \le j \le m+n$: $a_{j-m-1} b_{m+1} \, x^j + a_{j-m-1} \cdot 0 \cdot x^j = a_{j-m-1} b_{m+1} \, x^j$. For $j = m+n+1$, one has $\sum_{k=0}^{m+n+1} a_k b_{m+n+1-k} = a_n b_{m+1}$, since $b_{m+n+1-k} = 0$ for $n > k$ and $a_k = 0$ for $k > n$.

Finally, $\deg(PQ) = m + n$ follows from $c_{m+n} = a_m b_n \ne 0$.

Theorem 6.5. (a) *For each polynomial P given in the form of (6.3) and each $\xi \in \mathbb{K}$, we have the identity*

$$P(x) = \sum_{j=0}^{n} b_j \, (x - \xi)^j, \tag{6.5}$$

where

$$\underset{j \in \{0, \dots, n\}}{\forall} \quad b_j = \sum_{k=j}^{n} a_k \binom{k}{j} \xi^{k-j}, \quad \text{in particular} \quad b_0 = P(\xi), \quad b_n = a_n. \tag{6.6}$$

(b) *If P is a polynomial with $n := \deg(P) \ge 1$, then, for each $\xi \in \mathbb{K}$, there exists a polynomial Q with $\deg(Q) = n - 1$ such that*

$$P(x) = P(\xi) + (x - \xi) \, Q(x). \tag{6.7}$$

In particular, if ξ is a zero of P, then $P(x) = (x - \xi) \, Q(x)$.

Proof. (a): For $\xi = 0$, there is nothing to prove. For $\xi \neq 0$, defining the auxiliary variable $\eta := x - \xi$, we obtain $x = \xi + \eta$ and

$$P(x) = \sum_{k=0}^{n} a_k(\xi + \eta)^k \overset{(5.22)}{=} \sum_{k=0}^{n}\sum_{j=0}^{k} a_k \binom{k}{j} \xi^{k-j}\eta^j = \sum_{k=0}^{n}\sum_{j=0}^{n} a_k \binom{k}{j} \xi^{k-j}\eta^j$$

$$= \sum_{j=0}^{n}\sum_{k=0}^{n} a_k \binom{k}{j} \xi^{k-j}\eta^j = \sum_{j=0}^{n}\sum_{k=j}^{n} a_k \binom{k}{j} \xi^{k-j}\eta^j, \tag{6.8}$$

which is (6.5).

(b): According to (a), we have

$$P(x) = P(\xi) + (x - \xi)\,Q(x), \quad \text{with} \quad Q(x) = \sum_{j=1}^{n} b_j\,(x-\xi)^{j-1} = \sum_{j=0}^{n-1} b_{j+1}\,(x-\xi)^j, \tag{6.9}$$

proving (b). $\blacksquare$

Theorem 6.6. (a) *If P is a polynomial with $n := \deg(P) \geq 0$, then P has at most n zeros.*

(b) *Let P, Q be polynomials as in (6.3) with $n = m$, $\deg(P) \leq n$, and $\deg(Q) \leq n$. If $P(x_j) = Q(x_j)$ at $n + 1$ distinct points $x_0, x_1, \ldots, x_n \in \mathbb{K}$, then $a_j = b_j$ for each $j \in \{0, \ldots, n\}$.*

Consequence 1: If P, Q with degree $\leq n$ agree at $n+1$ distinct points, then $P = Q$.

Consequence 2: If we know $P = Q$, then they agree everywhere, in particular at $\max\{\deg(P), \deg(Q)\} + 1$ distinct points, which implies they have the same coefficients.

Proof. (a): For $n = 0$, P is constant, but not the zero polynomial, i.e. $P \equiv a_0 \neq 0$ with no zeros as claimed. For $n \in \mathbb{N}$, the proof is conducted by induction. The base case $(n = 1)$ is provided by the observation that $\deg(P) = 1$ implies P is the affine function with $P(x) = a_0 + a_1 x$, $a_1 \neq 0$, i.e. P has precisely one zero at $\xi = -a_0/a_1$. For the induction step, assume $\deg(P) = n + 1$. If P has no zeros, then the assertion of (a) holds true. Otherwise, P has at least one zero $\xi \in \mathbb{K}$, and, according to Th. 6.5(b), there exists a polynomial Q such that $\deg(Q) = n$ and

$$P(x) = (x - \xi)\,Q(x). \tag{6.10}$$

From the induction hypothesis, we gather that Q has at most n zeros, i.e. (6.10) implies P has at most $n + 1$ zeros, which completes the induction.

(b): If $P(x_j) = Q(x_j)$ at $n + 1$ distinct points x_j, then each of these points is a zero of $P - Q$. Thus $P - Q$ is a polynomial of degree $\leq n$ with at least $n + 1$ zeros. Then (a) implies $\deg(P - Q) = -1$, i.e. $P - Q$ is the zero polynomial, i.e. $a_j - b_j = 0$ for each $j \in \{0, \ldots, n\}$. $\blacksquare$

Remark 6.7. Let P be a polynomial with $n := \deg(P) \geq 0$. According to Th. 6.6(a), P has at most n zeros. Using Th. 6.5(b) for an induction shows there exists $k \in \{0, \ldots, n\}$ and a polynomial Q of degree $n - k$ such that

$$P(x) = Q(x) \prod_{j=1}^{k} (x - \xi_j) = (x - \xi_1)(x - \xi_2) \cdots (x - \xi_k) Q(x), \qquad (6.11a)$$

where Q does not have any zeros in $\mathbb{K}$ and $\{\xi_1, \ldots, \xi_k\} = \{\xi \in \mathbb{K} : P(\xi) = 0\}$ is the set of zeros of P. It can of course happen that P does not have any zeros and $P = Q$ (no ξ_j exist). It can also occur that some of the ξ_j in (6.11a) are identical. Thus, we can rewrite (6.11a) as

$$P(x) = Q(x) \prod_{j=1}^{l} (x - \lambda_j)^{m_j} = (x - \lambda_1)^{m_1} (x - \lambda_2)^{m_2} \cdots (x - \lambda_l)^{m_l} Q(x), \qquad (6.11b)$$

where $\lambda_1, \ldots, \lambda_l$, $l \in \{0, \ldots, k\}$, are the *distinct* zeros of P, and $m_j \in \mathbb{N}$ with $\sum_{j=1}^{l} m_j = k$. Then m_j is called the *multiplicity* of the zero λ_j of P.

7 Limits and Convergence in the Real and Complex Numbers

7.1 Sequences

Recall from Def. 2.15(b) that a sequence in $\mathbb{K}$ is a function $f : \mathbb{N} \longrightarrow \mathbb{K}$, in this context usually denoted as $f = (z_n)_{n \in \mathbb{N}}$ or $(z_1, z_2, \ldots)$ with $z_n := f(n)$. Sometimes the sequence also has the form $(z_n)_{n \in I}$, where $I \neq \emptyset$ is a countable index set (e.g. $I = \mathbb{N}_0$) different from $\mathbb{N}$ (in the context of convergence (see the following Def. 7.1), I must be $\mathbb{N}$ or it must have the same cardinality as $\mathbb{N}$, i.e. finite I are not permissible).

Definition 7.1. The sequence $(z_n)_{n \in \mathbb{N}}$ in $\mathbb{K}$ is said to be *convergent with limit* $z \in \mathbb{K}$ if, and only if, for each $\epsilon > 0$, there exists an index $N \in \mathbb{N}$ such that $|z_n - z| < \epsilon$ for every index $n > N$. The notation for $(z_n)_{n \in \mathbb{N}}$ converging to z is $\lim_{n \to \infty} z_n = z$ or $z_n \to z$ for $n \to \infty$. Thus, by definition,

$$\lim_{n \to \infty} z_n = z \quad \Leftrightarrow \quad \forall_{\epsilon \in \mathbb{R}^+} \ \exists_{N \in \mathbb{N}} \ \forall_{n > N} \ |z_n - z| < \epsilon. \qquad (7.1)$$

The sequence $(z_n)_{n \in \mathbb{N}}$ in $\mathbb{K}$ is called *divergent* if, and only if, it is not convergent.

Example 7.2. (a) For every constant sequence $(z_n)_{n \in \mathbb{N}} = (a)_{n \in \mathbb{N}}$ with $a \in \mathbb{K}$, one has $\lim_{n \to \infty} z_n = \lim_{n \to \infty} a = a$: Since, for each $n \in \mathbb{N}$, $|z_n - a| = |a - a| = 0$, one can choose $N = 1$ for each $\epsilon > 0$.

(b) $\lim_{n \to \infty} \frac{1}{n+a} = 0$ for each $a \in \mathbb{C}$: Here $z_n := 1/(n + a)$ (if $n = -a$, then set $z_n := w$ with $w \in \mathbb{C}$ arbitrary). Given $\epsilon > 0$, choose an arbitrary $N \in \mathbb{N}$ with $N \geq \epsilon^{-1} + |a|$. Then, for each $n \geq N$, we compute $|n + a| = |n - (-a)| \geq |n - |a|| = n - |a| > N - |a| \geq \epsilon^{-1}$, and, thus, $|z_n| = |n + a|^{-1} < \epsilon$ as desired.

(c) $((-1)^n)_{n \in \mathbb{N}}$ is *not* convergent: We have $z_n = 1$ for each even n and $z_n = -1$ for each odd n. Thus, for each $z \neq 1$ and each even n, $|z_n - z| = |1 - z| > |1 - z|/2 =: \epsilon > 0$, i.e. z is not a limit of $(z_n)_{n \in \mathbb{N}}$. However, $z = 1$ is also not a limit of the sequence, since, for each odd n, $|z_n - 1| = |-1 - 1| = 2 > 1 =: \epsilon > 0$, proving that the sequence has no limit.

Theorem 7.3. (a) *Let $(z_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{C}$. Then $(z_n)_{n \in \mathbb{N}}$ is convergent in $\mathbb{C}$ if, and only if, both $(\operatorname{Re} z_n)_{n \in \mathbb{N}}$ and $(\operatorname{Im} z_n)_{n \in \mathbb{N}}$ are convergent in $\mathbb{R}$. Moreover, in that case,*

$$\lim_{n \to \infty} z_n = z \quad \Leftrightarrow \quad \lim_{n \to \infty} \operatorname{Re} z_n = \operatorname{Re} z \wedge \lim_{n \to \infty} \operatorname{Im} z_n = \operatorname{Im} z. \tag{7.2}$$

(b) *Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}$ and $z \in \mathbb{C}$. Then*

$$\lim_{n \to \infty} x_n = z \quad \Rightarrow \quad z \in \mathbb{R}. \tag{7.3}$$

Proof. (a): Suppose $(z_n)_{n \in \mathbb{N}}$ converges to $z \in \mathbb{C}$. Then, given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for each $n > N$, $|z_n - z| < \epsilon$. In consequence, for each $n > N$,

$$|\operatorname{Re} z_n - \operatorname{Re} z| = |\operatorname{Re}(z_n - z)| \overset{\text{Th. 5.9(d)}}{\leq} |z_n - z| < \epsilon, \tag{7.4}$$

proving $\lim_{n \to \infty} \operatorname{Re} z_n = \operatorname{Re} z$. The proof of $\lim_{n \to \infty} \operatorname{Im} z_n = \operatorname{Im} z$ is completely analogous. Conversely, suppose there are $x, y \in \mathbb{R}$ such that $\lim_{n \to \infty} \operatorname{Re} z_n = x$ and $\lim_{n \to \infty} \operatorname{Im} z_n = y$. Here we encounter, for the first time, what is sometimes called an $\epsilon/2$ argument: Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for each $n > N$, $|\operatorname{Re} z_n - x| < \epsilon/2$ and $|\operatorname{Im} z_n - y| < \epsilon/2$, implying, for each $n > N$,

$$\begin{aligned} |z_n - (x + iy)| &= |\operatorname{Re} z_n + i \operatorname{Im} z_n - (x + iy)| \\ &\leq |\operatorname{Re} z_n - x| + |i| |\operatorname{Im} z_n - y| < \epsilon/2 + \epsilon/2 = \epsilon, \end{aligned} \tag{7.5}$$

proving $\lim_{n \to \infty} z_n = x + iy$.

(b) is a direct consequence of (a). ∎

Example 7.4. (a) According to Th. 7.3(a), we have

$$\lim_{n \to \infty} \left(\sqrt{2} + \frac{i}{n - 17} \right) \overset{\text{Ex. 7.2(a),(b)}}{=} \sqrt{2} + 0i = \sqrt{2}.$$

(b) According to Th. 7.3(a) and Ex. 7.2(c), the sequence $(\frac{1}{n} + (-1)^n i)_{n \in \mathbb{N}}$ is divergent.

Another important example relies on the following inequality:

Proposition 7.5 (Bernoulli's Inequality). *For each $n \in \mathbb{N}_0$ and each $x \in [-1, \infty[$, we have*

$$(1 + x)^n \geq 1 + nx, \tag{7.6}$$

with strict inequality whenever $n > 1$ and $x \neq 0$.

Proof. For $n = 0$, (7.6) reads $1 \geq 1$, for $n = 1$, (7.6) reads $1 + x \geq 1 + x$, for $n = 2$, (7.6) reads $(1 + x)^2 = 1 + 2x + x^2 \geq 1 + 2x$, all three statements being trivially true, in the case $n = 2$ with strict inequality for $x \neq 0$. We now proceed by induction for $n \geq 2$. For the induction step, one estimates

$$(1 + x)^{n+1} = (1 + x)^n (1 + x) \overset{\text{ind. hyp., } x \geq -1}{\geq} (1 + nx)(1 + x) = 1 + (n + 1)x + nx^2$$
$$\geq 1 + (n + 1)x, \tag{7.7}$$

with strict inequality for $x \neq 0$.   ∎

Example 7.6. We have, for each $q \in \mathbb{C}$,

$$|q| < 1 \quad \Rightarrow \quad \lim_{n \to \infty} q^n = 0 : \tag{7.8}$$

For $q = 0$, there is nothing to prove. For $0 < |q| < 1$, it is $|q|^{-1} > 1$, i.e. $h := |q|^{-1} - 1 > 0$. Thus, for each $\epsilon > 0$ and $N \geq 1/(\epsilon h)$, we obtain

$$n > N \quad \Rightarrow \quad |q|^{-n} = (1 + h)^n \overset{(7.6)}{\geq} 1 + nh > nh > 1/\epsilon \quad \Rightarrow \quad |q^n| = |q|^n < \epsilon. \tag{7.9}$$

Definition 7.7. (a) Given $z \in \mathbb{K}$ and $\epsilon \in \mathbb{R}^+$, we call the set $B_\epsilon(z) := \{w \in \mathbb{K} : |w - z| < \epsilon\}$ the ϵ-*neighborhood* of z or, in anticipation of Analysis II, the (open) ϵ-*ball* with center z (in fact, for $\mathbb{K} = \mathbb{C}$, $B_\epsilon(z)$ represents an open disk in the complex plane with center z and radius ϵ, whereas, for $\mathbb{K} = \mathbb{R}$, $B_\epsilon(z) =]z - \epsilon, z + \epsilon[$ is the open interval with center z and length 2ϵ). More generally, a set $U \subseteq \mathbb{K}$ is called a *neighborhood* of z if, and only if, there exists $\epsilon > 0$ with $B_\epsilon(z) \subseteq U$ (so, for example, for $\epsilon > 0$, $B_\epsilon(z)$ is always a neighborhood of z, whereas $\mathbb{R}$ and $[z - \epsilon, \infty[$ are neighborhoods of z for $\mathbb{K} = \mathbb{R}$, but not for $\mathbb{K} = \mathbb{C}$ ($[z - \epsilon, \infty[$ not even being defined for $z \notin \mathbb{R}$); the sets $\{z\}$, $\{w \in \mathbb{K} : \operatorname{Re} w \geq \operatorname{Re} z\}$, $\{w \in \mathbb{K} : \operatorname{Re} w \geq \operatorname{Re} z + \epsilon\}$ are never neighborhoods of z).

(b) If $\phi(n)$ is a statement for each $n \in \mathbb{N}$, then $\phi(n)$ is said to be true for *almost all* $n \in \mathbb{N}$ if, and only if, there exists a *finite* subset $A \subseteq \mathbb{N}$ such that $\phi(n)$ is true for each $n \in \mathbb{N} \setminus A$, i.e. if, and only if, $\phi(n)$ is always true, with the possible exception of finitely many cases.

Remark 7.8. In the language of Def. 7.7, the sequence $(z_n)_{n \in \mathbb{N}}$ converges to z if, and only if, every neighborhood of z contains almost all z_n.

Definition 7.9. The sequence $(z_n)_{n \in \mathbb{N}}$ in $\mathbb{K}$ is called *bounded* if, and only if, the *set* $\{|z_n| : n \in \mathbb{N}\}$ is bounded in the sense of Def. 2.27(a).

Proposition 7.10. *Let $(z_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{K}$.*

(a) *Limits are unique, that means if $z, w \in \mathbb{K}$ such that $\lim_{n \to \infty} z_n = z$ and $\lim_{n \to \infty} z_n = w$, then $z = w$.*

(b) *If $(z_n)_{n \in \mathbb{N}}$ is convergent, then it is bounded.*

Proof. (a): Exercise.

(b): If $\lim_{n\to\infty} z_n = z$, then $A := \{|z_n| : |z_n - z| \geq 1\} \cup \{|z_1|\}$ is nonempty and finite. According to Th. 3.13(a), A has an upper bound M. Then $\max\{M, |z| + 1\}$ is an upper bound for $\{|z_n| : n \in \mathbb{N}\}$, and 0 is always a lower bound, showing that the sequence is bounded. ∎

Proposition 7.11. *Let $(z_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{C}$ with $\lim_{n\to\infty} z_n = 0$.*

(a) *If $(b_n)_{n\in\mathbb{N}}$ is a sequences in $\mathbb{C}$ such that there exists $C \in \mathbb{R}^+$ with $|b_n| \leq C|z_n|$ for almost all n, then $\lim_{n\to\infty} b_n = 0$.*

(b) *If $(c_n)_{n\in\mathbb{N}}$ is a bounded sequence in $\mathbb{C}$, then $\lim_{n\to\infty}(c_n z_n) = 0$.*

Proof. (a): Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|z_n| < \epsilon/C$ and $|b_n| \leq C|z_n|$ for each $n > N$. Then, for each $n > N$, $|b_n| \leq C|z_n| < \epsilon$, proving $\lim_{n\to\infty} b_n = 0$.

(b): If $(c_n)_{n\in\mathbb{N}}$ is bounded, then there exists $C \in \mathbb{R}^+$ such that $|c_n| \leq C$ for each $n \in \mathbb{N}$. Thus, $|c_n z_n| \leq C|z_n|$ for each $n \in \mathbb{N}$, implying $\lim_{n\to\infty}(c_n z_n) = 0$ via (a). ∎

Example 7.12. The sequences $((-1)^n)_{n\in\mathbb{N}}$ and $(b)_{n\in\mathbb{N}}$ with $b \in \mathbb{C}$ are bounded. Since, for each $a \in \mathbb{C}$, $\lim_{n\to\infty} \frac{1}{n+a} = 0$ by Example 7.2(b), we obtain

$$\lim_{n\to\infty} \frac{(-1)^n}{n+a} = \lim_{n\to\infty} \frac{b}{n+a} = 0 \tag{7.10}$$

from Prop. 7.11(b).

Theorem 7.13. (a) *Let $(z_n)_{n\in\mathbb{N}}$ and $(w_n)_{n\in\mathbb{N}}$ be sequences in $\mathbb{C}$. Moreover, let $z, w \in \mathbb{C}$ with $\lim_{n\to\infty} z_n = z$ and $\lim_{n\to\infty} w_n = w$. We have the following identities:*

$$\lim_{n\to\infty} (\lambda z_n) = \lambda z \quad \text{for each } \lambda \in \mathbb{C}, \tag{7.11a}$$

$$\lim_{n\to\infty} (z_n + w_n) = z + w, \tag{7.11b}$$

$$\lim_{n\to\infty} (z_n w_n) = zw, \tag{7.11c}$$

$$\lim_{n\to\infty} z_n/w_n = z/w \quad \text{given all } w_n \neq 0 \text{ and } w \neq 0, \tag{7.11d}$$

$$\lim_{n\to\infty} |z_n| = |z|, \tag{7.11e}$$

$$\lim_{n\to\infty} \bar{z}_n = \bar{z}, \tag{7.11f}$$

$$\lim_{n\to\infty} z_n^p = z^p \quad \text{for each } p \in \mathbb{N}. \tag{7.11g}$$

(b) *Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be sequences in $\mathbb{R}$. Moreover, let $x, y \in \mathbb{R}$ with $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$. Then*

$$\lim_{n\to\infty} \max\{x_n, y_n\} = \max\{x, y\}, \tag{7.12a}$$

$$\lim_{n\to\infty} \min\{x_n, y_n\} = \min\{x, y\}. \tag{7.12b}$$

(c) *If, in the situation of* (b) *(i.e. for real sequences),* $x_n \leq y_n$ *holds for almost all* $n \in \mathbb{N}$*, then* $x \leq y$*. In particular, if almost all* $x_n \geq 0$*, then* $x \geq 0$*.*

Proof. We start with the identities of (a).

(7.11a): For $\lambda = 0$, there is nothing to prove. For $\lambda \neq 0$ and $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for each $n > N$, $|z_n - z| < \epsilon/|\lambda|$, implying

$$\mathop{\forall}_{n>N} |\lambda z_n - \lambda z| = |\lambda| |z_n - z| < \epsilon. \tag{7.13a}$$

(7.11b): Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for each $n > N$, $|z_n - z| < \epsilon/2$ and $|w_n - w| < \epsilon/2$, implying

$$\mathop{\forall}_{n>N} |z_n + w_n - (z + w)| \leq |z_n - z| + |w_n - w| < \epsilon/2 + \epsilon/2 = \epsilon. \tag{7.13b}$$

(7.11c): Let $M_1 := \max\{|z|, 1\}$. According to Prop. 7.10(b), there exists $M_2 \in \mathbb{R}^+$ such that M_2 is an upper bound for $\{|w_n| : n \in \mathbb{N}\}$. Moreover, given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for each $n > N$, $|z_n - z| < \epsilon/(2M_2)$ and $|w_n - w| < \epsilon/(2M_1)$, implying

$$\mathop{\forall}_{n>N} \left(\begin{matrix} |z_n w_n - zw| = |(z_n - z)w_n + z(w_n - w)| \\[2mm] \leq |w_n| \cdot |z_n - z| + |z| \cdot |w_n - w| < \dfrac{M_2 \epsilon}{2M_2} + \dfrac{M_1 \epsilon}{2M_1} = \epsilon. \end{matrix} \right) \tag{7.13c}$$

(7.11d): We first consider the case, where all $z_n = 1$. Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for each $n > N$, $|w_n - w| < \epsilon |w|^2/2$ and $|w_n - w| < |w|/2$ (since $w \neq 0$ for this case), implying $|w| \leq |w - w_n| + |w_n| < |w|/2 + |w_n|$ and $|w_n| > |w|/2$. Thus,

$$\mathop{\forall}_{n>N} \left| \frac{1}{w_n} - \frac{1}{w} \right| = \left| \frac{w_n - w}{w_n w} \right| \leq \frac{2|w_n - w|}{|w|^2} < \frac{2}{|w|^2} \frac{\epsilon |w|^2}{2} = \epsilon. \tag{7.13d}$$

The general case now follows from (7.11c).

(7.11e): This is a consequence of the inverse triangle inequality (5.13): Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for each $n > N$, $|z_n - z| < \epsilon$, implying

$$\mathop{\forall}_{n>N} \big| |z_n| - |z| \big| \leq |z_n - z| < \epsilon. \tag{7.13e}$$

(7.11f): Write $z_n = x_n + iy_n$ and $z = x + iy$ with $x_n, y_n, x, y \in \mathbb{R}$, $n \in \mathbb{N}$. Then we know $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$ from (7.2), and

$$\lim_{n\to\infty} \bar{z}_n = \lim_{n\to\infty} (x_n - iy_n) \overset{(7.11a),(7.11b)}{=} x - iy = \bar{z}, \tag{7.13f}$$

which establishes the case.

(7.11g) follows by induction from (7.11c) (cf. (7.16b) below).

The proofs for the two identities of (b) are left as exercises.

(c): Proceeding by contraposition, assume $x > y$ and set $s := (x+y)/2$. Then $y < s < x$ and $y_n < s < x_n$ holds for almost all n, i.e. $x_n \leq y_n$ does not hold for almost all n. ∎

Example 7.14. (a) $\lim_{n\to\infty} \frac{n+a}{n+b} = 1$ for each $a, b \in \mathbb{C}$: Here $z_n := (n+a)/(n+b)$ (if $n = -b$, then set $z_n := w$ with $w \in \mathbb{C}$ arbitrary). Using (7.11b) and (7.11d), one obtains

$$\lim_{n\to\infty} \frac{n+a}{n+b} = \lim_{n\to\infty} \frac{1 + a/n}{1 + b/n} = \frac{\lim\limits_{n\to\infty} 1 + \lim\limits_{n\to\infty} \frac{a}{n}}{\lim\limits_{n\to\infty} 1 + \lim\limits_{n\to\infty} \frac{b}{n}} = \frac{1+0}{1+0} = 1. \tag{7.14}$$

(b) Using (7.11b), (7.11d), and (7.11g), one obtains

$$\lim_{n\to\infty} \frac{2n^5 - 3in^3 + 2i}{3n^5 + 17n} = \lim_{n\to\infty} \frac{2 - 3i/n^2 + 2i/n^5}{3 + 17/n^4} = \frac{2+0+0}{3+0} = \frac{2}{3}. \tag{7.15}$$

Corollary 7.15. *For $k \in \mathbb{N}$, let $(z_n^{(1)})_{n\in\mathbb{N}}, \ldots, (z_n^{(k)})_{n\in\mathbb{N}}$ be sequences in $\mathbb{C}$. Moreover, let $z^{(1)}, \ldots, z^{(k)} \in \mathbb{C}$ with $\lim_{n\to\infty} z_n^{(j)} = z^{(j)}$ for each $j \in \{1, \ldots, k\}$. Then*

$$\lim_{n\to\infty} \sum_{j=1}^{k} z_n^{(j)} = \sum_{j=1}^{k} z^{(j)}, \tag{7.16a}$$

$$\lim_{n\to\infty} \prod_{j=1}^{k} z_n^{(j)} = \prod_{j=1}^{k} z^{(j)}. \tag{7.16b}$$

Proof. (7.16) follows by simple inductions from (7.11b) and (7.11c), respectively. ∎

Theorem 7.16 (Sandwich Theorem). *Let $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$, and $(a_n)_{n\in\mathbb{N}}$ be sequences in $\mathbb{R}$. If $x_n \leq a_n \leq y_n$ holds for almost all $n \in \mathbb{N}$, then*

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = x \in \mathbb{R} \quad \Rightarrow \quad \lim_{n\to\infty} a_n = x. \tag{7.17}$$

Proof. Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for each $n > N$, $x_n \leq a_n \leq y_n$, $|x_n - x| < \epsilon$, and $|y_n - x| < \epsilon$, implying

$$\underset{n>N}{\forall} \; x - \epsilon < x_n \leq a_n \leq y_n < x + \epsilon, \tag{7.18}$$

which establishes the case. ∎

Example 7.17. Since, $0 < \frac{1}{n!} \leq \frac{1}{n}$ holds for each $n \in \mathbb{N}$, the Sandwich Th. 7.16 implies

$$\lim_{n\to\infty} \frac{1}{n!} = 0. \tag{7.19}$$

Definition 7.18. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. The sequence is said to *diverge to* ∞ (resp. to $-\infty$), denoted $\lim_{n\to\infty} x_n = \infty$ (resp. $\lim_{n\to\infty} x_n = -\infty$) if, and only if, for each $K \in \mathbb{R}$, almost all x_n are bigger (resp. smaller) than K. Thus,

$$\lim_{n\to\infty} x_n = \infty \quad \Leftrightarrow \quad \underset{K\in\mathbb{R}}{\forall} \; \underset{N\in\mathbb{N}}{\exists} \; \underset{n>N}{\forall} \; x_n > K, \tag{7.20a}$$

$$\lim_{n\to\infty} x_n = -\infty \quad \Leftrightarrow \quad \underset{K\in\mathbb{R}}{\forall} \; \underset{N\in\mathbb{N}}{\exists} \; \underset{n>N}{\forall} \; x_n < K. \tag{7.20b}$$

Theorem 7.19. *Suppose $S := (x_n)_{n \in \mathbb{N}}$ is a monotone sequence in $\mathbb{R}$ (increasing or decreasing). Defining $A := \{x_n : n \in \mathbb{N}\}$, the following holds:*

$$\lim_{n \to \infty} x_n = \begin{cases} \sup A & \text{if } S \text{ is increasing and bounded,} \\ \infty & \text{if } S \text{ is increasing and not bounded,} \\ \inf A & \text{if } S \text{ is decreasing and bounded,} \\ -\infty & \text{if } S \text{ is decreasing and not bounded.} \end{cases} \tag{7.21}$$

Proof. We treat the increasing case; the decreasing case is proved completely analogously. If A is bounded and $\epsilon > 0$, let $K := \sup A - \epsilon$; if A is unbounded, then let $K \in \mathbb{R}$ be arbitrary. In both cases, since K can not be an upper bound, there exists $N \in \mathbb{N}$ such that $x_N > K$. Since the sequence is increasing, for each $n > N$, $x_N \leq x_n$, showing $|\sup A - x_n| < \epsilon$ in the bounded case, and $x_n > K$ in the unbounded case. ∎

Example 7.20. Theorem 7.19 implies

$$\underset{k \in \mathbb{N}}{\forall} \left(\lim_{n \to \infty} n^k = \infty, \quad \lim_{n \to \infty} (-n^k) = -\infty \right). \tag{7.22}$$

It is sometimes necessary to consider so-called subsequences and reorderings of a given sequence. Here, we are interested in sequences in $\mathbb{R}$ or $\mathbb{C}$, but for subsequences and reorderings it is irrelevant in which set A the sequence takes its values. As it presents virtually no extra difficulty to introduce the notions for general sequences, and since we will need to consider sequences with values in sets other than $\mathbb{R}$ or $\mathbb{C}$ in Analysis II, we admit general sequences in the following definition.

Definition 7.21. Let A be an arbitrary nonempty set. Consider a sequence $\sigma : \mathbb{N} \longrightarrow A$. Given a function $\phi : \mathbb{N} \longrightarrow \mathbb{N}$ (that means $(\phi(n))_{n \in \mathbb{N}}$ constitutes a sequence of indices), the new sequence $(\sigma \circ \phi) : \mathbb{N} \longrightarrow A$ is called a *subsequence* of σ if, and only if, ϕ is strictly increasing (i.e. $1 \leq \phi(1) < \phi(2) < \dots$). Moreover, $\sigma \circ \phi$ is called a *reordering* of σ if, and only if, ϕ is bijective. One can write σ in the form $(z_n)_{n \in \mathbb{N}}$ by setting $z_n := \sigma(n)$, and one can write $\sigma \circ \phi$ in the form $(w_n)_{n \in \mathbb{N}}$ by setting $w_n := (\sigma \circ \phi)(n) = z_{\phi(n)}$. Especially for a subsequence of $(z_n)_{n \in \mathbb{N}}$, it is also common to write $(z_{n_k})_{k \in \mathbb{N}}$. This notation corresponds to the one above if one lets $n_k := \phi(k)$. Analogous definitions work if the index set $\mathbb{N}$ of σ is replaced by a general countable nonempty index set I.

Example 7.22. Consider the sequence $(1, 2, 3, \dots)$. Then $(2, 4, 6, \dots)$ constitutes a subsequence and $(2, 1, 4, 3, 6, 5, \dots)$ constitutes a reordering. Using the notation of Def. 7.21, the original sequence is given by $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$, $\sigma(n) := n$; the subsequence is selected via $\phi_1 : \mathbb{N} \longrightarrow \mathbb{N}$, $\phi_1(n) := 2n$; and the reordering is accomplished via

$$\phi_2 : \mathbb{N} \longrightarrow \mathbb{N}, \quad \phi_2(n) := \begin{cases} n+1 & \text{if } n \text{ is odd,} \\ n-1 & \text{if } n \text{ is even.} \end{cases}$$

Proposition 7.23. *Let $(z_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{C}$. If $\lim_{n \to \infty} z_n = z$, then every subsequence and every reordering of $(z_n)_{n \in \mathbb{N}}$ is also convergent with limit z.*

Proof. Let $(w_n)_{n\in\mathbb{N}}$ be a subsequence of of $(z_n)_{n\in\mathbb{N}}$, i.e. there is a strictly increasing function $\phi : \mathbb{N} \longrightarrow \mathbb{N}$ such that $w_n = z_{\phi(n)}$. If $\lim_{n\to\infty} z_n = z$, then, given $\epsilon > 0$, there is $N \in \mathbb{N}$ such that $z_n \in B_\epsilon(z)$ for each $n > N$. For $\tilde{N}$ choose any number from $\mathbb{N}$ that is $\geq N$ and in $\phi(\mathbb{N})$. Take $M := \phi^{-1}(\tilde{N})$ (where $\phi^{-1} : \phi(\mathbb{N}) \longrightarrow \mathbb{N}$). Then, for each $n > M$, one has $\phi(n) > \tilde{N} \geq N$, and, thus, $w_n = z_{\phi(n)} \in B_\epsilon(z)$, showing $\lim_{n\to\infty} w_n = z$.

Let $(w_n)_{n\in\mathbb{N}}$ be a reordering of $(z_n)_{n\in\mathbb{N}}$, i.e. there is a bijective function $\phi : \mathbb{N} \longrightarrow \mathbb{N}$ such that $w_n = z_{\phi(n)}$. Let ϵ and N be as before. Define

$$M := \max\{\phi^{-1}(n) : n \leq N\}. \tag{7.23}$$

As ϕ is bijective, it is $\phi(n) > N$ for each $n > M$. Then, for each $n > M$, one has $w_n = z_{\phi(n)} \in B_\epsilon(z)$, showing $\lim_{n\to\infty} w_n = z$. ∎

Definition 7.24. Let $(z_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{K}$. A point $z \in \mathbb{K}$ is called a *cluster point* or an *accumulation point* of the sequence if, and only if, for each $\epsilon > 0$, $B_\epsilon(z)$ contains infinitely many members of the sequence (i.e. $\#\{n \in \mathbb{N} : z_n \in B_\epsilon(z)\} = \infty$).

Example 7.25. The sequence $((-1)^n)_{n\in\mathbb{N}}$ has cluster points 1 and -1.

Proposition 7.26. *A point $z \in \mathbb{K}$ is a cluster point of the sequence $(z_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ if, and only if, the sequence has a subsequence converging to z.*

Proof. If $(w_n)_{n\in\mathbb{N}}$ is a subsequence of $(z_n)_{n\in\mathbb{N}}$, $\lim_{n\to\infty} w_n = z$, then every $B_\epsilon(z)$, $\epsilon > 0$, contains infinitely many w_n, i.e. infinitely many z_n, i.e. z is a cluster point of $(z_n)_{n\in\mathbb{N}}$. Conversely, if z is a cluster point of $(z_n)_{n\in\mathbb{N}}$, then, inductively, define $\phi : \mathbb{N} \longrightarrow \mathbb{N}$ as follows: For $\phi(1)$, choose the index k of any point z_k in $B_1(z)$ (such a point exists, since z is a cluster point of the sequence). Now assume that $n > 1$ and that $\phi(m)$ have already been defined for each $m < n$. Let $M := \max\{\phi(m) : m < n\}$. Since $B_{\frac{1}{n}}(z)$ contains infinitely many z_k, there must be some $z_k \in B_{\frac{1}{n}}(z)$ such that $k > M$. Choose this k as $\phi(n)$. Thus, by construction, ϕ is strictly increasing, i.e. $(w_n)_{n\in\mathbb{N}}$ with $w_n := z_{\phi(n)}$ is a subsequence of $(z_n)_{n\in\mathbb{N}}$. Moreover, for each $\epsilon > 0$, there is $N \in \mathbb{N}$ such that $1/N < \epsilon$. Then, for each $n > N$, $w_n \in B_{\frac{1}{n}}(z) \subseteq B_{\frac{1}{N}}(z) \subseteq B_\epsilon(z)$, showing $\lim_{n\to\infty} w_n = z$. ∎

Theorem 7.27 (Bolzano-Weierstrass). *Every bounded sequence $S := (x_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ has at least one cluster point in $\mathbb{K}$. Moreover, for $\mathbb{K} = \mathbb{R}$, the set $A := \{x \in \mathbb{R} : x$ is cluster point of $S\}$ has a max $x^* \in \mathbb{R}$ and a min $x_* \in \mathbb{R}$, i.e. every bounded sequence in $\mathbb{R}$ has a largest and a smallest cluster point. In addition, for each $\epsilon > 0$, the inequality $x_* - \epsilon < x_n < x^* + \epsilon$ holds for almost all n.*

Proof. We first consider the case $\mathbb{K} = \mathbb{R}$. Define

$$A^* := \{x \in \mathbb{R} : x_n \leq x \text{ for almost all } n\}, \tag{7.24a}$$
$$A_* := \{x \in \mathbb{R} : x_n \geq x \text{ for almost all } n\}. \tag{7.24b}$$

We claim $A^* \neq \emptyset$ is bounded from below and $x^* = \max A = \inf A^*$; $A_* \neq \emptyset$ is bounded from above and $x_* = \min A = \sup A_*$. We prove the claim for A^* – the proof for A_* is

conducted completely analogous. Let $m, M \in \mathbb{R}$ be a lower and an upper bound for S, respectively. Then $M \in A^*$, showing $A^* \neq \emptyset$; and m is a lower bound for A^*. Since A^* is bounded from below, $a := \inf A^* \in \mathbb{R}$ by the completeness of $\mathbb{R}$. Moreover, for each $\epsilon > 0$, $a - \epsilon \notin A^*$, as a is a lower bound for A^*, i.e. $x_n > a - \epsilon$ holds for infinitely many $n \in \mathbb{N}$. On the other hand, $a + \epsilon/2 \in A^*$ follows from a being the largest lower bound of A^*, i.e. $x_n > a + \epsilon/2$ holds for only finitely many n (if any). In particular, we have shown $x_n < a + \epsilon$ holds for almost all n, and $a - \epsilon < x_n < a + \epsilon$ must hold for infinitely many n, showing a is a cluster point of S. To see that a is the largest cluster point of S (i.e. $a = \max A$), we have to show that $x > a$ implies x is not a cluster point of S. However, letting $\epsilon := x - a > 0$, we had seen above that $x_n > a + \epsilon/2$ holds for only finitely many n, i.e. $B_{\epsilon/2}(x)$ contains only finitely many x_n, showing x is not a cluster point of S.

It now remains to consider the complex case, i.e. a bounded sequence $S := (z_n)_{n \in \mathbb{N}}$ in $\mathbb{C}$. For each $n \in \mathbb{N}$, let $z_n = x_n + iy_n$ with $x_n, y_n \in \mathbb{R}$. Due to Th. 5.9(d), we have $|x_n| \leq |z_n|$ and $|y_n| \leq |z_n|$, i.e. the boundedness of S implies the boundedness of both $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$. Then we know that $(x_n)_{n \in \mathbb{N}}$ has a cluster point x and, by Prop. 7.26, S has a subsequence $(z_{n_j})_{j \in \mathbb{N}}$ such that $x = \lim_{j \to \infty} x_{n_j}$. As the subsequence $(y_{n_j})_{j \in \mathbb{N}}$ is still bounded, it must have a cluster point y and a subsequence $(y_{n_{j_k}})_{k \in \mathbb{N}}$ such that $y = \lim_{k \to \infty} y_{n_{j_k}}$. Since $x = \lim_{k \to \infty} x_{n_{j_k}}$ as well, we now have $\lim_{k \to \infty} z_{n_{j_k}} = x + iy =: z$, i.e. S has a subsequence converging to z. According to Prop. 7.26, z is a cluster point of S.   ∎

Definition 7.28. A sequence $(z_n)_{n \in \mathbb{N}}$ in $\mathbb{C}$ is defined to be a *Cauchy sequence* if, and only if, for each $\epsilon \in \mathbb{R}^+$, there exists $N \in \mathbb{N}$ such that $|z_n - z_m| < \epsilon$ for each $n, m > N$, i.e.

$$(z_n)_{n \in \mathbb{N}} \text{ Cauchy} \quad \Leftrightarrow \quad \underset{\epsilon \in \mathbb{R}^+}{\forall} \; \underset{N \in \mathbb{N}}{\exists} \; \underset{n,m > N}{\forall} \; |z_n - z_m| < \epsilon. \tag{7.25}$$

Theorem 7.29. *The sequence $(z_n)_{n \in \mathbb{N}}$ in $\mathbb{C}$ is convergent if, and only if, it is a Cauchy sequence.*

Proof. Suppose the sequence is convergent with $\lim_{n \to \infty} z_n = z$. Then, given $\epsilon > 0$, there is $N \in \mathbb{N}$ such that $z_n \in B_{\frac{\epsilon}{2}}(z)$ for each $n > N$. If $n, m > N$, then $|z_n - z_m| \leq |z_n - z| + |z - z_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$, establishing that $(z_n)_{n \in \mathbb{N}}$ is a Cauchy sequence.

Conversely, suppose the sequence is a Cauchy sequence. Using similar reasoning as in the proof of Prop. 7.10(b), we first show the sequence is bounded. If the sequence is Cauchy, then there exists $N \in \mathbb{N}$ such that $|z_n - z_m| < 1$ for all $n, m > N$. Thus, the set $A := \{|z_n| : |z_n - z_{N+1}| \geq 1\} \cup \{|z_1|\} \subseteq \mathbb{R}_0^+$ is nonempty and finite. According to Th. 3.13(a), A has an upper bound M. Then $\max\{M, |z_{N+1}| + 1\}$ is an upper bound for $\{|z_n| : n \in \mathbb{N}\}$, showing that the sequence is bounded. From Th. 7.27, we obtain that the sequence has a cluster point z. It remains to show $\lim_{n \to \infty} z_n = z$. Given $\epsilon > 0$, choose $N \in \mathbb{N}$ such that $|z_n - z_m| < \epsilon/2$ for all $n, m > N$. Since z is a cluster point, there exists $k > N$ such that $|z_k - z| < \epsilon/2$. Thus,

$$\underset{n > N}{\forall} \; |z_n - z| \leq |z_n - z_k| + |z_k - z| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \tag{7.26}$$

proving $\lim_{n \to \infty} z_n = z$.   ∎

Example 7.30. Consider the sequence $S := (s_n)_{n \in \mathbb{N}}$ defined by

$$s_n := \sum_{k=1}^{n} \frac{1}{k} = 1 + \frac{1}{2} + \cdots + \frac{1}{n}. \tag{7.27}$$

We claim S is *not* a Cauchy sequence and, thus, *not* convergent by Th. 7.29: For each $N \in \mathbb{N}$, we find $n, m > N$ such that $s_n - s_m > 1/2$, namely $m = N+1$ and $n = 2(N+1)$:

$$s_{2(N+1)} - s_{N+1} = \sum_{k=N+2}^{2(N+1)} \frac{1}{k} = \frac{1}{N+2} + \frac{1}{N+3} + \cdots + \frac{1}{2(N+1)}$$

$$> (N+1) \cdot \frac{1}{2(N+1)} = \frac{1}{2}. \tag{7.28}$$

While we have just seen that S is not convergent, it is clearly increasing, i.e. Th. 7.19 implies S is unbounded and $\lim_{n \to \infty} s_n = \infty$. Sequences defined by longer and longer sums are known as *series* and will be studied further in Sec. 7.3 below. The series of the present example is known as the *harmonic series*. It has become famous as the simplest example of a series that does *not* converge even though its summands converge to 0. In terms of the notation introduced in Sec. 7.3 below, we have shown

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots = \infty. \tag{7.29}$$

7.2 Continuity

7.2.1 Definitions and First Examples

Roughly, a function is continuous if a small change in its input results in a small change of its output. For functions defined on an interval, the notion of continuity makes precise the idea of a function having no jump – no discontinuity – at some point x in its domain. For example, we would say the sign function of (5.8) has precisely one jump – one discontinuity – at $x = 0$, whereas quadratic functions (or, more generally, polynomials) do not have any jumps – they are continuous.

Definition 7.31. Let $M \subseteq \mathbb{C}$. If $\zeta \in M$, then a function $f : M \longrightarrow \mathbb{K}$ is said to be *continuous* in ζ if, and only if, for each $\epsilon > 0$, there is $\delta > 0$ such that the distance between the values $f(z)$ and $f(\zeta)$ is less than ϵ, provided the distance between z and ζ is less than δ, i.e. if, and only if,

$$\underset{\epsilon \in \mathbb{R}^+}{\forall} \, \underset{\delta \in \mathbb{R}^+}{\exists} \, \underset{z \in M}{\forall} \, \left(|z - \zeta| < \delta \quad \Rightarrow \quad |f(z) - f(\zeta)| < \epsilon \right). \tag{7.30}$$

Moreover, f is called *continuous* if, and only if, f is continuous in every $\zeta \in M$. The set of all continuous functions from $f : M \longrightarrow \mathbb{K}$ is denoted by $C(M, \mathbb{K})$, $C(M) := C(M, \mathbb{R})$.

Example 7.32. (a) Every constant map $f : M \longrightarrow \mathbb{K}$, $\emptyset \neq M \subseteq \mathbb{C}$, is continuous: In this case, given ϵ, we can choose any $\delta > 0$ we want, say $\delta := 42$: If $\zeta, z \in M$, then $|f(\zeta) - f(z)| = 0 < \epsilon$, which holds independently of δ, in particular, if $|\zeta - z| < \delta$.

(b) Every affine function $f : \mathbb{K} \longrightarrow \mathbb{K}$, $f(z) := az + b$ is continuous: For $a = 0$, this follows from (a). For $a \neq 0$, given $\epsilon > 0$, choose $\delta := \epsilon/|a|$. Then,

$$\underset{\zeta, z \in \mathbb{K}}{\forall} \left(\begin{matrix} |z - \zeta| < \delta = \dfrac{\epsilon}{|a|} & \Rightarrow & |f(z) - f(\zeta)| = |az + b - a\zeta - b| \\ & & = |a| \, |z - \zeta| < |a| \dfrac{\epsilon}{|a|} = \epsilon \end{matrix} \right). \tag{7.31}$$

(c) The sign function of (5.8) is *not* continuous: It is continuous in each $\xi \in \mathbb{R} \setminus \{0\}$, but not continuous in 0: If $\xi \neq 0$, then, given $\epsilon > 0$, choose $\delta := |\xi|$. If $|x - \xi| < \delta$, then $\mathrm{sgn}(x) = \mathrm{sgn}(\xi)$, i.e. $|\mathrm{sgn}(x) - \mathrm{sgn}(\xi)| = 0 < \epsilon$, proving continuity in ξ. However, at 0, for $\epsilon := 1/2$, we have

$$\underset{\delta > 0}{\forall} \left| \mathrm{sgn}(0) - \mathrm{sgn}(\delta/2) \right| = |0 - 1| = 1 > \frac{1}{2} = \epsilon, \tag{7.32}$$

showing sgn is not continuous in 0.

Some subtleties arise from the possibility that f can be defined on subsets of $\mathbb{C}$ with very different properties. The notions introduced in Def. 7.33 help to deal with these subtleties.

Definition 7.33. Let $M \subseteq \mathbb{C}$.

(a) The point $z \in \mathbb{C}$ is called a *cluster point* or *accumulation point* of M if, and only if, each ϵ-neighborhood of z, $\epsilon \in \mathbb{R}^+$, contains infinitely many points of M, i.e. if, and only if,

$$\underset{\epsilon \in \mathbb{R}^+}{\forall} \#(M \cap B_\epsilon(z)) = \infty. \tag{7.33}$$

Note: A cluster point of M is not necessarily in M.

(b) The point z is called an *isolated point* of M if, and only if, there is $\epsilon \in \mathbb{R}^+$ such that $B_\epsilon(z) \cap M = \{z\}$. Note: An isolated point of M is always in M.

Proposition 7.34. *If $M \subseteq \mathbb{C}$, then each point of M is either a cluster point or an isolated point of M, i.e.*

$$M = \{z \in M : z \text{ cluster point of } M\} \, \dot\cup \, \{z \in M : z \text{ isolated point of } M\}. \tag{7.34}$$

Proof. Consider $z \in M$ that is not a cluster point of M. We have to show that z is an isolated point of M. Since z is not a cluster point of M, there exists $\tilde\epsilon > 0$ such that $A := (M \cap B_{\tilde\epsilon}(z)) \setminus \{z\}$ is finite. Define

$$\epsilon := \begin{cases} \min\{|a - z| : a \in A\} & \text{if } A \neq \emptyset, \\ \tilde\epsilon & \text{if } A = \emptyset. \end{cases} \tag{7.35}$$

Then $B_\epsilon(z) \cap M = \{z\}$, showing z is an isolated point of M. Finally, the union in (7.34) is clearly disjoint. ∎

Lemma 7.35. *Let $M \subseteq \mathbb{C}$, $f : M \longrightarrow \mathbb{K}$. If ζ is an isolated point of M, then f is always continuous in ζ.*

Proof. Independently of the concrete definition of f, we know there is $\delta > 0$ such that $B_\delta(\zeta) \cap M = \{\zeta\}$. In other words, if $z \in M$ with $|z - \zeta| < \delta$, then $z = \zeta$, implying $|f(z) - f(\zeta)| = 0 < \epsilon$ for each $\epsilon > 0$, showing f to be continuous in ζ. ∎

Example 7.36. (a) The sign function restricted to the set $M :=]-\infty, -1] \cup \{0\} \cup [1, \infty[$, i.e.

$$\mathrm{sgn}(x) = \begin{cases} 1 & \text{for } x \in [1, \infty[, \\ 0 & \text{for } x = 0, \\ -1 & \text{for } x \in]-\infty, -1] \end{cases}$$

is continuous: As in Ex. 7.32(c), one sees that sgn is continuous in each $\xi \in M \setminus \{0\}$. However, now it is also continuous in 0, since 0 is an isolated point of M.

(b) Every function $f : \mathbb{N} \longrightarrow \mathbb{K}$ is continuous, since every $n \in \mathbb{N}$ is an isolated point of $\mathbb{N}$ (due to $\{n\} = \mathbb{N} \cap B_{\frac{1}{2}}(n)$).

7.2.2   Continuity, Sequences, and Function Arithmetic

To make available the power of the results on convergent sequences from Sec. 7.1 to investigations regarding the continuity of functions, we need to understand the relationship between both notions. The core of this relationship is the contents of the following Th. 7.37, which provides a criterion allowing one to test continuity in terms of convergent sequences:

Theorem 7.37. *Let $M \subseteq \mathbb{C}$, $f : M \longrightarrow \mathbb{K}$. If $\zeta \in M$, then f is continuous in ζ if, and only if, for each sequence $(z_n)_{n \in \mathbb{N}}$ in M with $\lim_{n \to \infty} z_n = \zeta$, the sequence $(f(z_n))_{n \in \mathbb{N}}$ converges to $f(\zeta)$, i.e.*

$$\lim_{n \to \infty} z_n = \zeta \quad \Rightarrow \quad \lim_{n \to \infty} f(z_n) = f(\zeta). \tag{7.36}$$

Proof. If $\zeta \in M$ is an isolated point of M, then there is $\delta > 0$ such that $M \cap B_\delta(\zeta) = \{\zeta\}$. Then every $f : M \longrightarrow \mathbb{K}$ is continuous in ζ according to Lem. 7.35. On the other hand, every sequence in M converging to ζ must be finally constant and equal to ζ, i.e. (7.36) is trivially valid at ζ. Thus, the assertion of the theorem holds if $\zeta \in M$ is an isolated point of M.

If $\zeta \in M$ is not an isolated point of M, then ζ is a cluster point of M according to Prop. 7.34. So, for the remainder of the proof, let $\zeta \in M$ be a cluster point of M. Assume that f is continuous in ζ and $(z_n)_{n \in \mathbb{N}}$ is a sequence in M with $\lim_{n \to \infty} z_n = \zeta$. For each $\epsilon > 0$, there is $\delta > 0$ such that $z \in M$ and $|z - \zeta| < \delta$ implies $|f(z) - f(\zeta)| < \epsilon$. Since $\lim_{n \to \infty} z_n = \zeta$, there is also $N \in \mathbb{N}$ such that, for each $n > N$, $|z_n - \zeta| < \delta$. Thus, for each $n > N$, $|f(z_n) - f(\zeta)| < \epsilon$, proving $\lim_{n \to \infty} f(z_n) = f(\zeta)$. Conversely, assume that f is not continuous in ζ. We have to construct a sequence $(z_n)_{n \in \mathbb{N}}$ in M with

$\lim_{n\to\infty} z_n = \zeta$, but $(f(z_n))_{n\in\mathbb{N}}$ does not converge to $f(\zeta)$. Since f is not continuous in ζ, there must be some $\epsilon_0 > 0$ such that, for each $1/n$, $n \in \mathbb{N}$, there is at least one $z_n \in M$ satisfying $|z_n - \zeta| < 1/n$ and $|f(z_n) - f(\zeta)| \geq \epsilon_0$. Then $(z_n)_{n\in\mathbb{N}}$ is a sequence in M with $\lim_{n\to\infty} z_n = \zeta$ and $(f(z_n))_{n\in\mathbb{N}}$ does not converge to $f(\zeta)$. ∎

We can now apply the rules of Th. 7.13 to see that all the arithmetic operations defined in Not. 6.2 preserve continuity:

Theorem 7.38. *Let $M \subseteq \mathbb{C}$, $f,g : M \longrightarrow \mathbb{K}$, $\lambda \in \mathbb{K}$, $\zeta \in M$. If f,g are both continuous in ζ, then λf, $f + g$, fg, f/g for $g \neq 0$, $|f|$, $\operatorname{Re} f$, and $\operatorname{Im} f$ are all continuous in ζ. If $\mathbb{K} = \mathbb{R}$, then $\max(f,g)$, $\min(f,g)$, f^+ and f^-, are also all continuous in ζ.*

Proof. Let $(z_n)_{n\in\mathbb{N}}$ be a sequence in M such that $\lim_{n\to\infty} z_n = \zeta$. Then the continuity of f and g in ζ yields $\lim_{n\to\infty} f(z_n) = f(\zeta)$ and $\lim_{n\to\infty} g(z_n) = g(\zeta)$. Then

$$(7.11a) \qquad \Rightarrow \qquad \lim_{n\to\infty} (\lambda f)(z_n) = (\lambda f)(\zeta),$$

$$(7.11b) \qquad \Rightarrow \qquad \lim_{n\to\infty} (f + g)(z_n) = (f + g)(\zeta),$$

$$(7.11c) \qquad \Rightarrow \qquad \lim_{n\to\infty} (fg)(z_n) = (fg)(\zeta),$$

$$(7.11d) \qquad \Rightarrow \qquad \lim_{n\to\infty} (f/g)(z_n) = (f/g)(\zeta),$$

$$(7.11e) \qquad \Rightarrow \qquad \lim_{n\to\infty} |f|(z_n) = |f|(\zeta),$$

$$(7.2) \qquad \Rightarrow \qquad \lim_{n\to\infty} (\operatorname{Re} f)(z_n) = (\operatorname{Re} f)(\zeta),$$

$$(7.2) \qquad \Rightarrow \qquad \lim_{n\to\infty} (\operatorname{Im} f)(z_n) = (\operatorname{Im} f)(\zeta).$$

If f,g are both $\mathbb{R}$-valued, then we also have

$$(7.12a) \qquad \Rightarrow \qquad \lim_{n\to\infty} \max(f,g)(z_n) = \max(f,g)(\zeta),$$

$$(7.12b) \qquad \Rightarrow \qquad \lim_{n\to\infty} \min(f,g)(z_n) = \min(f,g)(\zeta),$$

and, finally, the continuity of f^+ and f^- follows from the continuity of $\max(f,g)$. ∎

Corollary 7.39. *A function $f : M \longrightarrow \mathbb{C}$, $M \subseteq \mathbb{C}$, is continuous in $\zeta \in M$ if, and only if, both $\operatorname{Re} f$ and $\operatorname{Im} f$ are continuous in ζ.*

Proof. If f is continuous in ζ, then $\operatorname{Re} f$ and $\operatorname{Im} f$ are both continuous in ζ by Th. 7.38. If $\operatorname{Re} f$ and $\operatorname{Im} f$ are both continuous in ζ, then, as

$$f = \operatorname{Re} f + i \operatorname{Im} f, \tag{7.37}$$

f is continuous in ζ, once again, by Th. 7.38. ∎

Example 7.40. (a) The continuity of the absolute value function $z \mapsto |z|$ on $\mathbb{K}$ can be concluded directly from (7.11e) and, alternatively, from combining the continuity of $f : \mathbb{K} \longrightarrow \mathbb{K}$, $f(z) = z$, according to Ex. 7.32(b), with the continuity of $|f|$ according to Th. 7.38.

(b) Every polynomial $P : \mathbb{K} \longrightarrow \mathbb{K}$, $P(x) = \sum_{j=0}^{n} a_j x^j$, $a_j \in \mathbb{K}$, is continuous: First note that every monomial $x \mapsto x^j$ is continuous on $\mathbb{K}$ by (7.11g). Then Th. 7.38 implies the continuity of $x \mapsto a_j x^j$ on $\mathbb{K}$. Now the continuity of P follows from (7.16a) or, alternatively, by an induction from the $f + g$ part of Th. 7.38.

(c) Let $P, Q : \mathbb{K} \longrightarrow \mathbb{K}$, be polynomials and let $A := Q^{-1}\{0\}$ the set of all zeros of Q (if any). Then the rational function $(P/Q) : \mathbb{K} \setminus A \longrightarrow \mathbb{K}$ is continuous as a consequence of (b) plus the f/g part of Th. 7.38.

Theorem 7.41. *Let $D_f, D_g \subseteq \mathbb{C}$, $f : D_f \longrightarrow \mathbb{C}$, $g : D_g \longrightarrow \mathbb{K}$, $f(D_f) \subseteq D_g$. If f is continuous in $\zeta \in D_f$ and g is continuous in $f(\zeta) \in D_g$, then $g \circ f : D_f \longrightarrow \mathbb{K}$ is continuous in ζ. In consequence, if f and g are both continuous, then the composition $g \circ f$ is also continuous.*

Proof. Let $\zeta \in D_f$ and assume f is continuous in ζ and g is continuous in $f(\zeta)$. If $(z_n)_{n \in \mathbb{N}}$ is a sequence in D_f such that $\lim_{n \to \infty} z_n = \zeta$, then the continuity of f in ζ implies that $\lim_{n \to \infty} f(z_n) = f(\zeta)$. Then the continuity of g in $f(\zeta)$ implies $\lim_{n \to \infty} g(f(z_n)) = g(f(\zeta))$, thereby establishing the continuity of $g \circ f$ in ζ. ∎

7.2.3   Bounded, Closed, and Compact Sets

Subsets A of $\mathbb{C}$ (and even subsets of $\mathbb{R}$) can be extremely complicated. If the set A has one or more of the benign properties defined in the following, then this can often be exploited in some useful way (we will see an important example in Th. 7.54 below).

Definition 7.42. Consider $A \subseteq \mathbb{C}$.

(a) A is called *bounded* if, and only if, $A = \emptyset$ or the set $\{|z| : z \in A\}$ is bounded in $\mathbb{R}$ in the sense of Def. 2.27(a), i.e. if, and only if,

$$\underset{M \in \mathbb{R}^+}{\exists} \quad A \subseteq B_M(0).$$

(b) A is called *closed* if, and only if, every sequence in A that converges in $\mathbb{C}$ has its limit in A (note that $\emptyset$ is, thus, closed).

(c) A is called *compact* if, and only if, A is both closed and bounded.

Example 7.43. (a) Clearly, $\emptyset$ and sets containing single points $\{z\}$, $z \in \mathbb{C}$ are compact. The sets $\mathbb{C}$ and $\mathbb{R}$ are simple examples of closed sets that are not bounded.

(b) Let $a, b \in \mathbb{R}$, $a < b$. Each bounded interval $]a, b[$, $]a, b]$, $[a, b[$, $[a, b]$ is, indeed, bounded (by $M := \max\{|a|, |b|\}$). If $(x_n)_{n \in \mathbb{N}}$ is a sequence in $[a, b]$, converging to $x \in \mathbb{R}$, then Th. 7.13(c) shows $a \le x \le b$, i.e. $x \in [a, b]$ and $[a, b]$ is, indeed, closed. Analogously, one sees that the unbounded intervals $[a, \infty[$ and $] - \infty, a]$ are also closed. On the other hand, open and half-open intervals are *not* closed: For sufficiently large n, the convergent sequence $(b - \frac{1}{n})_{n \in \mathbb{N}}$ is in $[a, b[$, but $\lim_{n \to \infty}(b - \frac{1}{n}) = b \notin [a, b[$, and the other cases are treated analogously. In particular, only intervals of the form $[a, b]$ (and trivial intervals) are compact.

(c) For each $\epsilon > 0$ and each $z \in \mathbb{C}$, the set $B_\epsilon(z)$ is bounded (since $B_\epsilon(z) \subseteq B_{\epsilon+|z|}(0)$ by the triangle inequality), but not closed (since, for sufficiently large $n \in \mathbb{N}$, $(z + \epsilon - \frac{1}{n})_{n \in \mathbb{N}}$ is a sequence in $B_\epsilon(z)$, converging to $z + \epsilon \notin B_\epsilon(z)$). In particular, $B_\epsilon(z)$ is not compact.

Proposition 7.44. (a) *Finite unions of bounded (resp. closed, resp. compact) sets are bounded (resp. closed, resp. compact), i.e. if $A_1, \ldots, A_n \subseteq \mathbb{C}$, $n \in \mathbb{N}$, are bounded (resp. closed, resp. compact), then $A := \bigcup_{j=1}^{n} A_j$ is also bounded (resp. closed, resp. compact).*

(b) *Arbitrary (i.e. finite or infinite) intersections of bounded (resp. closed, resp. compact) sets are bounded (resp. closed, resp. compact), i.e. if $I \neq \emptyset$ is an arbitrary index set and, for each $j \in I$, $A_j \subseteq \mathbb{C}$ is bounded (resp. closed, resp. compact), then $A := \bigcap_{j \in I} A_j$ is also bounded (resp. closed, resp. compact).*

Proof. (a): Exercise.

(b): Fix $j_0 \in I$. If all A_j, $j \in I$, are bounded, then, in particular, there is $M \in \mathbb{R}_0^+$ such that $A_{j_0} \subseteq B_M(0)$. Thus, $A = \bigcap_{j \in I} A_j \subseteq A_{j_0} \subseteq B_M(0)$ shows A is also bounded. If all A_j, $j \in I$, are closed and $(a_n)_{n \in \mathbb{N}}$ is a sequence in A that converges to some $z \in \mathbb{C}$, then $(a_n)_{n \in \mathbb{N}}$ is a sequence in each A_j, $j \in I$, and, since each A_j is closed, $z \in A_j$ for each $j \in I$, i.e. $z \in A = \bigcap_{j \in I} A_j$. If all A_j, $j \in I$, are compact, then they are all closed and bounded and, thus, A is closed and bounded, i.e. A is compact. $\blacksquare$

Example 7.45. (a) According to Prop. 7.44(a), all finite subsets of $\mathbb{C}$ are compact.

(b) $\mathbb{N} = \bigcup_{n \in \mathbb{N}} \{n\}$ shows that infinite unions of compact sets can be unbounded, and $]0, 1[= \bigcup_{n \in \mathbb{N}} [\frac{1}{n}, 1 - \frac{1}{n}]$ shows that infinite unions of compact sets are not always closed.

Many more examples of closed sets can be obtained as preimages of closed sets under continuous maps according to the following remark:

Remark 7.46. In Analysis II, it will be shown in the more general context of maps f between metric spaces that a map f is continuous if, and only if, all preimages $f^{-1}(A)$ under f of closed sets A are closed. Here, we will only prove the following special case:

$$f : \mathbb{C} \longrightarrow \mathbb{K} \text{ continuous and } A \subseteq \mathbb{K} \text{ closed} \quad \Rightarrow \quad f^{-1}(A) \subseteq \mathbb{C} \text{ closed.} \qquad (7.38)$$

Indeed, suppose f is continuous and $A \subseteq \mathbb{K}$ is closed. If $(z_n)_{n \in \mathbb{N}}$ is a sequence in $f^{-1}(A)$ with $\lim_{n \to \infty} z_n = z \in \mathbb{C}$, then $(f(z_n))_{n \in \mathbb{N}}$ is a sequence in A. The continuity of f then implies $\lim_{n \to \infty} f(z_n) = f(z)$ and, then, $f(z) \in A$, since A is closed. Thus, $z \in f^{-1}(A)$, showing $f^{-1}(A)$ is closed.

Example 7.47. (a) For each $z \in \mathbb{C}$ and each $r > 0$, the *closed disk* $\overline{B}_r(z) := \{w \in \mathbb{C} : |z - w| \leq r\}$ with radius r and center z is, indeed, closed by (7.38), since

$$\overline{B}_r(z) = f^{-1}[0, r], \qquad (7.39)$$

where f is the continuous map $f : \mathbb{C} \longrightarrow \mathbb{R}$, $f(w) := |z - w|$. Since $\overline{B}_r(z)$ is clearly bounded, it is also compact.

(b) For each $z \in \mathbb{C}$ and each $r > 0$, the *circle* (also called a 1-*sphere*) $S_r(z) := \{w \in \mathbb{C} : |z - w| = r\}$ with radius r and center z is closed by (7.38), since $S_r(z) = f^{-1}\{r\}$, where f is the same map as in (7.39). Moreover, $S_r(z)$ is also clearly bounded, and, thus, compact.

(c) According to (7.38), for each $x \in \mathbb{R}$, the closed *half-spaces* $\{z \in \mathbb{C} : \operatorname{Re} z \geq x\} = \operatorname{Re}^{-1}[x, \infty[$ and $\{z \in \mathbb{C} : \operatorname{Im} z \geq x\} = \operatorname{Im}^{-1}[x, \infty[$ are, indeed, closed.

Theorem 7.48. *A subset K of $\mathbb{C}$ is compact if, and only if, every sequence in K has a subsequence that converges to some limit $z \in K$.*

Proof. If K is closed and bounded, and $(z_n)_{n \in \mathbb{N}}$ is a sequence in K, then the boundedness, the Bolzano-Weierstrass Th. 7.27, and Prop. 7.26 yield a subsequence that converges to some $z \in \mathbb{C}$. However, since K is closed, $z \in K$.

Conversely, assume every sequence in K has a subsequence that converges to some limit $z \in K$. Let $(z_n)_{n \in \mathbb{N}}$ be a sequence in K that converges to some $w \in \mathbb{C}$. Then this sequence must have a subsequence that converges to some $z \in K$. However, according to Prop. 7.23, it must be $w = z \in K$, showing K is closed. If K is not bounded, then there exists a sequence $(z_n)_{n \in \mathbb{N}}$ in K such that $\lim_{n \to \infty} |z_n| = \infty$. Every subsequence $(z_{n_k})_{k \in \mathbb{N}}$ then still has the property that $\lim_{k \to \infty} |z_{n_k}| = \infty$, in particular, each subsequence is unbounded and can not converge to some $z \in \mathbb{C}$ (let alone in K). ∎

Caveat 7.49. In Analysis II, we will generalize the notion of compactness to subsets of so-called metric spaces. In metric spaces, it is still true that a set K is compact if, and only if, every sequence in K has a subsequence that converges to some limit in K. However, while it remains true that every compact set is closed and bounded, the converse does *not(!)* hold in general metric spaces (in general, even in closed sets, there exist bounded sequences that do not have convergent subsequences).

—

One reason that compact sets are useful is that real-valued continuous functions on compact sets assume a maximum and a minimum, which is the contents of Th. 7.54 below. In preparation, we now define maxima and minima for real-valued functions.

Definition 7.50. Let $M \subseteq \mathbb{C}$, $f : M \longrightarrow \mathbb{R}$.

(a) Given $z \in M$, f has a *(strict) global min* at z if, and only if, $f(z) \leq f(w)$ $(f(z) < f(w))$ for each $w \in M \setminus \{z\}$. Analogously, f has a *(strict) global max* at z if, and only if, $f(z) \geq f(w)$ $(f(z) > f(w))$ for each $w \in M \setminus \{z\}$. Moreover, f has a *(strict) global extreme value* at z if, and only if, f has a (strict) global min or a (strict) global max at z.

(b) Given $z \in M$, f has a *(strict) local min* at z if, and only if, there exists $\epsilon > 0$ such that $f(z) \leq f(w)$ $(f(z) < f(w))$ for each $w \in \{w \in M : |z - w| < \epsilon\} \setminus \{z\}$. Analogously, f has a *(strict) local max* at z if, and only if, there exists $\epsilon > 0$ such that $f(z) \geq f(w)$ $(f(z) > f(w))$ for each $w \in \{w \in M : |z - w| < \epsilon\} \setminus \{z\}$.

Moreover, f has a *(strict) local extreme value* at z if, and only if, f has a (strict) local min or a (strict) local max at z.

Remark 7.51. In the context of Def. 7.50, it is immediate from the respective definitions that f has a (strict) global min at $z \in M$ if, and only if, $-f$ has a (strict) global max at z. Moreover, the same holds if "global" is replaced by "local". It is equally obvious that every (strict) global min/max is a (strict) local min/max.

Theorem 7.52. *If $K \subseteq \mathbb{C}$ is compact, and $f : K \longrightarrow \mathbb{C}$ is continuous, then $f(K)$ is compact.*

Proof. If $(w_n)_{n \in \mathbb{N}}$ is a sequence in $f(K)$, then, for each $n \in \mathbb{N}$, there is some $z_n \in K$ such that $f(z_n) = w_n$. As K is compact, there is a subsequence $(a_n)_{n \in \mathbb{N}}$ of $(z_n)_{n \in \mathbb{N}}$ with $\lim_{n \to \infty} a_n = a$ for some $a \in K$. Then $(f(a_n))_{n \in \mathbb{N}}$ is a subsequence of $(w_n)_{n \in \mathbb{N}}$ and the continuity of f yields $\lim_{n \to \infty} f(a_n) = f(a) \in f(K)$, showing that $(w_n)_{n \in \mathbb{N}}$ has a convergent subsequence with limit in $f(K)$. By Th. 7.48, we have therefore established that $f(K)$ is compact. ∎

Lemma 7.53. *If K is a nonempty compact subset of $\mathbb{R}$, then K contains a smallest and a largest element, i.e. there exist $m, M \in K$ such that $m \leq x \leq M$ for each $x \in K$.*

Proof. Since the compact set K is bounded, we know that

$$-\infty < m := \inf K \leq \sup K =: M < \infty.$$

According to the definition of the inf and sup as largest lower bound and smallest upper bound, respectively, for each $n \in \mathbb{N}$, there must be elements $x_n, y_n \in K$ such that $m \leq x_n \leq m + \frac{1}{n}$ and $M - \frac{1}{n} \leq y_n \leq M$. Since the compact set K is also closed, we get $m = \lim_{n \to \infty} x_n \in K$ and $M = \lim_{n \to \infty} y_n \in K$. ∎

Theorem 7.54. *If $K \subseteq \mathbb{C}$ is compact, and $f : K \longrightarrow \mathbb{R}$ is continuous, then f assumes its max and its min, i.e. there are $z_m \in K$ and $z_M \in K$ such that f has a global min at z_m and a global max at z_M. In particular, the continuous function f assumes its max and min on each compact interval $K = [a, b] \subseteq \mathbb{R}$, $a, b \in \mathbb{R}$.*

Proof. Since K is compact and f is continuous, $f(K) \subseteq \mathbb{R}$ is compact according to Th. 7.52. Then, by Lem. 7.53, $f(K)$ contains a smallest element m and a largest element M. This, in turn, implies that there are $z_m, z_M \in K$ such that $f(z_m) = m$ and $f(z_M) = M$. ∎

Example 7.55. On an unbounded set, a continuous function does not necessarily have a global max or a global min, as one can already see from $x \mapsto x$. An example for a continuous function on a bounded, but not closed, interval, that does not have a global max is $f :]0, 1] \longrightarrow \mathbb{R}$, $f(x) := 1/x$, which is continuous by Th. 7.38.

7.2.4 Intermediate Value Theorem

Theorem 7.56 (Bolzano's Theorem). *Let $a, b \in \mathbb{R}$ with $a < b$. If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous with $f(a) > 0$ and $f(b) < 0$, then f has at least one zero in $]a, b[$. More precisely, the set $A := f^{-1}\{0\}$ has a min ξ_1 and a max ξ_2, $a < \xi_1 \leq \xi_2 < b$, where $f > 0$ on $[a, \xi_1[$ and $f < 0$ on $]\xi_2, b]$.*

Proof. Let $\xi_1 := \inf f^{-1}(\mathbb{R}_0^-)$.

(a): $f(\xi_1) \leq 0$: This is clear if $\xi_1 = b$. If $\xi_1 < b$, then, for each $n \in \mathbb{N}$ sufficiently large, there exists $x_n \in]\xi_1, \xi_1 + 1/n[\subseteq [a, b]$ such that $f(x_n) \leq 0$). Then $\lim_{n\to\infty} x_n = \xi_1$ and the continuity of f implies $\lim_{n\to\infty} f(x_n) = f(\xi_1)$. Now $f(\xi_1) \leq 0$ is a consequence of Th. 7.13(c). In particular, (a) yields $a < \xi_1$ and $f > 0$ on $[a, \xi_1[$.

(b): $f(\xi_1) \geq 0$: The continuity of f implies $\lim_{n\to\infty} f(\xi_1 - 1/n) = f(\xi_1)$ and, since we have already seen $f(\xi_1 - 1/n) > 0$ for each $n \in \mathbb{N}$ sufficiently large, $f(\xi_1) \geq 0$ is again a consequence of Th. 7.13(c). In particular, we have $\xi_1 < b$.

Combining (a) and (b), we have $f(\xi_1) = 0$ and $a < \xi_1 < b$.

Defining $\xi_2 := \sup f^{-1}(\mathbb{R}_0^+)$, $f(\xi_2) = 0$ and $a < \xi_2 < b$ is shown completely analogous. Then $f < 0$ on $]\xi_2, b]$ is also clear as well as $\xi_1 \leq \xi_2$. ∎

Theorem 7.57 (Intermediate Value Theorem). *Let $a, b \in \mathbb{R}$ with $a < b$. If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous, then f assumes every value between $f(a)$ and $f(b)$, i.e.*

$$\left[\min\{f(a), f(b)\}, \max\{f(a), f(b)\} \right] \subseteq f\big([a, b]\big). \tag{7.40}$$

Proof. If $f(a) = f(b)$, then there is nothing to prove. If $f(a) < f(b)$ and $\eta \in]f(a), f(b)[$, then consider the auxiliary function $g : [a, b] \longrightarrow \mathbb{R}$, $g(x) := \eta - f(x)$. Then g is continuous with $g(a) = \eta - f(a) > 0$ and $g(b) = \eta - f(b) < 0$. According to Bolzano's Th. 7.56, there exists $\xi \in]a, b[$ such that $g(\xi) = \eta - f(\xi) = 0$, i.e. $f(\xi) = \eta$ as claimed. If $f(b) < f(a)$ and $\eta \in]f(b), f(a)[$, then consider the auxiliary function $g : [a, b] \longrightarrow \mathbb{R}$, $g(x) := f(x) - \eta$. Then g is continuous with $g(a) = f(a) - \eta > 0$ and $g(b) = f(b) - \eta < 0$. Once again, according to Bolzano's Th. 7.56, there exists $\xi \in]a, b[$ such that $g(\xi) = f(\xi) - \eta = 0$, i.e. $f(\xi) = \eta$. ∎

Theorem 7.58. *If $I \subseteq \mathbb{R}$ is an interval (of one of the 8 types listed in (4.8)) and $f : I \longrightarrow \mathbb{R}$ is continuous, then $f(I)$ is also an interval (it can degenerate to a single point if f is constant). More precisely, if $\emptyset \neq I = [a, b]$ is a compact interval, then $\emptyset \neq f(I) = [\min f(I), \max f(I)]$; if I is not a compact interval, then one of the following 9 cases occurs:*

$$f(I) = \mathbb{R}, \tag{7.41a}$$

$$f(I) =]-\infty, \sup f(I)], \tag{7.41b}$$

$$f(I) =]-\infty, \sup f(I)[, \tag{7.41c}$$

$$f(I) = [\inf f(I), \infty[\tag{7.41d}$$

$$f(I) = [\inf f(I), \sup f(I)], \tag{7.41e}$$

$$f(I) = [\inf f(I), \sup f(I)[, \tag{7.41f}$$
$$f(I) =] \inf f(I), \infty[, \tag{7.41g}$$
$$f(I) =] \inf f(I), \sup f(I)], \tag{7.41h}$$
$$f(I) =] \inf f(I), \sup f(I)[. \tag{7.41i}$$

Proof. If I is a compact interval, then we merely combine Th. 7.54 with Th. 7.57. Otherwise, let $\eta \in f(I)$. If $f(I)$ has an upper bound, then Th. 7.57 implies $[\eta, \sup f(I)[\subseteq f(I)$ and $f(I) \cap [\eta, \infty[\subseteq [\eta, \sup f(I)]$. If $f(I)$ does not have an upper bound, then Th. 7.57 implies $f(I) \cap [\eta, \infty[= [\eta, \infty[$. Analogously, one obtains $f(I) \cap] - \infty, \eta] =] - \infty, \eta]$ or $f(I) \cap] - \infty, \eta] = [\inf f(I), \eta]$ or $f(I) \cap] - \infty, \eta] =] \inf f(I), \eta]$, showing that there are precisely the 9 possibilities of (7.41) for $f(I) = \big(f(I) \cap] - \infty, \eta]\big) \cup \big(f(I) \cap [\eta, \infty[\big)$. ∎

The above results will have striking consequences in the following Sec. 7.2.5.

Example 7.59. The piecewise affine function

$$f :]0,1] \longrightarrow \mathbb{R}, \quad f(x) := \begin{cases} (-1)^n \cdot n - \frac{2n+1}{\frac{1}{n-1} - \frac{1}{n}} \left(x - \frac{1}{n}\right) & \text{for } x \in [\frac{1}{n}, \frac{1}{n-1}], \; n \text{ even}, \\ (-1)^n \cdot n + \frac{2n+1}{\frac{1}{n-1} - \frac{1}{n}} \left(x - \frac{1}{n}\right) & \text{for } x \in [\frac{1}{n}, \frac{1}{n-1}], \; n \geq 3 \text{ odd}, \end{cases}$$

satisfies $f(1/n) = (-1)^n \cdot n$ for each $n \in \mathbb{N}$ and is an example of a continuous function on the bounded half-open interval $I :=]0,1]$ with $f(I) = \mathbb{R}$.

7.2.5 Inverse Functions, Existence of Roots, Exponential Function, Logarithm

Theorem 7.60. *Let $I \subseteq \mathbb{R}$ be an interval (of one of the 8 types listed in (4.8)). If $f : I \longrightarrow \mathbb{R}$ is strictly increasing (resp. decreasing), then f has an inverse function f^{-1} defined on $J := f(I)$, i.e. $f^{-1} : J \longrightarrow I$, and f^{-1} is continuous and strictly increasing (resp. decreasing). If f is also continuous, then J must be an interval.*

Proof. From Prop. 2.32(b), we know $f : I \longrightarrow \mathbb{R}$ is one-to-one. Then $f : I \longrightarrow f(I)$ is invertible and Prop. 2.32(c) shows f^{-1} is strictly monotone in the same sense as f. We need to prove the continuity of f^{-1}. We assume f to be strictly increasing (the case where f is strictly decreasing then follows by considering $-f$). Let $\eta \in J$, $\epsilon > 0$, and $\xi \in I$ with $f(\xi) = \eta$. If $I = \{\xi\}$, then $J = \{\eta\}$, and there is nothing to prove. It remains to consider three cases:
(a) $\xi = \min I$, i.e. ξ is the left endpoint of I (and $\xi \neq \max I$),
(b) $\xi = \max I$, i.e. ξ is the right endpoint of I (and $\xi \neq \min I$),
(c) ξ is neither the min nor the max of I.
We carry out the proof for (c) and leave the (very similar) cases (a) and (b) as exercises. In case (c), there are points $\xi_1, \xi_2 \in B_\epsilon(\xi) \cap I$ such that

$$\xi - \epsilon < \xi_1 < \xi < \xi_2 < \xi + \epsilon. \tag{7.42}$$

As f is strictly increasing, this implies

$$f(\xi_1) < \eta < f(\xi_2).$$

Choose $\delta > 0$ such that

$$f(\xi_1) < \eta - \delta < \eta < \eta + \delta < f(\xi_2).$$

Then

$$\underset{y \in J \cap B_\delta(\eta)}{\forall} \left(f(\xi_1) < y < f(\xi_2) \overset{f^{-1} \text{ str. inc.}}{\Rightarrow} \xi_1 < f^{-1}(y) < \xi_2 \overset{(7.42)}{\Rightarrow} f^{-1}(y) \in B_\epsilon(\xi) \right),$$

proving the continuity of f^{-1} in η. That J must be an interval if f is continuous was already shown in Th. 7.58. ∎

Remark and Definition 7.61 (Roots)**.** We are now in a position to fulfill the promise made in Def. and Rem. 5.6, i.e. to prove the existence of unique roots for nonnegative real numbers: For each $n \in \mathbb{N}$, the function $f : \mathbb{R}_0^+ \longrightarrow \mathbb{R}$, $f(x) := x^n$, is continuous and strictly increasing with $J := f(\mathbb{R}_0^+) = \mathbb{R}_0^+$. Then Th. 7.60 implies the existence of a continuous and strictly increasing inverse function $f^{-1} : \mathbb{R}_0^+ \longrightarrow \mathbb{R}_0^+$. For each $x \in \mathbb{R}_0^+$, we call $f^{-1}(x)$ the nth *root* of x and write $\sqrt[n]{x} := x^{\frac{1}{n}} := f^{-1}(x)$. Then $(\sqrt[n]{x})^n = (x^{\frac{1}{n}})^n = x$ is immediate from the definition. Caveat: By definition, roots are always *nonnegative* and they are only defined for *nonnegative* numbers (when studying complex numbers and $\mathbb{C}$-valued functions more deeply in the field of Complex Analysis, one typically extends the notion of root, but we will not pursue this route in this class). As anticipated in Def. and Rem. 5.6, one also writes $\sqrt{x}$ instead of $\sqrt[2]{x}$ and calls $\sqrt{x}$ the *square root* of x.

Remark and Definition 7.62. It turns out that $\sqrt{2}$ (and many other roots) are not rational numbers, i.e. $\sqrt{2} \notin \mathbb{Q}$. This is easily proved by contradiction: If $\sqrt{2} \in \mathbb{Q}$, then there exist natural numbers $m, n \in \mathbb{N}$ such that $\sqrt{2} = m/n$. Moreover, by canceling possible factors of 2, we may assume at least one of the numbers m, n is odd. Now $\sqrt{2} = m/n$ implies $m^2 = 2n^2$, i.e. m^2 and, thus, m must be even. In consequence, there exists $p \in \mathbb{N}$ such that $m = 2p$, implying $2n^2 = m^2 = 4p^2$ and $n^2 = 2p^2$. Thus n^2 and n must also be even, in contradiction to m, n not both being even.

The elements of $\mathbb{R} \setminus \mathbb{Q}$ are called *irrational* numbers. It turns out that most real numbers are irrational numbers – one can show that $\mathbb{Q}$ is countable, whereas $\mathbb{R} \setminus \mathbb{Q}$ is not countable (actually, every interval contains countably many rational and uncountably many irrational numbers, see Appendix F, in particular, Th. F.1(c) and Cor. F.4).

Theorem 7.63 (Inequality Between the Arithmetic Mean and the Geometric Mean)**.** *If $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathbb{R}_0^+$, then*

$$\sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \cdots + x_n}{n}, \tag{7.43}$$

where the left-hand side is called the geometric mean *and the right-hand side is called the* arithmetic mean *of the numbers $x_1, \ldots, x_n$. Equality occurs if, and only if, $x_1 = \cdots = x_n$.*

Proof. If at least one of the x_j is 0, then (7.43) becomes the true statement $0 \leq \frac{x_1 + \cdots + x_n}{n}$ with strict equality if at least one $x_j > 0$. If $x_1 = \cdots = x_n = x$, then (7.43) also holds since both sides are equal to x. Thus, for the remainder of the proof, we assume all $x_j > 0$ and not all x_j are equal. First, we consider the special case, where $\frac{x_1 + \cdots + x_n}{n} = 1$. Since not all x_j are equal, there exists k with $x_k \neq 1$. We prove (7.43) by induction for $n \in \{2, 3, \dots\}$ in the form

$$\left(\sum_{j=1}^{n} x_j = n \ \wedge \ \underset{k \in \{1, \dots, n\}}{\exists} \ x_k \neq 1 \right) \ \Rightarrow \ \prod_{j=1}^{n} x_j < 1.$$

Base Case ($n = 2$): Since $x_1 + x_2 = 2$, $0 < x_1, x_2$ and not both x_1 and x_2 are equal to 1, there is $\epsilon > 0$ such that $x_1 = 1 + \epsilon$ and $x_2 = 1 - \epsilon$, i.e. $x_1 x_2 = 1 - \epsilon^2 < 1$, which establishes the base case. Induction Step: We now have $n \geq 2$ and $0 < x_1, \dots, x_{n+1}$ with $\sum_{j=1}^{n+1} x_j = n + 1$ plus the existence of $k, l \in \{1, \dots, n+1\}$ such that $x_k = 1 + \alpha$, $x_l = 1 - \beta$ with $\alpha, \beta > 0$. Then define $y := x_k + x_l - 1 = 1 + \alpha - \beta$. One observes $y > 0$ (since $\beta < 1$) and

$$y + \sum_{\substack{j=1, \\ j \neq k, l}}^{n+1} x_j = -1 + \sum_{j=1}^{n+1} x_j = n \ \overset{\text{ind. hyp.}}{\Rightarrow} \ y \prod_{\substack{j=1, \\ j \neq k, l}}^{n+1} x_j \leq 1$$

(we can not exclude equality as y and all the remaining x_j might be equal to 1). Since $x_k x_l = (1 + \alpha)(1 - \beta) = 1 + \alpha - \beta - \alpha\beta = y - \alpha\beta < y$, we now infer $\prod_{j=1}^{n+1} x_j < 1$, concluding the induction proof. It remains to consider the case $\frac{x_1 + \cdots + x_n}{n} = \lambda > 0$, not all x_j equal. One estimates

$$\sqrt[n]{x_1 \cdots x_n} = \lambda \sqrt[n]{\frac{x_1}{\lambda} \cdots \frac{x_n}{\lambda}} \ \overset{\text{special case}}{<} \ \lambda \frac{x_1 + \cdots + x_n}{\lambda n} = \frac{x_1 + \cdots + x_n}{n},$$

completing the proof of the theorem. ∎

Corollary 7.64. *For each $a \in \mathbb{R}_0^+ \setminus \{1\}$, $n \in \{2, 3, \dots\}$, $p \in \{1, \dots, n-1\}$:*

$$\sqrt[n]{a^p} < 1 + \frac{p}{n}(a - 1); \quad p = 1 \ \textit{yields} \quad \sqrt[n]{a} < 1 + \frac{a - 1}{n}. \tag{7.44}$$

Proof. The simple application

$$\sqrt[n]{a^p} = \sqrt[n]{a^p \cdot \prod_{j=1}^{n-p} 1} \ \overset{\text{Th. 7.63}}{<} \ \frac{p a + n - p}{n} = 1 + \frac{p}{n}(a - 1) \tag{7.45}$$

of Th. 7.63 establishes the case. ∎

Example 7.65. We use (7.43) to show

$$\lim_{n \to \infty} \sqrt[n]{n} = 1 : \tag{7.46}$$

First note $0 < x < 1 \Rightarrow 0 < x^n < 1$, i.e. $\sqrt[n]{n} > 1$ for each $(\sqrt[n]{n})^n = n > 1$. Now write n as the product of n factors $n = \sqrt{n}\sqrt{n} \cdot \prod_{k=1}^{n-2} 1$. Then, for $n > 1$,

$$\sqrt[n]{n} = \sqrt[n]{\sqrt{n}\sqrt{n} \cdot \prod_{k=1}^{n-2} 1} \overset{\text{Th. 7.63}}{<} \frac{2\sqrt{n} + n - 2}{n} < 1 + \frac{2}{\sqrt{n}}. \tag{7.47}$$

It is an exercise to show

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0. \tag{7.48}$$

Now this together with $1 \leq \sqrt[n]{n} \leq 1 + \frac{2}{\sqrt{n}}$ and the Sandwich Th. 7.16 proves (7.46).

Example 7.66 (Euler's Number). We use Th. 7.63 to prove the limit

$$e := \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n \tag{7.49}$$

exists. It is known as *Euler's number*. One can show it is an irrational number (see Appendix H.1) and its first digits are $e = 2.71828\ldots$ It is of exceptional importance for analysis and mathematics in general, as it pops up in all kinds of different mathematical contexts. From Th. 7.63, we obtain

$$\underset{n\in\mathbb{N}}{\forall} \underset{\substack{x\in[-n,\infty[, \\ x\neq 0}}{\forall} \left(1 + \frac{x}{n}\right)^n = 1 \cdot \left(1 + \frac{x}{n}\right)^n < \left(1 + \frac{x}{n+1}\right)^{n+1}, \tag{7.50}$$

where we have used that, on both sides of the inequality in (7.50), there are $n+1$ factors having the same sum, namely $n + 1 + x$; and the inequality in (7.43) is strict, unless all factors are equal. We now apply (7.50) to the sequences $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, $(c_n)_{n\in\mathbb{N}}$, where

$$\underset{n\in\mathbb{N}}{\forall} \left(\begin{array}{c} a_n := \left(1 + \dfrac{1}{n}\right)^n, \quad b_n := \left(1 - \dfrac{1}{n}\right)^n, \\[2ex] c_n := b_{n+1}^{-1} = \left(\left(1 - \dfrac{1}{n+1}\right)^{-1}\right)^{n+1} = \left(1 + \dfrac{1}{n}\right)^{n+1} \end{array} \right) : \tag{7.51}$$

Applying (7.50) with $x = 1$ and $x = -1$, respectively, yields $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are strictly increasing, and $(c_n)_{n\in\mathbb{N}}$ is strictly decreasing. On the other hand, $a_n < c_n$ holds for each $n \in \mathbb{N}$, showing $(a_n)_{n\in\mathbb{N}}$ is bounded from above by c_1, and $(c_n)_{n\in\mathbb{N}}$ is bounded from below by a_1. In particular, Th. 7.19 implies the convergence of both $(a_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$. Moreover, $\lim_{n\to\infty} c_n = \lim_{n\to\infty} \left(a_n(1+1/n)\right) = e \cdot 1 = e$, which, together with $a_n < e < c_n$ for each $n \in \mathbb{N}$, can be used to compute e to an arbitrary precision.

Definition 7.67. Let $A \subseteq \mathbb{R}$ be a subset of the real numbers. Then A is called *dense* in $\mathbb{R}$ if, and only if, every ϵ-neighborhood of every real number contains a point from A, i.e. if, and only if,

$$\underset{x\in\mathbb{R}}{\forall} \underset{\epsilon\in\mathbb{R}^+}{\forall} A \cap B_\epsilon(x) \neq \emptyset.$$

Theorem 7.68. (a) $\mathbb{Q}$ *is dense in* $\mathbb{R}$.

(b) $\mathbb{R} \setminus \mathbb{Q}$ *is dense in* $\mathbb{R}$.

(c) *For each* $x \in \mathbb{R}$, *there exist sequences* $(r_n)_{n \in \mathbb{N}}$ *and* $(s_n)_{n \in \mathbb{N}}$ *in the rational numbers* $\mathbb{Q}$ *such that* $x = \lim_{n \to \infty} r_n = \lim_{n \to \infty} s_n$, $(r_n)_{n \in \mathbb{N}}$ *is strictly increasing and* $(s_n)_{n \in \mathbb{N}}$ *is strictly decreasing.*

Proof. (a): Since each $B_\epsilon(x)$ is an interval, it suffices to prove that every interval $]a, b[$, $a < b$, contains a rational number. If $0 \in]a, b[$, then there is nothing to prove. Suppose $0 < a < b$ and set $\delta := b - a > 0$. Choose $n \in \mathbb{N}$ such that $1/n < \delta$ and let

$$q := \max \left\{ \frac{k}{n} : k \in \mathbb{N} \wedge \frac{k}{n} < b \right\}.$$

Then $q \in \mathbb{Q}$ and $a < q < b$. If $a < b < 0$, choose δ and n as above, but let

$$q := \min \left\{ -\frac{k}{n} : k \in \mathbb{N} \wedge -\frac{k}{n} > a \right\}.$$

Then, once again, $q \in \mathbb{Q}$ and $a < q < b$.

(b): Analogous to (a), we show that every interval $]a, b[$, $a < b$, contains an irrational number: According to (a), we choose $q \in \mathbb{Q} \cap]a, b[$, $\delta := b - q > 0$ and $n \in \mathbb{N}$ such that $\sqrt{2}/n < \delta$. Then $a < \lambda := q + \sqrt{2}/n < b$ and also $\lambda \in \mathbb{R} \setminus \mathbb{Q}$ (otherwise, $\sqrt{2} = n(\lambda - q) \in \mathbb{Q}$).

(c): Using (a), for each $n \in \mathbb{N}$, we choose rational numbers r_n and s_n such that

$$r_n \in \left] x - \frac{1}{n}, x - \frac{1}{n+1} \right[, \quad s_n \in \left] x + \frac{1}{n+1}, x + \frac{1}{n} \right[.$$

Then, clearly, $(r_n)_{n \in \mathbb{N}}$ is strictly increasing, $(s_n)_{n \in \mathbb{N}}$ is strictly decreasing, and the Sandwich Th. 7.16 implies $x = \lim_{n \to \infty} r_n = \lim_{n \to \infty} s_n$. ∎

Definition and Remark 7.69 (Exponentiation). We have already used that Not. C.5 of the Appendix defines a^x for $(a, x) \in \mathbb{C} \times \mathbb{N}_0$ and for $(a, x) \in (\mathbb{C} \setminus \{0\}) \times \mathbb{Z}$ (using $G = \mathbb{C}$ and $G = \mathbb{C} \setminus \{0\}$, respectively). We will now extend the definition to $(a, x) \in \mathbb{R}^+ \times \mathbb{R}$ (later, we will further extend the definition to $(a, z) \in \mathbb{R}^+ \times \mathbb{C}$). The present extension to $(a, x) \in \mathbb{R}^+ \times \mathbb{R}$ is accomplished in two steps – first, in (a), for rational x, then, in (b), for irrational x.

(a) For rational $x = k/n$ with $k \in \mathbb{Z}$ and $n \in \mathbb{N}$, define

$$a^x := a^{\frac{k}{n}} := \sqrt[n]{a^k}. \tag{7.52}$$

For this definition to make sense, we have to check it does not depend on the special representation of x, i.e., we have to verify $x = \frac{k}{n} = \frac{km}{nm}$ with $k \in \mathbb{Z}$ and $m, n \in \mathbb{N}$ implies $a^{\frac{k}{n}} = a^{\frac{km}{nm}}$. To this end, observe, using Rem. and Def. 7.61,

$$\left(a^{\frac{k}{n}} \right)^{nm} = \left(\sqrt[n]{a^k} \right)^{nm} = a^{km} \quad \text{and} \quad \left(a^{\frac{km}{nm}} \right)^{nm} = \left(\sqrt[nm]{a^{km}} \right)^{nm} = a^{km}, \tag{7.53}$$

proving $a^{\frac{k}{n}} = a^{\frac{km}{nm}}$ (here, as in Rem. and Def. 7.61, we used that $\lambda \mapsto \lambda^N$ is one-to-one on $\mathbb{R}_0^+$ for each $N \in \mathbb{N}$). The exponentiation rules of Th. C.6 now extend to rational exponents in a natural way, i.e., for each $a, b > 0$ and each $x, y \in \mathbb{Q}$:

$$a^{x+y} = a^x \, a^y, \tag{7.54a}$$

$$(a^x)^y = a^{x\,y}, \tag{7.54b}$$

$$a^x \, b^x = (ab)^x. \tag{7.54c}$$

For the proof, by possibly multiplying numerator and denominator by some natural number, we can assume $x = k/n$ and $y = l/n$ with $k, l \in \mathbb{Z}$ and $n \in \mathbb{N}$. Then

$$(a^{x+y})^n = (a^{\frac{k+l}{n}})^n = a^{k+l} \overset{\text{Th. C.6(a)}}{=} a^k \, a^l = (a^{\frac{k}{n}})^n \, (a^{\frac{l}{n}})^n \overset{\text{Th. C.6(c)}}{=} (a^x \, a^y)^n,$$

proving (7.54a);

$$((a^x)^y)^{n^2} \overset{\text{Th. C.6(b)}}{=} \left(((a^x)^{\frac{l}{n}})^n \right)^n = ((a^{\frac{k}{n}})^l)^n \overset{\text{Th. C.6(b)}}{=} ((a^{\frac{k}{n}})^n)^l \overset{\text{Th. C.6(b)}}{=} a^{kl}$$

$$= (a^{\frac{kl}{n^2}})^{n^2} = (a^{x\,y})^{n^2},$$

proving (7.54b);

$$(a^x \, b^x)^n \overset{\text{Th. C.6(c)}}{=} (a^{\frac{k}{n}})^n \, (b^{\frac{k}{n}})^n a^k \, b^k \overset{\text{Th. C.6(c)}}{=} (ab)^k = (ab)^{\frac{k}{n} \cdot n} \overset{\text{Th. C.6(b)}}{=} ((ab)^x)^n,$$

proving (7.54c).

Moreover, we obtain the following monotonicity rules for each $a, b \in \mathbb{R}^+$ and each $x, y \in \mathbb{Q}$:

$$\underset{x>0}{\forall} \left(a < b \quad \Rightarrow \quad a^x < b^x \right), \tag{7.55a}$$

$$\underset{x<0}{\forall} \left(a < b \quad \Rightarrow \quad a^x > b^x \right), \tag{7.55b}$$

$$\underset{a>1}{\forall} \left(x < y \quad \Rightarrow \quad a^x < a^y \right), \tag{7.55c}$$

$$\underset{0<a<1}{\forall} \left(x < y \quad \Rightarrow \quad a^x > a^y \right). \tag{7.55d}$$

If $x = k/n$ with $k, n \in \mathbb{N}$ and $a < b$, then $a^{1/n} < b^{1/n}$ according to Rem. and Def. 7.61, which, in turn, implies $a^x = (a^{1/n})^k < (b^{1/n})^k = b^x$, proving (7.55a); and $a^{-1} > b^{-1}$ implies $a^{-x} = (a^{-1})^x > (b^{-1})^x = b^{-x}$, proving (7.55b). If $x < y$, set $q := y - x > 0$. Then $1 < a$ and (7.55a) imply $1 = 1^q < a^q$, i.e. $a^x < a^x \, a^q = a^y$, proving (7.55c). Similarly, $0 < a < 1$ and (7.55a) imply $a^q < 1^q = 1$, i.e. $a^y = a^x \, a^q < a^x$, proving (7.55d).

The following estimates will also come in handy: For $a \in \mathbb{R}^+$ and $x, y \in \mathbb{Q}$:

$$a > 1 \wedge x > 0 \quad \Rightarrow \quad a^x - 1 < x \cdot a^{x+1}, \tag{7.56}$$

$$\underset{m\in\mathbb{N}}{\forall} \left(x, y \in [-m, m] \quad \Rightarrow \quad |a^x - a^y| \leq L \, |x - y|, \right.$$

$$\left. \text{where} \quad L := \max\{a^{m+1}, (1/a)^{m+1}\} \right). \tag{7.57}$$

For $x \geq 1$, (7.56) is proved by $a^x < a^{x+1} < x \cdot a^{x+1} + 1$; for $x < 1$, write $x = p/n$ with $p, n \in \mathbb{N}$ and $p < n$, and apply (7.44) to obtain $a^x < 1 + x(a-1) < 1 + xa < 1 + x \cdot a^{x+1}$. For the proof of (7.57), first consider $a > 1$. Moreover, by possibly renaming x and y, we may assume $x < y$, i.e. $z := y - x > 0$. Thus, (7.56) holds with x replaced by z. Multiplying the resulting inequality by a^x yields

$$a^x\, a^z - a^x = a^y - a^x < z \cdot a^x\, a^{z+1} = (y-x)\, a^{y+1} \leq (y-x)\, a^{m+1},$$

proving (7.57) for $a > 1$. For $a = 1$, it is clearly true, and for $a < 1$, it is $a^{-1} > 1$, i.e.

$$|a^x - a^y| = |(a^{-1})^{-x} - (a^{-1})^{-y}| \leq |y - x|\, (a^{-1})^{m+1},$$

finishing the proof of (7.57).

(b) We now define a^x for irrational x by letting

$$a^x := \lim_{n \to \infty} a^{q_n}, \quad \text{where } (q_n)_{n \in \mathbb{N}} \text{ is a sequence in } \mathbb{Q} \text{ with } \lim_{n \to \infty} q_n = x. \quad (7.58)$$

For this definition to make sense, we have to know such sequences $(q_n)_{n \in \mathbb{N}}$ exist, which we do know from Th. 7.68(c). We also know from Th. 7.68(c) that there exists an *increasing* sequence $(q_n)_{n \in \mathbb{N}}$ in $\mathbb{Q}$ converging to x, in particular, bounded by x. Then, by (7.55c) and (7.55d), respectively, $(a^{q_n})_{n \in \mathbb{N}}$ is increasing for $a > 1$ and decreasing for $0 < a < 1$. Moreover, the sequence is bounded from above by a^N with $N \in \mathbb{N}$, $N > x$, for $a > 1$; and bounded from below by 0 for $0 < a < 1$. In both cases, Th. 7.19 implies convergence of the sequence to some limit that we may call a^x. However, we still need to verify that, for each sequence $(r_n)_{n \in \mathbb{N}}$ in $\mathbb{Q}$ with $\lim_{n \to \infty} r_n = x$, the sequence $(a^{r_n})_{n \in \mathbb{N}}$ converges to the same limit a^x in $\mathbb{R}$. If $\lim_{n \to \infty} r_n = x$, then $\lim_{n \to \infty} |q_n - r_n| = 0$. Since $(r_n)_{n \in \mathbb{N}}$ and $(q_n)_{n \in \mathbb{N}}$ are bounded, (7.57) implies

$$\underset{L \in \mathbb{R}^+}{\exists}\ \underset{n \in \mathbb{N}}{\forall}\ |a^{q_n} - a^{r_n}| \leq L\, |q_n - r_n|, \quad (7.59)$$

such that Prop. 7.11(a) implies $\lim_{n \to \infty} |a^{q_n} - a^{r_n}| = 0$ and

$$\lim_{n \to \infty} a^{r_n} = \lim_{n \to \infty} (a^{r_n} - a^{q_n} + a^{q_n}) = 0 + a^x = a^x, \quad (7.60)$$

showing (7.58) does not depend on the chosen sequence.

Proposition 7.70. *The exponentiation rules (7.54), the monotonicity rules (7.55), and the estimates (7.56) and (7.57) remain valid if $x, y \in \mathbb{Q}$ is replaced by $x, y \in \mathbb{R}$. Moreover, for each $a > 0$ and each sequence $(x_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$:*

$$\lim_{n \to \infty} x_n = x \in \mathbb{R} \quad \Rightarrow \quad \lim_{n \to \infty} a^{x_n} = a^x. \quad (7.61)$$

Proof. Given $x, y \in \mathbb{R}$, let $(p_n)_{n \in \mathbb{N}}$ and $(q_n)_{n \in \mathbb{N}}$ be sequences in $\mathbb{Q}$ such that $\lim_{n \to \infty} p_n = x$ and $\lim_{n \to \infty} q_n = y$.

We start by verifying (7.57). As we can assume $(p_n)_{n \in \mathbb{N}}$ and $(q_n)_{n \in \mathbb{N}}$ to be monotone, we may also assume $p_n, q_n \in [-m, m]$ for each $n \in \mathbb{N}$. Then the rational case of (7.57) implies

$$\underset{n \in \mathbb{N}}{\forall}\ |a^{p_n} - a^{q_n}| \leq L\, |p_n - q_n|,$$

and Th. 7.13(c) establishes the case. Then (7.61) also follows, since

$$0 \le |a^{x_n} - a^x| \le L\,|x_n - x| \to 0.$$

We deal with (7.54) next. For each $a, b > 0$:

$$a^{x+y} = \lim_{n\to\infty} a^{p_n+q_n} \overset{(7.54\mathrm{a})}{=} \lim_{n\to\infty} \left(a^{p_n}\, a^{q_n} \right) = a^x\, a^y,$$

$$a^x\, b^x = \lim_{n\to\infty} a^{p_n} \lim_{n\to\infty} b^{p_n} = \lim_{n\to\infty} \left(a^{p_n}\, b^{p_n} \right) \overset{(7.54\mathrm{c})}{=} \lim_{n\to\infty} (ab)^{p_n} = (ab)^x,$$

$$\underset{k\in\mathbb{N}}{\forall} \quad (a^x)^{q_k} = \lim_{n\to\infty} (a^{p_n})^{q_k} \overset{(7.54\mathrm{b})}{=} \lim_{n\to\infty} a^{p_n q_k} = a^{x\, q_k},$$

$$\Rightarrow \quad (a^x)^y = \lim_{n\to\infty} (a^x)^{q_n} = \lim_{n\to\infty} a^{x\, q_n} \overset{(7.57)}{=} a^{x\, y},$$

thereby proving (7.54).

Proceeding to (7.55c), let $a > 1$ and $h > 0$. If $(q_n)_{n\in\mathbb{N}}$ is a strictly increasing sequence in $\mathbb{Q}^+$ with $\lim_{n\to\infty} q_n = h$, then $a^h = \lim_{n\to\infty} a^{q_n} > a^{q_1} > 1$. Thus, if $x < y$, let $h := y - x > 0$ to obtain $a^y = a^x\, a^h > a^x$, i.e. (7.55c). If $0 < a < 1$ and $x < y$, then $(1/a)^x < (1/a)^y$, yielding (7.55d). For (7.55a), consider $x > 0$ and $0 < a < b$. Then

$$\frac{b}{a} > 1 \quad \Rightarrow \quad \frac{b^x}{a^x} = \left(\frac{b}{a}\right)^x > 1 \quad \Rightarrow \quad b^x > a^x,$$

proving (7.55a). If $x < 0$ and $0 < a < b$, then $a^x = (1/a)^{-x} > (1/b)^{-x} = b^x$, proving (7.55b).

Finally, it remains to verify (7.56). For $x \ge 1$, the proof for rational x still works for irrational x. For $0 < x < 1$, one uses the usual sequence $(q_n)_{n\in\mathbb{N}}$ in $\mathbb{Q}$ with $\lim_{n\to\infty} q_n = x$ and obtains (recalling $a > 1$)

$$a^x = \lim_{n\to\infty} a^{q_n} \overset{(7.44)}{\le} \lim_{n\to\infty} \left(1 + q_n(a-1) \right) = 1 + x(a-1) < 1 + x \cdot a^{x+1},$$

proving (7.56). ∎

Definition 7.71 (Exponential and Power Functions). **(a)** Each function of the form

$$f : \mathbb{R}^+ \longrightarrow \mathbb{R}, \quad f(x) := x^\alpha, \quad \alpha \in \mathbb{R}, \tag{7.62}$$

is called a *power function*. For $\alpha > 0$, the power function is extended to $x = 0$ by setting $0^\alpha := 0$; for $\alpha \in \mathbb{Z}$, it is defined on $\mathbb{R} \setminus \{0\}$; for $\alpha \in \mathbb{N}_0$ even on $\mathbb{R}$.

(b) Each function of the form

$$f : \mathbb{R} \longrightarrow \mathbb{R}^+, \quad f(x) := a^x, \quad a > 0, \tag{7.63}$$

is called a *(general) exponential function*. The case where $a = e$ with e being Euler's number from (7.49) is of particular interest and importance. Most of the time, when referring to an exponential function, one actually means $x \mapsto e^x$. It is also common to write $\exp(x)$ instead of e^x.

Theorem 7.72. (a) *Every power function as defined in Def. 7.71(a) is continuous on its respective domain. Moreover, for each $\alpha > 0$, it is strictly increasing on $[0, \infty[$; for each $\alpha < 0$, it is strictly decreasing on $]0, \infty[$.*

(b) *Every exponential function as defined in Def. 7.71(b) is continuous. Moreover, for each $a > 1$, it is strictly increasing; for each $0 < a < 1$, it is strictly decreasing.*

Proof. (a): The monotonicity claims are provided by (7.55a) and (7.55b), respectively. For each $\alpha \in \mathbb{N}_0$, the power function is a polynomial, for each $\alpha \in \mathbb{Z}$, a rational function, i.e. continuity is provided by Ex. 7.40(b) and Ex. 7.40(c), respectively. For a general $\alpha \in \mathbb{R}$, the continuity proof on $\mathbb{R}^+$ will be postponed to Ex. 7.76(a) below, where it can be accomplished more easily. So it remains to show the continuity in $x = 0$ for $\alpha > 0$. However, if $(x_n)_{n \in \mathbb{N}}$ is a sequence in $\mathbb{R}^+$ with $\lim_{n \to \infty} x_n = 0$ and $k \in \mathbb{N}$ with $1/k \leq \alpha$, then, at least for n sufficiently large such that $x_n \leq 1$, $0 < x_n^\alpha \leq x_n^{1/k}$ by (7.55d). Then the continuity of $x \mapsto x^{1/k}$ implies $\lim_{n \to \infty} x_n^{1/k} = 0$ and the Sandwich Th. 7.16 implies $\lim_{n \to \infty} x_n^\alpha = 0$, proving continuity in $x = 0$.

(b): Everything has already been proved – continuity is provided by (7.61), monotonicity is provided by (7.55c) and (7.55d). ∎

Remark and Definition 7.73 (Logarithm)**.** According to Th. 7.72(b), for each $a \in \mathbb{R}^+ \setminus \{1\}$, the exponential function $f : \mathbb{R} \longrightarrow \mathbb{R}^+$, $f(x) := a^x$, is continuous and strictly monotone with $f(\mathbb{R}) = \mathbb{R}^+$ (verify that the image is all of $\mathbb{R}^+$ as an exercise). Then Th. 7.60 implies the existence of a continuous and strictly monotone inverse function $f^{-1} : \mathbb{R}^+ \longrightarrow \mathbb{R}$. For each $x \in \mathbb{R}^+$, we call $f^{-1}(x)$ the *logarithm* of x to *base a* and write $\log_a x := f^{-1}(x)$. The most important special case is where the base is Euler's number, $a = e$. This is called the *natural* logarithm. Bases $a = 2$ and $a = 10$ also carry special names, *binary* and *common* logarithm, respectively. The notation is

$$\ln x := \log_e x, \quad \mathrm{lb}\, x := \log_2 x, \quad \lg x := \log_{10} x, \tag{7.64}$$

however, the notation in the literature varies – one finds log used instead of ln, lb, and lg; one also finds lg instead of lb. So you always need to verify what precisely is meant by either notation.

Corollary 7.74. *For each $a \in \mathbb{R}^+ \setminus \{1\}$, the logarithm function $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $f(x) = \log_a x$ is continuous. For $a > 1$, it is strictly increasing; for $0 < a < 1$, it is strictly decreasing.* ∎

Theorem 7.75. *One obtains the following* logarithm rules*:*

$$\mathop{\forall}_{a \in \mathbb{R}^+ \setminus \{1\}} \qquad \log_a 1 = 0, \tag{7.65a}$$

$$\mathop{\forall}_{a \in \mathbb{R}^+ \setminus \{1\}} \qquad \log_a a = 1, \tag{7.65b}$$

$$\mathop{\forall}_{a \in \mathbb{R}^+ \setminus \{1\}} \mathop{\forall}_{x \in \mathbb{R}^+} \qquad a^{\log_a x} = x, \tag{7.65c}$$

$$\mathop{\forall}_{a \in \mathbb{R}^+ \setminus \{1\}} \mathop{\forall}_{x \in \mathbb{R}} \qquad \log_a a^x = x, \tag{7.65d}$$

$$\mathop{\forall}_{a \in \mathbb{R}^+ \setminus \{1\}} \mathop{\forall}_{x,y \in \mathbb{R}^+} \qquad \log_a(xy) = \log_a x + \log_a y, \qquad (7.65\text{e})$$

$$\mathop{\forall}_{a \in \mathbb{R}^+ \setminus \{1\}} \mathop{\forall}_{x \in \mathbb{R}^+} \mathop{\forall}_{y \in \mathbb{R}} \qquad \log_a(x^y) = y \log_a x, \qquad (7.65\text{f})$$

$$\mathop{\forall}_{a \in \mathbb{R}^+ \setminus \{1\}} \mathop{\forall}_{x,y \in \mathbb{R}^+} \qquad \log_a(x/y) = \log_a x - \log_a y, \qquad (7.65\text{g})$$

$$\mathop{\forall}_{a \in \mathbb{R}^+ \setminus \{1\}} \mathop{\forall}_{x \in \mathbb{R}^+} \mathop{\forall}_{n \in \mathbb{N}} \qquad \log_a \sqrt[n]{x} = \frac{1}{n} \log_a x, \qquad (7.65\text{h})$$

$$\mathop{\forall}_{a,b \in \mathbb{R}^+ \setminus \{1\}} \mathop{\forall}_{x \in \mathbb{R}^+} \qquad \log_b x = (\log_b a) \log_a x. \qquad (7.65\text{i})$$

Proof. All the rules are easy consequences of the logarithm being defined as the inverse function to $f : \mathbb{R} \longrightarrow \mathbb{R}^+$, $f(x) := a^x$.

(7.65a): It is $\log_a 1 = f^{-1}(1) = 0$, as $f(0) = a^0 = 1$.

(7.65b): It is $\log_a a = f^{-1}(a) = 1$, as $f(1) = a^1 = a$.

(7.65c): It is $a^{\log_a x} = f(f^{-1}(x)) = x$.

(7.65d): It is $\log_a a^x = f^{-1}(f(x)) = x$.

(7.65e): It is $\log_a(xy) = f^{-1}(xy) = f^{-1}\big(f(\log_a x + \log_a y)\big) = \log_a x + \log_a y$, since

$$f(\log_a x + \log_a y) = a^{\log_a x + \log_a y} = a^{\log_a x} a^{\log_a y} \overset{(7.65\text{c})}{=} xy.$$

(7.65f): It is $\log_a(x^y) = f^{-1}(x^y) = f^{-1}\big(f(y \log_a x)\big) = y \log_a x$, since

$$f(y \log_a x) = a^{y \log_a x} = (a^{\log_a x})^y \overset{(7.65\text{c})}{=} x^y.$$

(7.65g) is just a combination of (7.65e) and (7.65f): $\log_a(x/y) = \log_a(xy^{-1}) = \log_a x - \log_a y$.

(7.65h) is just a special case of (7.65f): $\log_a \sqrt[n]{x} = \log_a x^{1/n} = \frac{1}{n} \log_a x$.

(7.65i): One computes

$$(\log_b a) \log_a x \overset{(7.65\text{f})}{=} \log_b a^{\log_a x} \overset{(7.65\text{c})}{=} \log_b x.$$

Thus, we have verified all the rules and concluded the proof. ∎

Example 7.76. (a) For each $\alpha \in \mathbb{R}$, the power function

$$f : \mathbb{R}^+ \longrightarrow \mathbb{R}, \quad f(x) := x^\alpha = e^{\alpha \ln x}, \qquad (7.66)$$

is continuous, which follows from Th. 7.41, since $f = \exp \circ (\alpha \ln)$, ln is continuous by Cor. 7.74, and exp is continuous by Th. 7.72(b).

(b) As a consequence of Th. 7.41, each of the following functions f_1, f_2, f_3, where

$$f_1 : \mathbb{R} \longrightarrow \mathbb{R}, \qquad\qquad f_1(x) := \big(\exp(\lambda + x^2)\big)^\alpha,$$

$$f_2 : \mathbb{R} \longrightarrow \mathbb{R}, \qquad\qquad f_2(x) := \frac{1}{e^{\alpha x} + \lambda},$$

$$f_3 : \mathbb{R} \longrightarrow \mathbb{R}, \qquad\qquad f_3(x) := \frac{x^5}{(\lambda + |x|)^\alpha},$$

is continuous for each $\alpha \in \mathbb{R}$ and each $\lambda \in \mathbb{R}^+$.

7.3   Series

7.3.1   Definition and Convergence

Series are a special type of sequences, namely sequences whose members arise from summing up the members of another sequence. We have, on occasion, already encountered series, for example the harmonic series $(s_n)_{n \in \mathbb{N}}$, whose members s_n were defined in (7.27). In the present section, we will study series more systematically.

Definition 7.77. Given a sequence $(a_n)_{n \in \mathbb{N}}$ in $\mathbb{K}$ (or, more generally, in any set A, where an addition is defined), the sequence $(s_n)_{n \in \mathbb{N}}$, where

$$\underset{n \in \mathbb{N}}{\forall} \; s_n := \sum_{j=1}^{n} a_j, \tag{7.67}$$

is called an *(infinite) series* and is denoted by

$$\sum_{j=1}^{\infty} a_j := \sum_{j \in \mathbb{N}} a_j := (s_n)_{n \in \mathbb{N}}. \tag{7.68}$$

The a_n are called the *summands* of the series, the s_n its *partial sums*. Moreover, each series $\sum_{j=k}^{\infty} a_j$ with $k \in \mathbb{N}$ is called a *remainder (series)* of the series $(s_n)_{n \in \mathbb{N}}$.

The example of the remainder series already shows that it is useful to allow countable index sets other than $\mathbb{N}$. Thus, if $(a_j)_{j \in I}$, where I is a countable index set and $\phi : \mathbb{N} \longrightarrow I$ a bijective map, then define

$$\sum_{j \in I} a_j := \sum_{j=1}^{\infty} a_{\phi(j)} \tag{7.69}$$

(compare the definition in (3.19c) for finite sums). Note that the definition depends on ϕ, which is suppressed in the notation $\sum_{j \in I} a_j$.

———

For sequences in $\mathbb{K}$, the notion of convergence is available, and, thus, it is also available for series arising from real or complex sequences (as such series are, again, sequences in $\mathbb{K}$).

Definition 7.78. If $(s_n)_{n \in \mathbb{N}}$ is a series with the s_n defined as in (7.67) and with summands $a_j \in \mathbb{K}$, then the series is called *convergent* with *limit* $s \in \mathbb{K}$ if, and only if, $\lim_{n \to \infty} s_n = s$ in the sense of (7.1). In that case, one writes

$$\sum_{j=1}^{\infty} a_j = s \tag{7.70}$$

and calls s the *sum* of the series. The series is called *divergent* if, and only if, it is not convergent. We write $\sum_{j=1}^{\infty} a_j = \infty$ (resp. $\sum_{j=1}^{\infty} a_j = -\infty$) if, and only if, $(s_n)_{n \in \mathbb{N}}$ diverges to ∞ (resp. $-\infty$) in the sense of Def. 7.18.

Caveat 7.79. One has to use care as the symbol $\sum_{j=1}^{\infty} a_j$ is used with two completely different meanings. If it is used according to (7.68), then it means a *sequence*; if it is used according to (7.70), then it means a *real or complex number* (or, possibly, ∞ or $-\infty$). It should always be clear from the context, if it means a sequence or a number. For example, in the statement "the series $\sum_{j=1}^{\infty} 2^{-j}$ is convergent", it means a sequence; whereas in the statement "$\sum_{j=1}^{\infty} 2^{-j} = 1$", it means a number.

Example 7.80. (a) For each $q \in \mathbb{C}$ with $|q| < 1$, $\sum_{j=0}^{\infty} q^j$ is called a *geometric series*. From (3.22b) (the reader is asked to go back and check that (3.22b) and its proof, indeed, remain valid for each $q \in \mathbb{C}$), we obtain the partial sums $s_n = \sum_{j=0}^{n} q^j = \frac{1-q^{n+1}}{1-q}$. Since $|q| < 1$, we know $\lim_{n \to \infty} q^{n+1} = 0$ from Ex. 7.6. Thus, the series is convergent with

$$\underset{|q|<1}{\forall} \quad \sum_{j=0}^{\infty} q^j = \lim_{n \to \infty} s_n = \lim_{n \to \infty} \frac{1-q^{n+1}}{1-q} = \frac{1}{1-q}. \tag{7.71}$$

(b) In Ex. 7.30, we obtained the divergence of the harmonic series:

$$\sum_{k=1}^{\infty} \frac{1}{k} = \infty. \tag{7.72}$$

Corollary 7.81. *Let $\sum_{j=1}^{\infty} a_j$ and $\sum_{j=1}^{\infty} b_j$ be convergent series in $\mathbb{C}$.*

(a) *Linearity:*

$$\underset{\lambda,\mu \in \mathbb{C}}{\forall} \quad \sum_{j=1}^{\infty} (\lambda\, a_j + \mu\, b_j) = \lambda \sum_{j=1}^{\infty} a_j + \mu \sum_{j=1}^{\infty} b_j. \tag{7.73}$$

(b) *Complex Conjugation:*

$$\sum_{j=1}^{\infty} \overline{a_j} = \overline{\sum_{j=1}^{\infty} a_j}. \tag{7.74}$$

(c) *Monotonicity:*

$$\left(\underset{j \in \mathbb{N}}{\forall} \quad a_j, b_j \in \mathbb{R} \wedge a_j \leq b_j \right) \quad \Rightarrow \quad \sum_{j=1}^{\infty} a_j \leq \sum_{j=1}^{\infty} b_j. \tag{7.75}$$

(d) *Each remainder series* $\sum_{j=n+1}^{\infty} a_j$, $n \in \mathbb{N}$, *converges, and, letting* $S := \sum_{j=1}^{\infty} a_j$, $s_n := \sum_{j=1}^{n} a_j$, $r_n := \sum_{j=n+1}^{\infty} a_j$, *one has*

$$\left(\underset{n \in \mathbb{N}}{\forall}\ S = s_n + r_n \right), \qquad \lim_{n \to \infty} a_n = \lim_{n \to \infty} r_n = 0. \tag{7.76}$$

Proof. (a) follows from the first two identities of Th. 7.13(a), (b) is due to

$$\sum_{j=1}^{\infty} \overline{a_j} = \lim_{n \to \infty} \sum_{j=1}^{n} \overline{a_j} \overset{\text{Def. and Rem. 5.5(a)}}{=} \overline{\lim_{n \to \infty} \sum_{j=1}^{n} a_j} \overset{(7.11f)}{=} \overline{\lim_{n \to \infty} \sum_{j=1}^{n} a_j} = \overline{\sum_{j=1}^{\infty} a_j},$$

(c) follows from Th. 7.13(c), and, for (d), one computes

$$\lim_{n \to \infty} a_n = \lim_{n \to \infty} (s_n - s_{n-1}) = S - S = 0,$$

$$\underset{n \in \mathbb{N}}{\forall}\quad r_n = \lim_{k \to \infty} \sum_{j=n+1}^{k} a_j = \lim_{k \to \infty} (s_k - s_n) = S - s_n,$$

$$\lim_{n \to \infty} r_n = \lim_{n \to \infty} (S - s_n) = S - S = 0,$$

completing the proof. ∎

7.3.2 Convergence Criteria

Corollary 7.82. *Let* $\sum_{j=1}^{\infty} a_j$ *be series such that all* $a_j \in \mathbb{R}_0^+$. *If* $s_n := \sum_{j=1}^{n} a_j$ *are the partial sums of* $\sum_{j=1}^{\infty} a_j$, *then*

$$\lim_{n \to \infty} s_n = \begin{cases} \sup\{s_n : n \in \mathbb{N}\} & \text{if } (s_n)_{n \in \mathbb{N}} \text{ is bounded,} \\ \infty & \text{if } (s_n)_{n \in \mathbb{N}} \text{ is not bounded.} \end{cases} \tag{7.77}$$

Proof. Since $(s_n)_{n \in \mathbb{N}}$ is increasing, (7.77) is a consequence of (7.21). ∎

Theorem 7.83. *Let* $\sum_{j=1}^{\infty} a_j$ *and* $\sum_{j=1}^{\infty} b_j$ *be series in* $\mathbb{C}$ *such that* $|a_j| \leq |b_j|$ *holds for each* $j \geq k$ *for some fixed* $k \in \mathbb{N}$.

(a) *If* $\sum_{j=1}^{\infty} |b_j|$ *is convergent, then* $\sum_{j=1}^{\infty} a_j$ *is convergent as well, and, moreover,*

$$\left| \sum_{j=k}^{\infty} a_j \right| \leq \sum_{j=k}^{\infty} |b_j|. \tag{7.78}$$

(b) *If* $\sum_{j=1}^{\infty} a_j$ *is divergent, then* $\sum_{j=1}^{\infty} |b_j|$ *is divergent as well.*

Proof. Since (b) is merely the contraposition of (a), it suffices to prove (a). To this end, let $s_n := \sum_{j=1}^n a_j$ and $t_n := \sum_{j=1}^n |b_j|$ be the partial sums of $\sum_{j=1}^\infty a_j$ and $\sum_{j=1}^\infty |b_j|$, respectively. Since $(t_n)_{n \in \mathbb{N}}$ converges, it must be a Cauchy sequence by Th. 7.29. Thus,

$$\underset{\substack{N \in \mathbb{N}, \ n>m>N \\ N \geq k}}{\underset{\epsilon \in \mathbb{R}^+}{\forall} \ \exists} \ \forall \ |t_n - t_m| = |b_{m+1}| + \cdots + |b_n| < \epsilon$$

and the triangle inequality for finite sums implies

$$\underset{\substack{N \in \mathbb{N}, \ n>m>N \\ N \geq k}}{\underset{\epsilon \in \mathbb{R}^+}{\forall} \ \exists} \ \forall \ \begin{aligned} |s_n - s_m| &= |a_{m+1} + \cdots + a_n| \leq |a_{m+1}| + \cdots + |a_n| \\ &\leq |b_{m+1}| + \cdots + |b_n| < \epsilon, \end{aligned}$$

showing $(s_n)_{n \in \mathbb{N}}$ is a Cauchy sequence as well, i.e. convergent by Th. 7.29. Since the triangle inequality for finite sums also implies $\left| \sum_{j=k}^n a_j \right| \leq \sum_{j=k}^n |b_j|$ for each $n \geq k$, (7.78) is now a consequence of Th. 7.13(c). ∎

Definition 7.84. A series $\sum_{j=1}^\infty a_j$ in $\mathbb{R}$ is called *alternating* if, and only if, its summands alternate between positive and negative signs, i.e. if $\operatorname{sgn}(a_{j+1}) = -\operatorname{sgn}(a_j) \neq 0$ for each $j \in \mathbb{N}$.

Theorem 7.85 (Leibniz Criterion). *Let $\sum_{j=1}^\infty a_j$ be an alternating series. If the sequence $(|a_n|)_{n \in \mathbb{N}}$ of absolute values is strictly decreasing and $\lim_{n \to \infty} a_n = 0$, then the series is convergent and*

$$\underset{n \in \mathbb{N}}{\forall} \ \underset{0 < \theta_n < 1}{\exists} \quad r_n := \sum_{j=n+1}^\infty a_j = \theta_n \, a_{n+1}, \tag{7.79}$$

that means the error made when approximating the limit by the partial sum s_n has the same sign as the first neglected summand a_{n+1}, and its absolute value is less than $|a_{n+1}|$.

Proof. We first consider the case where $a_1 > 0$, i.e. where there exists a strictly decreasing sequence of positive numbers $(b_n)_{n \in \mathbb{N}}$ such that $a_n = (-1)^{n+1} b_n$. As the b_n are strictly decreasing, we obtain $b_n - b_{n+1} > 0$ for each $n \in \mathbb{N}$, such that the sequences $(u_n)_{n \in \mathbb{N}}$ and $(v_n)_{n \in \mathbb{N}}$, defined by

$$\underset{n \in \mathbb{N}}{\forall} \quad u_n := s_{2n} = \sum_{j=1}^n (b_{2j-1} - b_{2j}) = (b_1 - b_2) + (b_3 - b_4) + \cdots + (b_{2n-1} - b_{2n}),$$

$$\underset{n \in \mathbb{N}}{\forall} \quad v_n := s_{2n+1} = b_1 - \sum_{j=1}^n (b_{2j} - b_{2j+1})$$

$$= b_1 - (b_2 - b_3) - (b_4 - b_5) - \cdots - (b_{2n} - b_{2n+1}),$$

are strictly monotone, namely $(u_n)_{n \in \mathbb{N}}$ strictly increasing and $(v_n)_{n \in \mathbb{N}}$ strictly decreasing. Since, $0 < u_n < u_n + b_{2n+1} = v_n < b_1$ for each $n \in \mathbb{N}$, both sequences $(u_n)_{n \in \mathbb{N}}$ and $(v_n)_{n \in \mathbb{N}}$ are also bounded, and, thus, convergent by Th. 7.19, i.e. $U := \lim_{n \to \infty} u_n \in \mathbb{R}$ and $V := \lim_{n \to \infty} v_n \in \mathbb{R}$. Since

$$V - U = \lim_{n \to \infty} (v_n - u_n) = \lim_{n \to \infty} (s_{2n+1} - s_{2n}) = \lim_{n \to \infty} a_{2n+1} = 0,$$

we obtain $U = V$ and $\lim_{n \to \infty} s_n = U$ and $0 < U < b_1 = a_1$. In particular, there is $\theta \in]0, 1[$ satisfying $\sum_{j=1}^{\infty} a_j = \theta \, a_1$.

In the case $a_1 < 0$, the above proof yields convergence of $-\sum_{j=1}^{\infty} a_j = \sum_{j=1}^{\infty} (-a_j)$ with $\sum_{j=1}^{\infty} (-a_j) = \theta \, (-a_1)$ for a suitable $\theta \in]0, 1[$. However, this then yields, as before, $\sum_{j=1}^{\infty} a_j = \theta \, a_1$.

Applying the above result to each remainder series $\sum_{j=n+1}^{\infty} a_j$, $n \in \mathbb{N}$, completes the proof of (7.79) and the theorem.   ∎

Example 7.86. (a) Each of the following alternating series clearly converges, as the Leibniz criterion of Th. 7.85 clearly applies in each case:

$$\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} = 1 - \frac{1}{2} + \frac{1}{3} - + \ldots, \tag{7.80a}$$

$$\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{2\,j - 1} = 1 - \frac{1}{3} + \frac{1}{5} - + \ldots, \tag{7.80b}$$

$$\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{\ln(j+1)} = \frac{1}{\ln 2} - \frac{1}{\ln 3} + \frac{1}{\ln 4} - + \ldots \tag{7.80c}$$

(b) To see that Th. 7.85 is false without its monotonicity requirement, take any divergent series with $\sum_{j=1}^{\infty} a_j = \infty$, $0 < a_j$, $\lim_{j \to \infty} a_j = 0$ (for example the harmonic series), any convergent series with $\sum_{j=1}^{\infty} c_j = s \in \mathbb{R}^+$ and $0 < c_j$ (for example any geometric series with $0 < q < 1$), and define

$$d_n := \begin{cases} a_{(n+1)/2} & \text{for } n \text{ odd,} \\ -c_{n/2} & \text{for } n \text{ even.} \end{cases}$$

It is an exercise to show that $\sum_{j=1}^{\infty} d_j$ is an alternating series with $\lim_{n \to \infty} d_n = 0$ and $\sum_{j=1}^{\infty} d_j = \infty$.

Definition 7.87. The series $\sum_{j=1}^{\infty} a_j$ in $\mathbb{C}$ is said to be *absolutely convergent* if, and only if, $\sum_{j=1}^{\infty} |a_j|$ is convergent.

Corollary 7.88. *Every absolutely convergent series $\sum_{j=1}^{\infty} a_j$ is also convergent and satisfies the* triangle inequality *for infinite series:*

$$\left| \sum_{j=1}^{\infty} a_j \right| \leq \sum_{j=1}^{\infty} |a_j|. \tag{7.81}$$

Proof. The corollary is given by the special case $a_j = b_j$ for each $j \in \mathbb{N}$ of Th. 7.83(a).   ∎

Theorem 7.89. *We consider the series $\sum_{j=1}^{\infty} a_j$ in $\mathbb{C}$.*

(a) *If $\sum_{j=1}^{\infty} c_j$ is a convergent series such that $c_j \in \mathbb{R}_0^+$ and $|a_j| \leq c_j$ for each $j \in \mathbb{N}$, then $\sum_{j=1}^{\infty} a_j$ is absolutely convergent.*

(b) Root Test:

$$\left(\underset{0<q<1}{\exists} \left(\sqrt[n]{|a_n|} \leq q < 1 \text{ for almost all } n \in \mathbb{N} \right) \right)$$

$$\Rightarrow \quad \sum_{j=1}^{\infty} a_j \text{ is absolutely convergent}, \tag{7.82a}$$

$$\# \left\{ n \in \mathbb{N} : \sqrt[n]{|a_n|} \geq 1 \right\} = \infty \quad \Rightarrow \quad \sum_{j=1}^{\infty} a_j \text{ is divergent}. \tag{7.82b}$$

(c) Ratio Test: *If all $a_n \neq 0$, then*

$$\left(\underset{0<q<1}{\exists} \left(\left| \frac{a_{n+1}}{a_n} \right| \leq q < 1 \text{ for almost all } n \in \mathbb{N} \right) \right)$$

$$\Rightarrow \quad \sum_{j=1}^{\infty} a_j \text{ is absolutely convergent}, \tag{7.83a}$$

$$\left| \frac{a_{n+1}}{a_n} \right| \geq 1 \text{ for almost all } n \in \mathbb{N} \quad \Rightarrow \quad \sum_{j=1}^{\infty} a_j \text{ is divergent}. \tag{7.83b}$$

Proof. (a) is just another special case of Th. 7.83(a).

(b): If there is $q \in]0,1[$ and $N \in \mathbb{N}$ such that $\sqrt[n]{|a_n|} \leq q$ for each $n > N$, i.e. $|a_n| \leq q^n$ for each $n > N$, then, by (7.71), $\sum_{j=1}^{\infty} |a_j|$ is bounded by $\frac{1}{1-q} + \sum_{j=1}^{N} |a_j|$ and, thus, convergent. If $\sqrt[n]{|a_n|} \geq 1$ for infinitely many $n \in \mathbb{N}$, then $|a_n| \geq 1$ for infinitely many $n \in \mathbb{N}$, showing that $(a_n)_{n\in\mathbb{N}}$ does not converge to 0, proving the divergence of $\sum_{j=1}^{\infty} a_j$.

(c): If there is $q \in]0,1[$ and $N \in \mathbb{N}$ such that $\left| \frac{a_{n+1}}{a_n} \right| \leq q$ for each $n > N$, then, letting $C := |a_{N+1}|$, an induction shows $|a_{N+1+k}| \leq Cq^k$ for each $k \in \mathbb{N}$, i.e., by (7.71), $\sum_{j=1}^{\infty} |a_j|$ is bounded by $\frac{C}{1-q} + \sum_{j=1}^{N+1} |a_j|$ and, thus, convergent. If there is $N \in \mathbb{N}$ such that $\left| \frac{a_{n+1}}{a_n} \right| \geq 1$ for each $n > N$, then $|a_n| \geq |a_{N+1}| > 0$ for each $n > N$, showing $(a_n)_{n\in\mathbb{N}}$ does not converge to 0 and proving the divergence of $\sum_{j=1}^{\infty} a_j$. ∎

Caveat 7.90. In (7.82a), it does not suffice to have $\sqrt[n]{|a_n|} < 1$ to conclude convergence, and, likewise, $\left| \frac{a_{n+1}}{a_n} \right| < 1$ does not suffice in (7.83a): As a counterexample, consider the harmonic series, which does not converge, but $\sqrt[n]{1/n} < 1$ for each $n \geq 2$ and $\frac{1/(n+1)}{1/n} = \frac{n}{n+1} < 1$ for each $n \in \mathbb{N}$.

Example 7.91. **(a)** For each $z \in \mathbb{C}$ with $|z| < 1$ and each $p \in \mathbb{N}_0$, the series $\sum_{n=1}^{\infty} n^p z^n$ is absolutely convergent: We have $\lim_{n\to\infty} \sqrt[n]{n^p} = 1$ as a consequence of Ex. 7.65. This implies $\lim_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sqrt[n]{n^p |z|^n} = |z| < 1$. Thus, the root test of (7.82a) applies and proves convergence of the series.

(b) Let $z \in \mathbb{C}$. The series $\sum_{n=1}^{\infty} \frac{z^n n!}{n^n}$ is absolutely convergent for $|z| < e$ and divergent for $|z| \geq e$, where e is Euler's number from (7.49). We have, for each $n \in \mathbb{N}$,

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{|z| (n+1) n^n}{(n+1)^{n+1}} = \frac{|z|}{\left(1 + \frac{1}{n}\right)^n} \to \frac{|z|}{e} \quad \text{for } n \to \infty. \tag{7.84}$$

Thus, the ratio test of (7.83a) applies and proves absolute convergence of the series for $|z| < e$. For $|z| > e$, (7.83b) applies and proves divergence. Since, according to Ex. 7.66, $\left(1 + \frac{1}{n}\right)^n < e$ for each $n \in \mathbb{N}$, (7.83b) applies to prove divergence also for $|z| = e$.

7.3.3 Absolute Convergence and Rearrangements

In general, one has to use care when dealing with infinite series, as convergence properties and even the limit in case of convergence can depend on the *order* of the summands (in obvious contrast to the situation of finite sums). For real series that are convergent, but not absolutely convergent, one has the striking Riemann rearrangement Th. 7.93, that states one can choose an arbitrary number $S \in \mathbb{R} \cup \{-\infty, \infty\}$ and reorder the summands such that the new series converges to S (actually, Th. 7.93 says even more, namely that one can prescribe an entire interval of cluster points for the rearranged series). However, the situation is better for absolutely convergent series. In Th. 7.95, we will see that the sum of absolutely convergent series does not depend on the order of the summands.

Proposition 7.92. *Let $\sum_{j=1}^{\infty} a_j$ be a series in $\mathbb{R}$. Defining*

$$\mathop{\forall}_{j \in \mathbb{N}} \left(a_j^+ := \max\{a_j, 0\}, \quad a_j^- := \max\{-a_j, 0\} \right), \tag{7.85}$$

the following assertions (a) *and* (b) *hold:*

(a) $\sum_{j=1}^{\infty} a_j$ *is absolutely convergent if, and only if, both series $\sum_{j=1}^{\infty} a_j^+$ and $\sum_{j=1}^{\infty} a_j^-$ are convergent.*

(b) *If $\sum_{j=1}^{\infty} a_j$ is convergent, but not absolutely convergent, then*

$$\sum_{j=1}^{\infty} a_j^+ = \sum_{j=1}^{\infty} a_j^- = \infty. \tag{7.86}$$

Proof. The key observation is that (7.85) implies, for each $j \in \mathbb{N}$,

$$a_j^+ + a_j^- = |a_j|, \tag{7.87a}$$

$$a_j^+ - a_j^- = a_j, \tag{7.87b}$$

$$0 \leq a_j^+, a_j^- \leq |a_j|. \tag{7.87c}$$

(a): If $\sum_{j=1}^{\infty} a_j^+$ and $\sum_{j=1}^{\infty} a_j^-$ are convergent, then

$$\sum_{j=1}^{\infty} |a_j| \overset{(7.87a),(7.73)}{=} \sum_{j=1}^{\infty} a_j^+ + \sum_{j=1}^{\infty} a_j^-, \tag{7.88}$$

and, in particular, $\sum_{j=1}^{\infty} a_j$ is absolutely convergent. Conversely, if $\sum_{j=1}^{\infty} a_j$ is absolutely convergent, then $\sum_{j=1}^{\infty} a_j^+$ and $\sum_{j=1}^{\infty} a_j^-$ are convergent by (7.87c) and Th. 7.83(a).

(b): If $\sum_{j=1}^{\infty} a_j$ and $\sum_{j=1}^{\infty} a_j^+$ are convergent, then (7.87b) implies that $\sum_{j=1}^{\infty} a_j^-$ is also convergent and, thus, $\sum_{j=1}^{\infty} a_j$ absolutely convergent by (a). Likewise, if $\sum_{j=1}^{\infty} a_j$ and $\sum_{j=1}^{\infty} a_j^-$ are convergent, then (7.87b) implies that $\sum_{j=1}^{\infty} a_j^+$ is also convergent and, once again, $\sum_{j=1}^{\infty} a_j$ absolutely convergent by (a). Therefore, if $\sum_{j=1}^{\infty} a_j$ is convergent, but not absolutely convergent, then (7.86) must hold by (7.77). ∎

Theorem 7.93 (Riemann Rearrangement Theorem a.k.a. Riemann Series Theorem). *Let $\sum_{j=1}^{\infty} a_j$ be a series in $\mathbb{R}$. If $\sum_{j=1}^{\infty} a_j$ is convergent, but not absolutely convergent, then, given $x, y \in \mathbb{R} \cup \{-\infty, \infty\}$ with $x \leq y$, there exists a rearrangement $\sum_{j=1}^{\infty} b_j$ of the series (i.e. a reordering $(b_j)_{j \in \mathbb{N}}$ of $(a_j)_{j \in \mathbb{N}}$) such that $\sum_{j=1}^{\infty} b_j$ has precisely all elements of $[x, y]$ as cluster points (where we call $-\infty$ (resp. ∞) a cluster point of the real sequence $(t_n)_{n \in \mathbb{N}}$ if, and only if, $\#\{n \in \mathbb{N} : t_n < -N\} = \infty$ (resp. $\#\{n \in \mathbb{N} : t_n > N\} = \infty$) for each $N \in \mathbb{N}$). In particular, choosing $S := x = y \in \mathbb{R} \cup \{-\infty, \infty\}$, one can prescribe an arbitrary limit S such that $\sum_{j=1}^{\infty} b_j = S$.*

Sketch of Proof. Here we just give a sketch of the proof to convey its fairly simple idea; a detailed proof is provided in Appendix E.1. According to Prop. 7.92(b), (7.86) must hold, where the a_j^+ and a_j^- are as defined in (7.85). Thus, we can define

$$
\underset{k \in \mathbb{N}}{\forall} \quad x_k := \begin{cases} -k & \text{for } x = -\infty, \\ x & \text{for } x \in \mathbb{R}, \\ k & \text{for } x = \infty, \end{cases} \qquad y_k := \begin{cases} -k & \text{for } y = -\infty, \\ y & \text{for } y \in \mathbb{R}, \\ k & \text{for } y = \infty, \end{cases}
$$

and, noting $x_k \leq y_k$ for almost all $k \in \mathbb{N}$, alternate between adding summands a_j^+ until the partial sum exceeds y_k and subtracting summands a_j^- until the partial sum falls below x_k. If k is sufficiently large such that $x_k \leq y_k$, then, at each switching point (from adding to subtracting or vice versa), the absolute value of the difference between the last partial sum and x_k or y_k, respectively, is less than the value of the last contributing nonzero summand. Since

$$
\lim_{j \to \infty} a_j^+ = \lim_{j \to \infty} a_j^- = 0,
$$

the partial sums corresponding to the switching points converge to the respective endpoints x or y, respectively, and precisely all points between x and y are cluster points. ∎

We will now study the more benign situation of absolutely convergent series.

Theorem 7.94. *Let $\sum_{j=1}^{\infty} a_j$ and $\sum_{j=1}^{\infty} b_j$ be series in $\mathbb{C}$ such that $(b_n)_{n \in \mathbb{N}}$ is a reordering of $(a_n)_{n \in \mathbb{N}}$ in the sense of Def. 7.21. If $\sum_{j=1}^{\infty} a_j$ is absolutely convergent, then so is $\sum_{j=1}^{\infty} b_j$ and $\sum_{j=1}^{\infty} a_j = \sum_{j=1}^{\infty} b_j$.*

Proof. Let $s_n := \sum_{j=1}^{n} a_j$, $\tilde{s}_n := \sum_{j=1}^{n} |a_j|$, and $t_n := \sum_{j=1}^{n} b_j$ denote the respective partial sums. We will show that $\lim_{n \to \infty}(s_n - t_n) = 0$. Given $\epsilon > 0$, since $(\tilde{s}_n)_{n \in \mathbb{N}}$ is a Cauchy sequence by Th. 7.29, there exists $N \in \mathbb{N}$, such that

$$
\underset{n > m > N}{\forall} |\tilde{s}_n - \tilde{s}_m| = |a_{m+1}| + \cdots + |a_n| < \epsilon.
$$

Since $(b_n)_{n\in\mathbb{N}}$ is a reordering of $(a_n)_{n\in\mathbb{N}}$, there exists a bijective map $\phi : \mathbb{N} \longrightarrow \mathbb{N}$ such that $b_n = a_{\phi(n)}$ for each $n \in \mathbb{N}$. Since ϕ is bijective, there exists $M \in \mathbb{N}$ such that $\{1, 2, \ldots, N+1\} \subseteq \phi\{1, 2, \ldots, M\}$. Then $n > M$ implies $\phi(n) > N+1$, and

$$\underset{n>M}{\forall} \ \underset{k\in\mathbb{N}}{\exists} \ |s_n - t_n| \leq |a_{N+2}| + \cdots + |a_{N+k}| < \epsilon,$$

since all a_j with $j \leq N+1$ occur in both s_n and t_n and cancel in $s_n - t_n$ (i.e. all a_j that do not cancel must have an index $j > N+1$). So we have shown that $\lim_{n\to\infty}(s_n - t_n) = 0$, which, in turn, implies

$$\sum_{j=1}^{\infty} b_j = \lim_{n\to\infty} t_n = \lim_{n\to\infty} (t_n - s_n + s_n) = 0 + \sum_{j=1}^{\infty} a_j = \sum_{j=1}^{\infty} a_j.$$

Applying this to $\tilde{s}_n := \sum_{j=1}^{n} |a_j|$ yields $\sum_{j=1}^{\infty} |b_j| = \sum_{j=1}^{\infty} |a_j|$, proving absolute convergence of $\sum_{j=1}^{\infty} b_j$. ∎

Theorem 7.95. *Let I be an arbitrary infinite countable index set and let*

$$I = \overset{\cdot}{\bigcup_{n\in\mathbb{N}}} I_n \tag{7.89}$$

be a disjoint decomposition of I into (empty, finite, or infinite) countable index sets I_n.

(a) *If the series $\sum_{j\in I} a_j$ (cf. (7.69)) is absolutely convergent, then*

$$\sum_{j\in I} a_j = \sum_{n=1}^{\infty} \sum_{\alpha\in I_n} a_\alpha. \tag{7.90}$$

(b) *The following statements are equivalent:*

(i) *$\sum_{j\in I} a_j$ is absolutely convergent.*

(ii) *There exists a constant $C \in \mathbb{R}_0^+$ such that $\sum_{j\in J} |a_j| \leq C$ for each finite subset J of I.*

(iii) *$\sum_{n=1}^{\infty} \sum_{\alpha\in I_n} |a_\alpha| < \infty$.*

Proof. (a): First note that Th. 7.94 implies that, for absolutely convergent $\sum_{j\in I} a_j$, the limit $\sum_{j\in I} a_j = \sum_{j=1}^{\infty} a_{\phi(j)}$ does *not* depend on the bijective map $\phi : \mathbb{N} \longrightarrow I$: For each bijective map $\psi : \mathbb{N} \longrightarrow I$, $(a_{\psi(j)})_{j\in\mathbb{N}}$ is a reordering of $(a_{\phi(j)})_{j\in\mathbb{N}}$ and, thus, $\sum_{j=1}^{\infty} a_{\psi(j)} = \sum_{j=1}^{\infty} a_{\phi(j)}$.

Analogously, the sums $\sum_{\alpha\in I_n} a_\alpha$ do not depend on the order of the indices in I_n.

Claim 1. If $M \subseteq I$, then $S(I) = S(M) + S(I \setminus M)$, where $S(J) := \sum_{j\in J} a_j$ for each $J \subseteq I$.

Proof. If $M = \emptyset$, then there is nothing to prove. If $\#M = n \in \mathbb{N}$, then let $\phi_1 : \{1, \ldots, n\} \longrightarrow M$ and $\phi_2 : \{n+1, n+2, \ldots\} \longrightarrow I \setminus M$ be bijective maps. Then

$$\phi : \mathbb{N} \longrightarrow I, \quad \phi(j) := \begin{cases} \phi_1(j) & \text{for } j \leq n, \\ \phi_2(j) & \text{for } j > n, \end{cases}$$

is a bijective map. Moreover,

$$S(I) = \sum_{j=1}^{\infty} a_{\phi(j)} \overset{(7.76)}{=} \sum_{j=1}^{n} a_{\phi(j)} + \sum_{j=n+1}^{\infty} a_{\phi(j)} = S(M) + S(I \setminus M),$$

establishing the case.

If $\#M = \#(I \setminus M) = \#\mathbb{N}$, then let $\phi_1 : \{1, 3, 5, \ldots\} \longrightarrow M$ and $\phi_2 : \{2, 4, 6, \ldots\} \longrightarrow I \setminus M$ be bijective maps. Then

$$\phi : \mathbb{N} \longrightarrow I, \quad \phi(j) := \begin{cases} \phi_1(j) & \text{for } j \text{ odd}, \\ \phi_2(j) & \text{for } j \text{ even}, \end{cases}$$

is a bijective map. Define,

$$b_{\phi(j)} := \begin{cases} a_{\phi(j)} & \text{for } j \text{ odd}, \\ 0 & \text{for } j \text{ even}, \end{cases} \quad c_{\phi(j)} := \begin{cases} a_{\phi(j)} & \text{for } j \text{ even}, \\ 0 & \text{for } j \text{ odd}. \end{cases}$$

One then obtains

$$S(I) = \sum_{j=1}^{\infty} a_{\phi(j)} \overset{(7.73)}{=} \sum_{j=1}^{\infty} b_{\phi(j)} + \sum_{j=1}^{\infty} c_{\phi(j)} = S(M) + S(I \setminus M),$$

establishing the case. ▲

Claim 2. If $I = \dot{\bigcup}_{n=1}^{k} M_n$ with $k \in \mathbb{N}$ is a decomposition of I, then, using the notation introduced in Cl. 1, $S(I) = \sum_{n=1}^{k} S(M_n)$.

Proof. Follows by an induction from Cl. 1. ▲

Coming back to (7.89), Cl. 2 implies

$$\underset{k \in \mathbb{N}}{\forall} \left(S(I) = S(I_1) + S(I_2) + \cdots + S(I_k) + S(M_k), \quad \text{where} \quad M_k := I \setminus \bigcup_{j=1}^{k} I_j \right).$$

To prove the equality in (7.90), fix a bijective $\phi : \mathbb{N} \longrightarrow I$, and let $\epsilon > 0$. Due to Cor. 7.81(d), the sums $r_n := \sum_{j=n+1}^{\infty} |a_{\phi(j)}|$ of the remainder series converge to 0, i.e. there exists $N \in \mathbb{N}$ such that $r_n < \epsilon$ for each $n > N$. More generally, for each (empty, finite, or infinite) subset $J \subseteq \{N+2, N+3, \ldots\}$,

$$\sum_{j \in J} |a_{\phi(j)}| \leq \sum_{j=N+2}^{\infty} |a_{\phi(j)}| = r_{N+1} < \epsilon.$$

Next, we choose $M \in \mathbb{N}$ sufficiently large such that $\{\phi(1), \ldots, \phi(N+1)\} \subseteq I_1 \cup \cdots \cup I_M$. Then, for each $k > M$,

$$|S(M_k)| = \left| \sum_{j \in M_k} a_j \right| \overset{(7.81)}{\leq} \sum_{j \in M_k} |a_j| \leq \sum_{j=N+2}^{\infty} |a_{\phi(j)}| = r_{N+1} < \epsilon,$$

proving

$$\sum_{j \in I} a_j = S(I) = \lim_{k \to \infty} \sum_{n=1}^{k} S(I_n) = \sum_{n=1}^{\infty} \sum_{\alpha \in I_n} a_\alpha,$$

which is (7.90).

(b): (i) implies (ii) with $C := \sum_{j \in I} |a_j|$ using Cl. 1 (with a_j replaced by $|a_j|$). (i) implies (iii) using (7.90) (with a_j replaced by $|a_j|$). (ii) implies (i) via (7.77), as C is an upper bound for $(\sum_{j=1}^{n} |a_{\phi(j)}|)_{n \in \mathbb{N}}$ for each bijection $\phi : \mathbb{N} \longrightarrow I$. Finally, (iii) implies (ii) with $C := \sum_{n=1}^{\infty} \sum_{\alpha \in I_n} |a_\alpha|$, since, given a finite $J \subseteq I$, there exists $k \in \mathbb{N}$ such that $J \subseteq I_1 \cup \cdots \cup I_k$, i.e.

$$\sum_{j \in J} |a_j| \leq \sum_{n=1}^{k} \sum_{\alpha \in I_n} |a_\alpha| \leq \sum_{n=1}^{\infty} \sum_{\alpha \in I_n} |a_\alpha| = C,$$

thereby completing the proof. ∎

Example 7.96. We apply Th. 7.95 to so-called *double series*, i.e. to series with index set $I := \mathbb{N} \times \mathbb{N}$. The following notation is common:

$$\sum_{m,n=1}^{\infty} a_{mn} := \sum_{(m,n) \in \mathbb{N} \times \mathbb{N}} a_{(m,n)}, \tag{7.91}$$

where one writes a_{mn} (also $a_{m,n}$) instead of $a_{(m,n)}$. Recall from Th. 3.16 that $\mathbb{N} \times \mathbb{N}$ is countable. In general, the convergence properties of the double series and, if it exists, the value of the sum, will depend on the chosen bijection $\phi : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$.

However, we will now assume our double series to be absolutely convergent. Then Th. 7.94 guarantees the sum does not depend on the chosen bijection and we can apply Th. 7.95. Applying Th. 7.95 to the decompositions

$$\mathbb{N} \times \mathbb{N} = \overset{\text{\Large .}}{\bigcup_{m \in \mathbb{N}}} \{(m,n) : n \in \mathbb{N}\}, \tag{7.92a}$$

$$\mathbb{N} \times \mathbb{N} = \overset{\text{\Large .}}{\bigcup_{n \in \mathbb{N}}} \{(m,n) : m \in \mathbb{N}\}, \tag{7.92b}$$

$$\mathbb{N} \times \mathbb{N} = \overset{\text{\Large .}}{\bigcup_{k \in \mathbb{N}}} \{(m,n) \in \mathbb{N} \times \mathbb{N} : m + n = k\}, \tag{7.92c}$$

yields

$$\sum_{(m,n)\in\mathbb{N}\times\mathbb{N}} a_{(m,n)} \stackrel{(7.92\text{a})}{=} \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} \stackrel{(7.92\text{b})}{=} \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{mn}$$

$$\stackrel{(7.92\text{c})}{=} \sum_{k=2}^{\infty}\sum_{m+n=k} a_{mn} := \sum_{k=2}^{\infty}\sum_{m=1}^{k-1} a_{m,k-m}. \tag{7.93}$$

Theorem 7.97. *It is possible to compute the product of two absolutely convergent (real or complex) series $\sum_{m=1}^{\infty} a_m$ and $\sum_{m=1}^{\infty} b_m$ as a double series:*

$$\left(\sum_{m=1}^{\infty} a_m\right)\left(\sum_{m=1}^{\infty} b_m\right) = \sum_{m,n=1}^{\infty} a_m b_n = \sum_{k=2}^{\infty}\sum_{m=1}^{k-1} a_m b_{k-m} = \sum_{k=2}^{\infty} c_k,$$

$$where \quad c_k := \sum_{m=1}^{k-1} a_m b_{k-m} = a_1 b_{k-1} + a_2 b_{k-2} + \cdots + a_{k-1} b_1. \tag{7.94}$$

This form of computing the product is known as a Cauchy product.

Proof. We first show that $\sum_{m,n=1}^{\infty} a_m b_n$ is absolutely convergent: By letting $A := \sum_{m=1}^{\infty} |a_m|$ and $B := \sum_{m=1}^{\infty} |b_m|$, we obtain

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} |a_m b_n| = \sum_{m=1}^{\infty} \left(|a_m|\,B\right) = AB < \infty,$$

i.e. $\sum_{m,n=1}^{\infty} a_m b_n$ is absolutely convergent according to Th. 7.95(b)(iii). Now the second equality in (7.94) is just the third equality in (7.93), and the first equality in (7.94) also follows from (7.93):

$$\sum_{m,n=1}^{\infty} a_m b_n = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_m b_n = \sum_{m=1}^{\infty} a_m \sum_{n=1}^{\infty} b_n = \left(\sum_{m=1}^{\infty} a_m\right)\left(\sum_{m=1}^{\infty} b_m\right),$$

completing the proof.   ∎

Theorem 7.97 will be useful in Sec. 8.2 below.

7.3.4 b-**Adic Representations of Real Numbers**

We are mostly used to representing real numbers in the decimal system. For example, we write

$$x = \frac{395}{3} = 131.\overline{6} = 1\cdot 10^2 + 3\cdot 10^1 + 1\cdot 10^0 + \sum_{n=1}^{\infty} 6\cdot 10^{-n}, \tag{7.95a}$$

where

$$\sum_{n=1}^{\infty} 6\cdot 10^{-n} \stackrel{(7.71)}{=} 6\cdot\left(\frac{1}{1-\frac{1}{10}} - 1\right) = 6\cdot\frac{1}{9} = \frac{2}{3}.$$

The decimal system represents real numbers as, in general, infinite series of decimal fractions. Digital computers represent numbers in the dual system, using base 2 instead of 10. For example, the number from (7.95a) has the dual representation

$$x = 10000011.\overline{10} = 2^7 + 2^1 + 2^0 + \sum_{n=0}^{\infty} 2^{-(2n+1)}, \qquad (7.95b)$$

where it is an exercise to verify

$$\sum_{n=0}^{\infty} 2^{-(2n+1)} = \frac{2}{3}.$$

Representations with base 16 (hexadecimal) and 8 (octal) are also of importance when working with digital computers. More generally, each natural number $b \geq 2$ can be used as a base.

Definition 7.98. Let $b \geq 2$ be a natural number.

(a) Given an integer $N \in \mathbb{Z}$ and a sequence $(d_N, d_{N-1}, d_{N-2}, \ldots)$ in $\{0, \ldots, b-1\}$, the series

$$\sum_{\nu=0}^{\infty} d_{N-\nu} \, b^{N-\nu} \qquad (7.96)$$

is called a *b-adic series*. The number b is called the *base* or the *radix*, and the numbers d_ν are called *digits*.

(b) If $x \in \mathbb{R}_0^+$ is the sum of the b-adic series given by (7.96), than one calls the b-adic series a b-adic *representation* or a b-adic *expansion* of x.

Theorem 7.99. *Given a natural number $b \geq 2$ and a nonnegative real number $x \in \mathbb{R}_0^+$, there exists a b-adic series representing x, i.e. there is $N \in \mathbb{Z}$ and a sequence $(d_N, d_{N-1}, d_{N-2}, \ldots)$ in $\{0, \ldots, b-1\}$ such that*

$$x = \sum_{\nu=0}^{\infty} d_{N-\nu} \, b^{N-\nu}. \qquad (7.97)$$

If one introduces the additional requirement that $0 \neq d_N$, then each $x > 0$ has either a unique b-adic representation or precisely two b-adic representations. More precisely, for $0 \neq d_N$ and $x > 0$, the following statements are equivalent:

(i) *The b-adic representation of x is not unique.*

(ii) *There are precisely two b-adic representations of x.*

(iii) *There exists a b-adic representation of x such that $d_n = 0$ for each $n \leq n_0$ for some $n_0 < N$.*

(iv) *There exists a b-adic representation of x such that $d_n = b - 1$ for each $n \leq n_0$ for some $n_0 \leq N$.*

Proof. The proof is a bit lengthy and is provided in Appendix E.2.   ∎

Example 7.100. Every natural number has precisely two decimal (i.e. 10-adic) representations. For instance,

$$2 = 2.\overline{0} = 1.\overline{9} = 1 + \sum_{n=1}^{\infty} 9 \cdot 10^{-n} \overset{(7.71)}{=} 1 + 9 \cdot \left(\frac{1}{1 - \frac{1}{10}} - 1 \right) = 1 + 9 \cdot \frac{1}{9}, \qquad (7.98)$$

and analogously for all other natural numbers.

8   Convergence of $\mathbb{K}$-Valued Functions

8.1   Pointwise and Uniform Convergence

So far we have studied the convergence of sequences in $\mathbb{K}$. We will now also need to study the convergence of sequences $(f_n)_{n \in \mathbb{N}}$, where each member f_n of the sequence is a function $f_n : M \longrightarrow \mathbb{K}$, $M \subseteq \mathbb{C}$. Here, for the first time, we encounter the situation that there exist different useful notions of convergence for such sequences.

Definition 8.1. Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of functions, $f_n : M \longrightarrow \mathbb{K}$, $\emptyset \neq M \subseteq \mathbb{C}$.

(a) We say $(f_n)_{n \in \mathbb{N}}$ converges *pointwise* to $f : M \longrightarrow \mathbb{K}$ if, and only if, $\lim_{n \to \infty} f_n(z) = f(z)$ for each $z \in M$, i.e. if, and only if,

$$\underset{z \in M}{\forall} \, \underset{\epsilon \in \mathbb{R}^+}{\forall} \, \underset{N \in \mathbb{N}}{\exists} \, \underset{n > N}{\forall} \, |f_n(z) - f(z)| < \epsilon. \qquad (8.1)$$

So, in general, N in (8.1) depends on both z and ϵ.

(b) We say $(f_n)_{n \in \mathbb{N}}$ converges *uniformly* to $f : M \longrightarrow \mathbb{K}$ if, and only if,

$$\underset{\epsilon \in \mathbb{R}^+}{\forall} \, \underset{N \in \mathbb{N}}{\exists} \, \underset{n > N}{\forall} \, \underset{z \in M}{\forall} \, |f_n(z) - f(z)| < \epsilon. \qquad (8.2)$$

In (8.2), N is still allowed to depend on ϵ, but, in contrast to the situation of (8.1), not on z – in that sense, the convergence is uniform in z.

Remark 8.2. It is immediate from Def. 8.1(a),(b) that uniform convergence implies pointwise convergence, but Ex. 8.3(b) below will show the converse is not true.

Example 8.3. (a) Let $\emptyset \neq M \subseteq \mathbb{C}$ (for example $M = [0,1]$ or $M = B_1(0)$), and $f_n : M \longrightarrow \mathbb{K}$, $f_n(z) = 1/n$ for each $n \in \mathbb{N}$. Then, clearly, $(f_n)_{n \in \mathbb{N}}$ converges uniformly to $f \equiv 0$.

(b) The sequence $(f_n)_{n \in \mathbb{N}}$, where $f_n : [0,1] \longrightarrow \mathbb{R}$, $f_n(x) := x^n$, converges pointwise, but not uniformly, to

$$f : [0,1] \longrightarrow \mathbb{R}, \quad f(x) := \begin{cases} 0 & \text{for } 0 \leq x < 1, \\ 1 & \text{for } x = 1 : \end{cases} \qquad (8.3)$$

For $x = 1$, $\lim_{n\to\infty} x^n = \lim_{n\to\infty} 1 = 1$, and, for $0 \leq x < 1$, $\lim_{n\to\infty} x^n = 0$ by Ex. 7.6. To see that the convergence is not uniform, consider $\epsilon := \frac{1}{2}$. Then, for every $n \in \mathbb{N}$, according to the intermediate value Th. 7.57, there exists $\xi_n \in]0, 1[$ such that $f_n(\xi_n) = \xi_n^n = \frac{1}{2}$, i.e.

$$\underset{n\in\mathbb{N}}{\forall} |f_n(\xi_n) - f(\xi_n)| = \xi_n^n = \frac{1}{2} = \epsilon, \qquad (8.4)$$

proving the convergence is not uniform.

Theorem 8.4. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions, $f_n : M \longrightarrow \mathbb{K}$, $\emptyset \neq M \subseteq \mathbb{C}$. If $(f_n)_{n\in\mathbb{N}}$ converges uniformly to $f : M \longrightarrow \mathbb{K}$ and all f_n are continuous at $\zeta \in M$, then f is also continuous at ζ. In particular, if each f_n is continuous, then so is f (uniform limits of continuous functions are continuous).*

Proof. Let $\epsilon > 0$. Due to the uniform convergence of $(f_n)_{n\in\mathbb{N}}$,

$$\underset{m\in\mathbb{N}}{\exists} \underset{z\in M}{\forall} |f_m(z) - f(z)| < \frac{\epsilon}{3}. \qquad (8.5)$$

Due to the continuity of f_m in ζ,

$$\underset{\delta>0}{\exists} \underset{z\in M\cap B_\delta(\zeta)}{\forall} |f_m(z) - f_m(\zeta)| < \frac{\epsilon}{3}. \qquad (8.6)$$

Thus,

$$\underset{z\in M\cap B_\delta(\zeta)}{\forall} |f(z)-f(\zeta)| \leq |f(z)-f_m(z)|+|f_m(z)-f_m(\zeta)|+|f_m(\zeta)-f(\zeta)| < 3\cdot\frac{\epsilon}{3} = \epsilon, \ \ (8.7)$$

proving continuity of f in ζ. ∎

8.2 Power Series

Definition 8.5. (a) In Def. 7.77, it was mentioned that series can be formed from each sequence in a set A, where an addition is defined. Letting $\emptyset \neq M \subseteq \mathbb{C}$, we now consider $A := \mathcal{F}(M, \mathbb{K})$, i.e. the set of functions from M into $\mathbb{K}$. Then the addition on A is defined according to (6.1a) and, given a sequence of functions $(f_n)_{n\in\mathbb{N}}$ in A, the series

$$\sum_{j=1}^{\infty} f_j := (s_n)_{n\in\mathbb{N}} \qquad (8.8)$$

is defined as the sequence of partial sums $s_n := \sum_{j=1}^{n} f_j$.

(b) Given a sequence of functions $(f_n)_{n\in\mathbb{N}}$, where $f_n : \mathbb{K} \longrightarrow \mathbb{K}$, $f_n(z) = a_n z^n$ with $a_n \in \mathbb{K}$, then

$$\sum_{j=0}^{\infty} a_j z^j := \sum_{j=0}^{\infty} f_j \qquad (8.9)$$

is called a *power series* and the a_j are called the *coefficients* of the power series. Note: The notation $\sum_{j=0}^{\infty} a_j z^j$ introduced in (8.9) is very common, but not entirely correct, since one writes $a_j z^j = f_j(z)$ for the summands, even though one actually means f_j. Moreover, one uses the same notation if one actually does mean the series $\sum_{j=0}^{\infty} f_j(z)$ in $\mathbb{K}$, so one has to see from the context if $\sum_{j=0}^{\infty} a_j z^j$ means a series of $\mathbb{K}$-valued functions or a series of numbers.

Definition 8.6. Consider a series of $\mathbb{K}$-valued functions $\sum_{j=1}^{\infty} f_j$ as in Def. 8.5(a), in particular, $s_n := \sum_{j=1}^{n} f_j$ for each $n \in \mathbb{N}$.

(a) The series converges *pointwise* to $f : M \longrightarrow \mathbb{K}$ if, and only if, it (i.e. $(s_n)_{n \in \mathbb{N}}$) converges pointwise in the sense of Def. 8.1(a). In that case, we use the notation

$$f = \sum_{j=1}^{\infty} f_j. \tag{8.10}$$

If (8.10) holds, then the series is sometimes called a *series expansion* of f, in particular, a *power series expansion* if the series happens to be a power series.

Analogous to the situation of series in $\mathbb{K}$, the notation $\sum_{j=1}^{\infty} f_j$ is also used with two different meanings – it can mean the sequence of partial sums as in (8.8) or, in the case of convergent series, the limit function as in (8.10) (cf. Caveat 7.79).

(b) The series converges *uniformly* to $f : M \longrightarrow \mathbb{K}$ if, and only if, it converges uniformly in the sense of Def. 8.1(b).

Corollary 8.7. *Consider a function series* $\sum_{j=1}^{\infty} f_j$ *with* $f_j : M \longrightarrow \mathbb{K}$, $\emptyset \neq M \subseteq \mathbb{C}$.

(a) *The series converges uniformly to some* $f : M \longrightarrow \mathbb{K}$ *if, and only if, for each* $n \in \mathbb{N}$ *and each* $z \in M$, *the remainder series* $\sum_{j=n+1}^{\infty} f_j(z)$ *in* $\mathbb{K}$ *converges to some* $r_n(z) \in \mathbb{K}$ *such that*

$$\underset{\epsilon \in \mathbb{R}^+}{\forall} \; \underset{N \in \mathbb{N}}{\exists} \; \underset{n > N}{\forall} \; \underset{z \in M}{\forall} \; |r_n(z)| < \epsilon. \tag{8.11}$$

(b) *If* $\sum_{j=1}^{\infty} a_j$ *is a convergent series in* $\mathbb{R}_0^+$, *then the condition*

$$\underset{z \in M}{\forall} \; \underset{j \in \mathbb{N}}{\forall} \; |f_j(z)| \leq a_j \tag{8.12}$$

implies uniform convergence of $\sum_{j=1}^{\infty} f_j$.

(c) *If each* f_j *is continuous in* $\zeta \in M$ *and the series converges uniformly to* $f : M \longrightarrow \mathbb{K}$, *then* f *is continuous in* ζ. *In particular, if each* f_j *is continuous, then* f *is continuous.*

Proof. (a): If $\sum_{j=1}^{\infty} f_j$ converges uniformly to f, then $f(z) = \sum_{j=1}^{\infty} f_j(z)$ holds for each $z \in M$, $r_n(z) = f(z) - s_n(z)$ for each $n \in \mathbb{N}$, $z \in M$ according to (7.76), where $s_n(z) := \sum_{j=1}^{n} f_j(z)$. Then (8.11) is just (8.2), where the s_n now play the role of the f_n in (8.2). Conversely, if the remainder series converge for each $z \in M$, then we can define

$f : M \longrightarrow \mathbb{K}$, $f(z) := f_1(z) + r_1(z) = \sum_{j=1}^{\infty} f_j(z)$. Then, once again, $r_n(z) = f(z) - s_n(z)$ for each $n \in \mathbb{N}$, $z \in M$, and (8.11) is just (8.2), yielding the uniform convergence of the series.

(b): First, (8.12) implies each remainder series $\sum_{j=n+1}^{\infty} f_j(z)$ converges absolutely. Thus, with $r_n(z)$ as in (a),

$$\underset{z \in M}{\forall} |r_n(z)| \overset{(7.81)}{\leq} \sum_{j=n+1}^{\infty} |f_j(z)| \leq \sum_{j=n+1}^{\infty} a_j \to 0 \quad \text{for } n \to \infty,$$

such that (a) yields uniform convergence.

(c) is immediate from Th. 8.4. ∎

Remark 8.8. Given a function series $\sum_{j=1}^{\infty} f_j$ with $f_j : M \longrightarrow \mathbb{K}$, $\emptyset \neq M \subseteq \mathbb{C}$; for each $z \in M$, $\sum_{j=1}^{\infty} f_j(z)$ constitutes a series in $\mathbb{K}$. Typically, one will only have convergence of $\sum_{j=1}^{\infty} f_j(z)$ in $\mathbb{K}$ on a subset $C \subseteq M$. The series then converges pointwise in the sense of Def. 8.6(a) if all f_j are restricted to C. It can be very difficult to determine if $\sum_{j=1}^{\infty} f_j(z)$ converges or diverges for some $z \in M$, and such investigations are often of particular interest in the context of function series. Even for power series, studying convergence can still be difficult, but the availability of the following Th. 8.9 does help to (at least partially) settle the question in many cases.

Theorem 8.9. *For each power series* $\sum_{j=0}^{\infty} a_j z^j$, $a_j \in \mathbb{K}$, *there exists a number* $r \in [0, \infty] := \mathbb{R}_0^+ \cup \{\infty\}$, *called the* radius of convergence *of the power series, such that*

$$\left(z \in \mathbb{K} \wedge |z| < r \right) \quad \Rightarrow \quad \sum_{j=0}^{\infty} a_j z^j \quad \text{converges absolutely in } \mathbb{K}, \tag{8.13a}$$

$$\left(z \in \mathbb{K} \wedge |z| > r \right) \quad \Rightarrow \quad \sum_{j=0}^{\infty} a_j z^j \quad \text{diverges in } \mathbb{K} \tag{8.13b}$$

(for $r = \infty$, *(8.13a) claims absolute convergence for each* $z \in \mathbb{K}$). *In particular,* $\sum_{j=0}^{\infty} a_j z^j$ *converges pointwise in the sense of Def. 8.6(a) for each* $z \in B_r(0)$ *(cf. Def. 7.7(a)). Moreover,*

$$\underset{0 < r_0 < r}{\forall} \left(\sum_{j=0}^{\infty} a_j z^j \quad \begin{array}{l} \text{converges uniformly on } \overline{B}_{r_0}(0) \text{ (cf. Ex. 7.47(a))} \\ \text{in the sense of Def. 8.6(b)} \end{array} \right). \tag{8.14}$$

For the radius of convergence, one has the formula

$$r = \frac{1}{L}, \quad \text{where} \quad L := \limsup_{n \to \infty} \sqrt[n]{|a_n|}. \tag{8.15}$$

In (8.15), $\limsup$ *denotes the so-called* limit superior, *which is defined as the largest cluster point of the sequence* $(\sqrt[n]{|a_n|})_{n \in \mathbb{N}}$ *if the sequence is bounded (cf. Th. 7.27) and* ∞ *if the sequence is unbounded. As the limit superior can be 0 or* ∞, *we also define* $1/0 := \infty$ *and* $1/\infty := 0$ *in (8.15).*

One has the simpler formula

$$r = \lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right|, \tag{8.16}$$

provided all a_n are nonzero and provided the limit in (8.16) either exists in $\mathbb{R}_0^+$ or is ∞.

Proof. For the proof of (8.15), we apply the root test from Th. 7.89(b). Here, for the root test, we have to consider the sequence $(\sqrt[n]{|a_n||z|^n})_{n \in \mathbb{N}}$. As a consequence of (7.11a) and Prop. 7.26, $\limsup_{n \to \infty}(\lambda x_n) = \lambda \limsup_{n \to \infty} x_n$ for each $\lambda > 0$ and each sequence $(x_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ (with $\lambda \cdot \infty := \infty$, this also holds if the limit superior is infinite). Thus,

$$\limsup_{n \to \infty} \sqrt[n]{|a_n||z|^n} = |z| \limsup_{n \to \infty} \sqrt[n]{|a_n|} = |z| \, L.$$

If $|z| > 1/L$, then $|z| L > 1$ and (7.82b) applies, i.e. (8.13b) holds for $r = 1/L$. If $|z| < 1/L$, then $|z| L < 1$, and, recalling the Bolzano-Weierstrass Th. 7.27, one sees that (7.82a) applies, i.e. (8.13a) holds for $r = 1/L$.

Next, if $0 < r_0 < r$, then $\sum_{j=0}^{\infty} |a_j r_0^j|$ converges according to (8.13a). Since, for each $z \in B_{r_0}(0)$ and each $j \in \mathbb{N}$, we have $|a_j z^j| \leq |a_j r_0^j|$, (8.14) is a consequence of Cor. 8.7(b).

The validity of (8.16) follows from the ratio test of Th. 7.89(c): If all $a_n \neq 0$ and $z \neq 0$, then

$$\lim_{n \to \infty} \left| \frac{a_{n+1} z^{n+1}}{a_n z^n} \right| = |z| \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = \frac{|z|}{\lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right|}.$$

If $|z| < l := \lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right|$, then $|z|/l < 1$, i.e. (7.83a) applies, proving (8.13a) for $r = l$. If $|z| > l$, then $|z|/l > 1$, i.e. (7.83b) applies, proving (8.13b) for $r = l$. $\blacksquare$

Corollary 8.10. *If $\sum_{j=0}^{\infty} a_j z^j$, $a_j \in \mathbb{K}$, is a power series with radius of convergence $r \in]0, \infty]$, then the function*

$$f : B_r(0) \longrightarrow \mathbb{K}, \quad f(z) := \sum_{j=0}^{\infty} a_j z^j, \tag{8.17}$$

is continuous. In particular, if $r = \infty$, then f is continuous on $\mathbb{K}$.

Proof. Each partial sum $z \mapsto \sum_{j=0}^{n} a_j z^j$ is a polynomial, i.e. continuous on $\mathbb{K}$. Moreover, if $\zeta \in B_r(0)$, then the power series converges uniformly on $M := B_{|\zeta|}(0)$ by (8.14), i.e. it is continuous at $\zeta \in M$ by Th. 8.4. $\blacksquare$

Example 8.11. (a) For each $\alpha \in \mathbb{R}$, the radius of convergence of $\sum_{n=1}^{\infty} n^\alpha z^n$ is $r = 1$, since

$$\limsup_{n \to \infty} \sqrt[n]{|a_n|} = \lim_{n \to \infty} \sqrt[n]{n^\alpha} = 1, \tag{8.18}$$

which, for each $\alpha \in \mathbb{Z}$, follows from (7.46) and Th. 7.13(a), and, then, for all $\alpha \in \mathbb{R}$ from the Sandwich Th. 7.16.

Let us investigate what can happen for $|z| = r = 1$ for some cases: The series $\sum_{n=1}^{\infty} z^n$ ($\alpha = 0$) is divergent for each $z \in \mathbb{C}$ with $z = 1$ by the observation that $(z^n)_{n \in \mathbb{N}}$ does not converge to 0 for $n \to \infty$ (as $|z^n| = 1$ for each $n \in \mathbb{N}$); the series $\sum_{n=1}^{\infty} n^{-1} z^n$ ($\alpha = -1$) is the harmonic series, i.e. divergent, for $z = 1$, but convergent for $z = -1$ according to Ex. 7.86(a).

(b) The radius of convergence of both $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ and $\sum_{n=0}^{\infty} \frac{z^n}{n^n}$ is $r = \infty$ by (8.16) and (8.15), respectively, since

$$\lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n \to \infty} \frac{(n+1)!}{n!} = \lim_{n \to \infty} (n+1) = \infty, \qquad (8.19a)$$

$$\limsup_{n \to \infty} \sqrt[n]{|a_n|} = \lim_{n \to \infty} \sqrt[n]{\frac{1}{n^n}} = \lim_{n \to \infty} \frac{1}{n} = 0. \qquad (8.19b)$$

(c) The radius of convergence of $\sum_{n=0}^{\infty} n! \, z^n$ is $r = 0$ by (8.16), since

$$\lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n \to \infty} \frac{n!}{(n+1)!} = \lim_{n \to \infty} \frac{1}{n+1} = 0. \qquad (8.20)$$

Caveat 8.12. Theorem 8.9 does *not* claim the uniform convergence of $\sum_{j=0}^{\infty} a_j z^j$ on $B_r(0)$, which is usually not true (e.g., it is an exercise to show that $\sum_{j=0}^{\infty} z^j$ does not converge uniformly on $B_1(0)$). Theorem 8.9 also claims nothing about the convergence or divergence of $\sum_{j=0}^{\infty} a_j z^j$ for $|z| = r$, which has to be determined case by case (cf. Ex. 8.11(a)).

Definition and Remark 8.13. Given two power series $p := \sum_{j=0}^{\infty} a_j z^j$ and $q := \sum_{j=0}^{\infty} b_j z^j$ in $\mathbb{K}$, we define their *Cauchy product*

$$p * q := \sum_{j=0}^{\infty} c_j z^j, \quad \text{where} \quad c_j := \sum_{k=0}^{j} a_k b_{j-k} = a_0 b_j + a_1 b_{j-1} + \cdots + a_j b_0. \qquad (8.21)$$

Note that we have not assumed any convergence of the series so far, i.e. p, q, and $p * q$ are not $\mathbb{K}$-valued functions, but *sequences* of $\mathbb{K}$-valued functions according to Def. 8.5 (sequences of polynomials, actually). Sometimes one also calls the Cauchy product $p * q$ the *convolution* of p and q.

Now, if we do assume p and q to have some nonzero radii of convergence, say $r_p, r_q \in \,]0, \infty]$, respectively, then, by (8.13a), both series are absolutely convergent for each $z \in B_r(0)$, where $r := \min\{r_p, r_q\}$. Thus, the functions

$$f : B_r(0) \longrightarrow \mathbb{K}, \quad f(z) := \sum_{j=0}^{\infty} a_j z^j, \quad g : B_r(0) \longrightarrow \mathbb{K}, \quad g(z) := \sum_{j=0}^{\infty} b_j z^j, \qquad (8.22)$$

are well-defined, and (7.94) implies

$$\underset{z \in B_r(0)}{\forall} \quad f(z)g(z) = \sum_{j=0}^{\infty} c_j z^j \quad \text{with } c_j \text{ as in (8.21).} \qquad (8.23)$$

8.3   Exponential Functions

The notion of power series allows us to extend the definition of exponential functions to complex arguments:

Definition and Remark 8.14. We define the *exponential function*

$$\exp : \mathbb{C} \longrightarrow \mathbb{C}, \quad \exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \tag{8.24}$$

From Ex. 8.11(b), we already know the radius of convergence of the power series in (8.24) is ∞, such that the function in (8.24) is well-defined.

For the time being, we also *redefine* Euler's number as $e := \exp(1) > 1 > 0$ and, for each $x \in \mathbb{R}^+$, $\ln x := \log_{\exp(1)}(x)$. This, as well as calling the function of (8.24) exponential function, will be justified as soon as we will have proved

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots \tag{8.25}$$

and

$$\underset{x \in \mathbb{R}}{\forall} \quad e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \tag{8.26}$$

in (8.36) of Th. 8.18 and in Th. 8.16(c) below, respectively.

Proposition 8.15. *If a continuous function $E : \mathbb{R} \longrightarrow \mathbb{R}$ satisfies*

$$a := E(1) > 0 \quad and \tag{8.27a}$$

$$\underset{x,y \in \mathbb{R}}{\forall} \quad E(x + y) = E(x)E(y), \tag{8.27b}$$

then f is an exponential function – more precisely

$$\underset{x \in \mathbb{R}}{\forall} \quad E(x) = a^x. \tag{8.28}$$

Proof. First, $a = E(1) = E(0 + 1) = E(0)E(1) = E(0)\,a$ and $a > 0$ shows $E(0) = 1$. Then, for each $x \in \mathbb{R}$, $1 = E(0) = E(x - x) = E(x)E(-x)$, i.e. $E(-x) = \big(E(x)\big)^{-1}$, showing $E(x) \neq 0$ for each $x \in \mathbb{R}$. Thus, $E(1) > 0$, the continuity of E, and the intermediate value Th. 7.57 imply $E(x) > 0$ for each $x \in \mathbb{R}$. Next, an induction shows

$$\underset{x \in \mathbb{R}}{\forall} \quad \underset{n \in \mathbb{N}}{\forall} \quad E(n \cdot x) = \big(E(x)\big)^n : \tag{8.29}$$

The base case is trivially true and the induction step is

$$E((n+1)x) = E(nx)E(x) \overset{\text{ind. hyp.}}{=} (E(x))^n \, E(x) = (E(x))^{n+1}.$$

Applying (8.29) with $x = 1$ shows $E(n) = a^n$ for each $n \in \mathbb{N}$. Applying (8.29) with $x = 1/n$, $n \in \mathbb{N}$, shows $a = E(1) = (E(1/n))^n$, i.e. $E(1/n) = a^{1/n}$ since $E(1/n) > 0$. Next,

$$\underset{n,k\in\mathbb{N}}{\forall} \quad E(k/n) \overset{(8.29)}{=} \left(E(1/n)\right)^k = (a^{1/n})^k = a^{\frac{k}{n}},$$

showing (8.28) holds for each $x \in \mathbb{Q}^+$. Then (8.28) also holds for each $x \in \mathbb{R}^+$, since, if $(q_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{Q}^+$ with $\lim_{n\to\infty} q_n = x$, then the continuity of E implies

$$a^x = \lim_{n\to\infty} a^{q_n} = \lim_{n\to\infty} E(q_n) = E(x).$$

Finally, if $x \in \mathbb{R}^-$, then

$$a^x = (a^{-x})^{-1} = \left(E(-x)\right)^{-1} = E(x),$$

completing the proof that (8.28) holds for each $x \in \mathbb{R}$.   ∎

Theorem 8.16. *We consider the exponential function* exp *as defined in (8.24). The following holds:*

(a) exp *is continuous on $\mathbb{C}$.*

(b) $\exp(z + w) = \exp(z)\exp(w)$ *is valid for all $z, w \in \mathbb{C}$.*

(c) *With $e := \exp(1)$ (cf. Def. and Rem. 8.14), it is*

$$\underset{x\in\mathbb{R}}{\forall} \quad e^x = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Proof. (a) holds by Cor. 8.10; for (b), we compute (using (7.94)),

$$\underset{z,w\in\mathbb{C}}{\forall} \left(\begin{array}{l} \exp(z)\exp(w) = \sum_{n=0}^{\infty} c_n, \\[2mm] \text{where} \quad c_n = \sum_{j=0}^{n} \frac{z^j}{j!} \frac{w^{n-j}}{(n-j)!} = \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} z^j w^{n-j} \overset{(5.22)}{=} \frac{(z+w)^n}{n!} \end{array} \right);$$

$$\tag{8.30}$$

and then (c) is an immediate consequence of (a), (b), and Prop. 8.15.   ∎

Definition 8.17. Let $M \subseteq \mathbb{C}$. If $\zeta \in \mathbb{C}$ is a cluster point of M, then a function $f : M \longrightarrow \mathbb{K}$ is said to tend to $\eta \in \mathbb{K}$ (or to have the *limit* $\eta \in \mathbb{K}$) for $z \to \zeta$ (denoted by $\lim_{z\to\zeta} f(z) = \eta$) if, and only if, for each sequence $(z_k)_{k\in\mathbb{N}}$ in $M \setminus \{\zeta\}$ with $\lim_{k\to\infty} z_k = \zeta$, the sequence $(f(z_k))_{k\in\mathbb{N}}$ converges to $\eta \in \mathbb{K}$, i.e.

$$\lim_{z\to\zeta} f(z) = \eta \quad \Leftrightarrow \quad \underset{(z_k)_{k\in\mathbb{N}} \text{ in } M\setminus\{\zeta\}}{\forall} \left(\lim_{k\to\infty} z_k = \zeta \quad \Rightarrow \quad \lim_{k\to\infty} f(z_k) = \eta \right). \tag{8.31}$$

Theorem 8.18. *We consider the exponential function* exp *as defined in* (8.24)*. With* $e^z := \exp(z)$ *for each* $z \in \mathbb{C}$ *and* $\ln x := \log_{\exp(1)}(x)$ *for each* $x \in \mathbb{R}^+$ *(cf. Th. 8.16(c) and Def. and Rem. 8.14), we have the following limits:*

$$\lim_{z \to 0} \frac{e^z - 1}{z} = 1 \quad \left(z \in M := \mathbb{C} \setminus \{0\} \right), \tag{8.32}$$

$$\lim_{x \to 0} \frac{\ln(1 + x)}{x} = 1 \quad \left(x \in M :=]-1, \infty[\setminus \{0\} \right), \tag{8.33}$$

$$\underset{\xi \in \mathbb{R}}{\forall} \lim_{x \to 0} \ln(1 + \xi x)^{\frac{1}{x}} = \xi \quad \left(x \in M := \{x \in \mathbb{R} : 1 + \xi x > 0\} \setminus \{0\} \right), \tag{8.34}$$

$$\underset{\xi \in \mathbb{R}}{\forall} \lim_{x \to 0} (1 + \xi x)^{\frac{1}{x}} = e^\xi \quad \left(x \in M := \{x \in \mathbb{R} : 1 + \xi x > 0\} \setminus \{0\} \right), \tag{8.35}$$

$$\underset{x \in \mathbb{R}}{\forall} \lim_{n \to \infty} \left(1 + \frac{x}{n} \right)^n = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}. \tag{8.36}$$

Proof. (8.32): From (8.24) and $e^z = \exp(z)$, we obtain

$$\underset{z \neq 0}{\forall} \quad \frac{e^z - 1}{z} = \sum_{n=0}^{\infty} \frac{z^n}{(n+1)!} = 1 + \frac{z}{2!} + \frac{z^2}{3!} + \cdots,$$

which, since $z \mapsto \sum_{n=0}^{\infty} \frac{z^n}{(n+1)!}$ is continuous on $\mathbb{C}$ by Cor. 8.10, implies (8.32).

(8.33): Consider the auxiliary function $f :]-1, \infty[\longrightarrow \mathbb{R}$, $f(x) := \ln(x+1)$, with $f^{-1}(x) = e^x - 1$. Now, given a sequence $(x_k)_{k \in \mathbb{N}}$ in $]-1, \infty[\setminus \{0\}$ with $\lim_{k \to \infty} x_k = 0$, one obtains

$$\lim_{k \to \infty} \frac{\ln(1 + x_k)}{x_k} = \lim_{k \to \infty} \frac{\ln\left(1 + f^{-1}(f(x_k))\right)}{f^{-1}(f(x_k))} = \lim_{k \to \infty} \frac{\ln\left(1 + e^{f(x_k)} - 1\right)}{e^{f(x_k)} - 1}$$

$$= \lim_{k \to \infty} \frac{f(x_k)}{e^{f(x_k)} - 1} \overset{(8.32)}{=} 1,$$

where, in the last step, it was used that $\lim_{k \to \infty} x_k = 0$ and the continuity of f implies $\lim_{k \to \infty} f(x_k) = \ln 1 = 0$.

Similarly, but simpler, one obtains (8.34) and (8.35) (exercise). Finally, for the sequence $(x_n)_{n \in \mathbb{N}}$ with $x_n := 1/n$, (8.35) implies (8.36). ∎

Definition 8.19 (Exponentiation with Complex Exponents)**.** For each $(a, z) \in \mathbb{R}^+ \times \mathbb{C}$, we define

$$a^z := \exp(z \ln a), \tag{8.37}$$

where exp is the function defined in (8.24). For $a = e$, (8.37) yields $e^z = \exp(z)$, i.e. (8.37) is consistent with (8.26).

Theorem 8.20. (a) *The first two exponentiation rules of* (7.54) *still hold for each* $a, b > 0$ *and each* $z, w \in \mathbb{C}$:

$$a^{z+w} = a^z \, a^w, \tag{8.38a}$$

$$a^z \, b^z = (ab)^z. \tag{8.38b}$$

(b) *For each $a \in \mathbb{R}^+$, the exponential function*

$$f : \mathbb{C} \longrightarrow \mathbb{C}, \quad f(z) := a^z, \tag{8.39a}$$

is continuous, and, for each $\zeta \in \mathbb{C}$, the power function

$$g : \mathbb{R}^+ \longrightarrow \mathbb{C}, \quad g(x) := x^\zeta, \tag{8.39b}$$

is continuous.

(c) *The limit in (8.36) extends to complex numbers:*

$$\underset{z \in \mathbb{C}}{\forall} \quad \lim_{n \to \infty} \left(1 + \frac{z}{n}\right)^n = e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}. \tag{8.40}$$

Proof. (a): We compute

$$a^{z+w} \overset{(8.37)}{=} \exp((z+w) \ln a) = \exp(z \ln a + w \ln a)$$
$$\overset{\text{Th. } 8.16(b)}{=} \exp(z \ln a) \exp(w \ln a) \overset{(8.37)}{=} a^z \, a^w,$$

proving (8.38a), and

$$a^z \, b^z \overset{(8.37)}{=} \exp(z \ln a) \exp(z \ln b) \overset{\text{Th. } 8.16(b)}{=} \exp(z \ln a + z \ln b)$$
$$\overset{(7.65e)}{=} \exp\left(z \ln(ab)\right) \overset{(8.37)}{=} (ab)^z,$$

proving (8.38b).

(b): The continuity of both functions follows from the continuity of exp (according to Th. 8.16(a)) and from the fact that continuity is preserved by compositions (according to Th. 7.41): The exponential function f, given by $f(z) = e^{z \ln a}$, is the composition of the continuous functions $z \mapsto z \ln a$ and $w \mapsto e^w$, whereas (analogous to Ex. 7.76(a)), the power function g, given by $g(x) = e^{\zeta \ln x}$, is the composition $g = \exp \circ (\zeta \ln)$, where ln is continuous by Cor. 7.74.

(c): Exercise.   ∎

8.4   Trigonometric Functions

The first "definition" of the trigonometric functions sine and cosine is the one based on geometric visualization usually given in high school: $\cos x$ and $\sin x$ are the coordinates of the point $p = (p_1, p_2) \in \mathbb{R}^2$ on the unit circle, such that x is the angle measured in radian between the line segment between $(0,0)$ and $(1,0)$ and the line segment between $(0,0)$ and p.

While this "definition" allows to obtain many important properties of sine and cosine using geometric arguments, it is not mathematically rigorous, and, for example, provides

no clue how to compute values like $\sin 1$. The problem is related to the fact that the angle measured in radian between the line segment between $(0,0)$ and $(1,0)$ and the line segment between $(0,0)$ and p is supposed to be the length of the segment of the unit circle between $(1,0)$ and p (taken in the counter-clockwise direction).

In the following Def. and Rem. 8.21, we will provide a mathematically rigorous definition of sine and cosine using power series, and we will then verify that the functions have the familiar properties one learns in high school. However, as the computation of lengths of curved paths is actually beyond the scope of this lecture, we will not be able to see that our sine and cosine functions are precisely the same we visualized in high school (the interested reader is referred to Ex. 1 in Sec. 5.14 of [Wal02]).

Definition and Remark 8.21. We define the *sine function*, denoted sin, and the *cosine function*, denoted cos by

$$\sin : \mathbb{C} \longrightarrow \mathbb{C}, \qquad \sin z := \sum_{n=0}^{\infty} \frac{(-1)^n \, z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - + \dots, \qquad (8.41\text{a})$$

$$\cos : \mathbb{C} \longrightarrow \mathbb{C}, \qquad \cos z := \sum_{n=0}^{\infty} \frac{(-1)^n \, z^{2n}}{(2n)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - + \dots. \qquad (8.41\text{b})$$

(a) sin and cos are well-defined and continuous: For both series and each $z \in \mathbb{C}$, we can estimate the absolute value of the nth summand by the nth summand of the series for the exponential function $e^{|z|}$ (cf. (8.36)), which we know to be convergent from Ex. 8.11(b). Thus, by Th. 8.9, both series in (8.41) have radius of convergence ∞ and are continuous by Cor. 8.10.

(b) $\cos : \mathbb{R} \longrightarrow \mathbb{R}$ (i.e. $\cos\!\restriction_{\mathbb{R}}$) has a smallest positive zero $\alpha \in \mathbb{R}^+$. We *define* $\pi := 2\alpha$. One can show π is an irrational number (see Appendix H.2) and its first digits are $\pi = 3.14159\dots$

To see cos has a smallest positive zero and to obtain a first (very coarse) estimate, note

$$\underset{x \in \mathbb{R}^+}{\forall}\ \underset{k \in \mathbb{N}}{\forall}\ \left(\frac{x^k}{k!} > \frac{x^{k+1}}{(k+1)!} \quad \Leftrightarrow \quad 1 > \frac{x}{k+1} \quad \Leftrightarrow \quad k+1 > x \right),$$

showing $\frac{x^k}{k!} > \frac{x^{k+1}}{(k+1)!}$ holds for each $k \geq 2$ and each $x \in\ ]0,3[$. In particular, the summands of the series in (8.41) converge monotonically to 0 (for $k \geq 2$) and, since the series are alternating for $x \neq 0$, Th. 7.85 applies and (7.79) yields

$$\underset{0<x<3}{\forall}\ \left(\begin{array}{l} f(x) := 1 - \dfrac{x^2}{2} < \cos x < 1 - \dfrac{x^2}{2} + \dfrac{x^4}{24} =: g(x), \\[2mm] x - \dfrac{x^3}{6} < \sin x < x - \dfrac{x^3}{6} + \dfrac{x^5}{120}. \end{array} \right) \qquad (8.42)$$

The zeros of $x \mapsto f(x)$ are $-\sqrt{2}, \sqrt{2}$, i.e. $\sqrt{2}$ is its smallest positive zero; the zeros of $x \mapsto g(x)$ are $-\sqrt{6-2\sqrt{3}}, -\sqrt{6+2\sqrt{3}}, \sqrt{6-2\sqrt{3}}, \sqrt{6+2\sqrt{3}}$, i.e. $\sqrt{6-2\sqrt{3}}$

is its smallest positive zero. Thus, as $f(0) = g(0) = 1$, the intermediate value Th. 7.57 implies cos has a smallest positive zero α and

$$1.4 < \sqrt{2} < \frac{\pi}{2} := \alpha < \sqrt{6 - 2\sqrt{3}} < 1.6 \tag{8.43}$$

Theorem 8.22. *We have the following identities:*

$$\sin 0 = 0, \quad \cos 0 = 1, \tag{8.44a}$$

$$\underset{z \in \mathbb{C}}{\forall} \quad \sin z = -\sin(-z), \quad \cos z = \cos(-z), \tag{8.44b}$$

$$\underset{z,w \in \mathbb{C}}{\forall} \quad \sin(z + w) = \sin z \cos w + \cos z \sin w, \tag{8.44c}$$

$$\underset{z,w \in \mathbb{C}}{\forall} \quad \cos(z + w) = \cos z \cos w - \sin z \sin w, \tag{8.44d}$$

$$\underset{z \in \mathbb{C}}{\forall} \quad (\sin z)^2 + (\cos z)^2 = 1, \tag{8.44e}$$

$$\cos \frac{\pi}{2} = 0, \quad \sin \frac{\pi}{2} = 1, \quad \underset{x \in [0, \frac{\pi}{2}[}{\forall} \cos x > 0, \tag{8.44f}$$

$$\underset{z \in \mathbb{C}}{\forall} \quad \sin\left(z + \frac{\pi}{2}\right) = \cos z, \quad \cos\left(z + \frac{\pi}{2}\right) = -\sin z, \tag{8.44g}$$

$$\underset{z \in \mathbb{C}}{\forall} \quad \sin(z + \pi) = -\sin z, \quad \cos(z + \pi) = -\cos z, \tag{8.44h}$$

$$\underset{z \in \mathbb{C}}{\forall} \quad \sin(z + 2\pi) = \sin z, \quad \cos(z + 2\pi) = \cos z, \tag{8.44i}$$

$$\lim_{z \to 0} \frac{\sin z}{z} = 1, \quad \lim_{z \to 0} \frac{\cos z - 1}{z^2} = -\frac{1}{2}. \tag{8.44j}$$

Identities (8.44i) *can be restated as sine and cosine being periodic functions with period* 2π.

Proof. (8.44a) is immediate from (8.41) since, for $z = 0$, all summands of the sine series are 0 and all summands of the cosine series are 0, except the first one, which is $\frac{(-1)^0 \, 0^0}{0!} = 1$.

(8.44b) is also immediate from (8.41), since $(-z)^{2n+1} = (-1)^{2n+1} z^{2n+1} = -z^{2n+1}$ and $(-z)^{2n} = (-1)^{2n} z^{2n} = z^{2n}$.

(8.44c) and (8.44d) can be verified using the Cauchy product: According to (7.94),

$$\underset{z,w \in \mathbb{C}}{\forall} \left(\begin{array}{l} \sin z \cos w = \displaystyle\sum_{n=0}^{\infty} c_n, \quad \cos z \sin w = \displaystyle\sum_{n=0}^{\infty} d_n, \\[2mm] \text{where} \quad c_n = \displaystyle\sum_{j=0}^{n} \frac{(-1)^j \, z^{2j+1}}{(2j+1)!} \frac{(-1)^{n-j} \, w^{2(n-j)}}{(2(n-j))!}, \\[2mm] \quad d_n = \displaystyle\sum_{j=0}^{n} \frac{(-1)^j \, z^{2j}}{(2j)!} \frac{(-1)^{n-j} \, w^{2(n-j)+1}}{(2(n-j)+1)!}, \end{array} \right),$$

that means, for each $z, w \in \mathbb{C}$,

$$
\begin{aligned}
c_n + d_n &= \sum_{j=0}^{n} \frac{(-1)^n \, z^{2j+1}}{(2j+1)!} \frac{w^{2(n-j)}}{(2(n-j))!} + \sum_{j=0}^{n} \frac{(-1)^n \, z^{2j}}{(2j)!} \frac{w^{2(n-j)+1}}{(2(n-j)+1)!} \\
&= \sum_{j=0}^{n} \frac{(-1)^n \, z^{2j+1}}{(2j+1)!} \frac{w^{2n+1-(2j+1)}}{(2n+1-(2j+1))!} + \sum_{j=0}^{n} \frac{(-1)^n \, z^{2j}}{(2j)!} \frac{w^{2n+1-2j}}{(2n+1-2j)!} \\
&= (-1)^n \sum_{j=0}^{2n+1} \frac{z^j}{j!} \frac{w^{2n+1-j}}{(2n+1-j)!} = \frac{(-1)^n}{(2n+1)!} \sum_{j=0}^{2n+1} \binom{2n+1}{j} z^j \, w^{2n+1-j} \\
&= \frac{(-1)^n \, (z+w)^{2n+1}}{(2n+1)!},
\end{aligned}
$$

proving (8.44c). Similarly, according to (7.94),

$$
\underset{z,w\in\mathbb{C}}{\forall} \left(\begin{array}{l} \cos z \cos w = \displaystyle\sum_{n=0}^{\infty} c_n, \quad \sin z \sin w = \sum_{n=0}^{\infty} d_n, \\[2mm] \text{where} \quad c_n = \displaystyle\sum_{j=0}^{n} \frac{(-1)^j \, z^{2j}}{(2j)!} \frac{(-1)^{n-j} \, w^{2(n-j)}}{(2(n-j))!}, \\[2mm] \qquad\qquad d_n = \displaystyle\sum_{j=0}^{n} \frac{(-1)^j \, z^{2j+1}}{(2j+1)!} \frac{(-1)^{n-j} \, w^{2(n-j)+1}}{(2(n-j)+1)!}, \end{array} \right),
$$

that means, for each $z, w \in \mathbb{C}$,

$$c_0 = 1 \quad \text{and}$$

$$
\begin{aligned}
\underset{n\in\mathbb{N}}{\forall} \quad c_n - d_{n-1} &= \sum_{j=0}^{n} \frac{(-1)^n \, z^{2j}}{(2j)!} \frac{w^{2(n-j)}}{(2(n-j))!} - \sum_{j=0}^{n-1} \frac{(-1)^{n-1} \, z^{2j+1}}{(2j+1)!} \frac{w^{2(n-1-j)+1}}{(2(n-1-j)+1)!} \\
&= \sum_{j=0}^{n} \frac{(-1)^n \, z^{2j}}{(2j)!} \frac{w^{2n-2j}}{(2n-2j)!} + \sum_{j=0}^{n-1} \frac{(-1)^n \, z^{2j+1}}{(2j+1)!} \frac{w^{2n-(2j+1)}}{(2n-(2j+1))!} \\
&= (-1)^n \sum_{j=0}^{2n} \frac{z^j}{j!} \frac{w^{2n-j}}{(2n-j)!} = \frac{(-1)^n}{(2n)!} \sum_{j=0}^{2n} \binom{2n}{j} z^j \, w^{2n-j} \\
&= \frac{(-1)^n \, (z+w)^{2n}}{(2n)!},
\end{aligned}
$$

proving (8.44d).

(8.44e): One computes for each $z \in \mathbb{C}$:

$$
(\sin z)^2 + (\cos z)^2 = \cos z \cos(-z) - \sin z \sin(-z) \overset{(8.44\mathrm{d})}{=} \cos(z - z) = \cos 0 = 1.
$$

(8.44f): $\cos \frac{\pi}{2} = 0$ and $\cos x > 0$ for $0 \leq x < \frac{\pi}{2}$ hold according to the definition of π in Def. and Rem. 8.21(b). Then

$$
\left(\sin \frac{\pi}{2}\right)^2 \overset{(8.44\mathrm{e})}{=} 1 - \left(\cos \frac{\pi}{2}\right)^2 = 1 \quad \text{and} \quad \sin \frac{\pi}{2} > \frac{\pi}{2} - \frac{(\pi/2)^3}{6} \overset{(8.43)}{>} 1.4 - \frac{(1.6)^3}{6} > 0.7 > 0.
$$

(8.44g) is immediate from (8.44c), (8.44d), and (8.44f).

(8.44h): One obtains

$$\sin \pi \quad = \quad \sin\left(\frac{\pi}{2} + \frac{\pi}{2}\right) \overset{(8.44c)}{=} 1 \cdot 0 + 0 \cdot 1 = 0,$$

$$\cos \pi \quad = \quad \cos\left(\frac{\pi}{2} + \frac{\pi}{2}\right) \overset{(8.44d)}{=} 0 \cdot 0 - 1 \cdot 1 = -1,$$

$$\underset{z \in \mathbb{C}}{\forall} \quad \sin(z + \pi) \overset{(8.44c)}{=} -\sin z + 0 = -\sin z,$$

$$\underset{z \in \mathbb{C}}{\forall} \quad \cos(z + \pi) \overset{(8.44d)}{=} -\cos z + 0 = -\cos z.$$

(8.44i): One obtains

$$\sin(2\pi) \quad = \quad \sin(\pi + \pi) \overset{(8.44c)}{=} 0 + 0 = 0,$$

$$\cos(2\pi) \quad = \quad \cos(\pi + \pi) \overset{(8.44d)}{=} (-1)(-1) - 0 = 1,$$

$$\underset{z \in \mathbb{C}}{\forall} \quad \sin(z + 2\pi) \overset{(8.44c)}{=} \sin z + 0 = \sin z,$$

$$\underset{z \in \mathbb{C}}{\forall} \quad \cos(z + 2\pi) \overset{(8.44d)}{=} \cos z - 0 = \cos z.$$

(8.44j): One obtains

$$\underset{z \in \mathbb{C} \setminus \{0\}}{\forall} \left(\begin{array}{l} \dfrac{\sin z}{z} = \displaystyle\sum_{n=0}^{\infty} \dfrac{(-1)^n \, z^{2n}}{(2n+1)!} = 1 - \dfrac{z^2}{3!} + \dfrac{z^4}{5!} - + \ldots, \\[2ex] \dfrac{\cos z - 1}{z^2} = \displaystyle\sum_{n=0}^{\infty} \dfrac{(-1)^{n+1} \, z^{2n}}{(2(n+1))!} = -\dfrac{1}{2!} + \dfrac{z^2}{4!} - \dfrac{z^4}{6!} + - \ldots \end{array} \right).$$

For both series on the right-hand side and each $z \in \mathbb{C}$, we can estimate the absolute value of each summand by the corresponding summand of the exponential series for $e^{|z|}$ (cf. (8.36)), showing they have radius of convergence ∞ and are continuous by Cor. 8.10. In particular, their continuity in $z = 0$ proves (8.44j). ∎

Theorem 8.23. *One has* $\sin(\mathbb{R}) = \cos(\mathbb{R}) = [-1, 1]$*, i.e. the range of both sine and cosine is* $[-1, 1]$*. Moreover, for each* $k \in \mathbb{Z}$*:*

$$\sin \text{ is strictly increasing on} \quad \left[-\frac{\pi}{2} + 2k\pi, \frac{\pi}{2} + 2k\pi\right], \tag{8.45a}$$

$$\sin \text{ is strictly decreasing on} \quad \left[\frac{\pi}{2} + 2k\pi, \frac{3\pi}{2} + 2k\pi\right], \tag{8.45b}$$

$$\cos \text{ is strictly increasing on} \quad [(2k-1)\pi, 2k\pi], \tag{8.45c}$$

$$\cos \text{ is strictly decreasing on} \quad [2k\pi, (2k+1)\pi], \tag{8.45d}$$

which, due to (8.44e), can be summarized (and visualized) by saying that, if x *runs from* $2k\pi$ *to* $2(k+1)\pi$*, then* $(\cos x, \sin x)$ *runs once counterclockwise through the unit circle, starting at* $(1, 0)$*.*

Proof. From (8.44e), we know $\sin(\mathbb{R}) \subseteq [-1,1]$ and $\cos(\mathbb{R}) \subseteq [-1,1]$. As

$$\sin \frac{\pi}{2} \stackrel{(8.44f)}{=} 1, \quad \sin\left(-\frac{\pi}{2}\right) \stackrel{(8.44b)}{=} -1, \quad \cos 0 \stackrel{(8.44a)}{=} 1, \quad \cos \pi \stackrel{(8.44h)}{=} -\cos 0 = -1,$$

the continuity of sine and cosine together with the intermediate value Th. 7.57 implies $\sin(\mathbb{R}) = \cos(\mathbb{R}) = [-1,1]$.

From (8.42), we know $0 < x - \frac{x^3}{6} < \sin x$ and $\cos x < 1 - \frac{x^2}{2} + \frac{x^4}{24} < 1$ for each $x \in]0, \frac{\pi}{2}]$, implying

$$\forall_{0 \leq x < x+y \leq \frac{\pi}{2}} \quad \cos(x+y) = \cos x \cos y - \sin x \sin y \leq \cos x \cos y < \cos x,$$

showing cos is strictly decreasing on $[0, \frac{\pi}{2}]$. Then cos is strictly increasing on $[-\frac{\pi}{2}, 0]$ by (8.44b), sin is strictly increasing on $[0, \frac{\pi}{2}]$ and strictly decreasing on $[\frac{\pi}{2}, \pi]$ by (8.44g), implying sin is strictly increasing on $[-\frac{\pi}{2}, 0]$ and strictly decreasing on $[-\pi, -\frac{\pi}{2}]$ by (8.44b), i.e. sin is strictly increasing on $[\frac{3\pi}{2}, 2\pi]$ and strictly decreasing on $[\pi, \frac{3\pi}{2}]$ by (8.44i), implying cos is strictly decreasing on $[\frac{\pi}{2}, \pi]$ and strictly increasing on $[-\pi, -\frac{\pi}{2}]$ by (8.44g). Since this fixes the monotonicity properties of both sine and cosine over more than one period, the general statements in (8.45) are provided by (8.44i). ∎

We now come to important complex number relations between sine, cosine, and the exponential function.

Theorem 8.24. *One has the following formulas, relating the (complex) sine, cosine, and exponential function:*

$$\forall_{z \in \mathbb{C}} \quad e^{iz} = \cos z + i \sin z \quad \text{(Euler formula)}, \tag{8.46a}$$

$$\forall_{z \in \mathbb{C}} \quad \cos z = \frac{e^{iz} + e^{-iz}}{2}, \tag{8.46b}$$

$$\forall_{z \in \mathbb{C}} \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}. \tag{8.46c}$$

Proof. Let $z \in \mathbb{C}$. For (8.46a), one computes

$$e^{iz} \stackrel{(8.37),(8.24)}{=} \sum_{n=0}^{\infty} \frac{(iz)^n}{n!} = \sum_{n=0}^{\infty} \left(\frac{(-1)^n z^{2n}}{(2n)!} + \frac{i (-1)^n z^{2n+1}}{(2n+1)!} \right)$$

$$\stackrel{\text{Th. 7.95(a)}}{=} \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!} \stackrel{(8.41)}{=} \cos z + i \sin z.$$

Then

$$e^{iz} + e^{-iz} \stackrel{(8.46a)}{=} \cos z + i \sin z + \cos(-z) + i \sin(-z) = 2 \cos z$$

proves (8.46b), and

$$e^{iz} - e^{-iz} \stackrel{(8.46a)}{=} \cos z + i \sin z - \cos(-z) - i \sin(-z) = 2i \sin z$$

proves (8.46c). ∎

As a first application of (8.46), we can now determine all solutions to the equation $e^z = 1$ and all zeros (if any) of exp, sin, and cos:

Theorem 8.25. *The set of (complex) solutions to the equation $e^z = 1$ consists precisely of all integer multiples of $2\pi i$, the exponential function has no zeros (neither in $\mathbb{R}$ nor in $\mathbb{C}$), and the set of all (real or complex) zeros of sine and cosine consists of a discrete set of real numbers. More precisely:*

$$\exp^{-1}\{1\} = \{2k\pi i : k \in \mathbb{Z}\}, \tag{8.47a}$$
$$\exp^{-1}\{0\} = \emptyset, \tag{8.47b}$$
$$\sin^{-1}\{0\} = \{k\pi : k \in \mathbb{Z}\}, \tag{8.47c}$$
$$\cos^{-1}\{0\} = \left\{(2k+1)\tfrac{\pi}{2} : k \in \mathbb{Z}\right\}. \tag{8.47d}$$

Proof. Exercise.   ∎

Definition and Remark 8.26. We define *tangent* and *cotangent* by

$$\tan : \quad \underbrace{\mathbb{C} \setminus \cos^{-1}\{0\}}_{\mathbb{C} \setminus \{(2k+1)\frac{\pi}{2} \,:\, k \in \mathbb{Z}\} \text{ by } (8.47d)} \quad \longrightarrow \mathbb{C}, \qquad \tan z := \frac{\sin z}{\cos z}, \tag{8.48a}$$

$$\cot : \quad \underbrace{\mathbb{C} \setminus \sin^{-1}\{0\}}_{\mathbb{C} \setminus \{k\pi \,:\, k \in \mathbb{Z}\} \text{ by } (8.47c)} \quad \longrightarrow \mathbb{C}, \qquad \cot z := \frac{\cos z}{\sin z}, \tag{8.48b}$$

respectively. Since sine and cosine are both continuous, tangent and cotangent are also both continuous on their respective domains. Both functions have period π, since, for each z in the respective domains,

$$\tan(z+\pi) = \frac{\sin(z+\pi)}{\cos(z+\pi)} \stackrel{(8.44h)}{=} \frac{-\sin z}{-\cos z} = \tan z, \quad \cot(z+\pi) \stackrel{(8.44h)}{=} \frac{-\cos z}{-\sin z} = \cot z. \tag{8.49}$$

Since

$$\lim_{n\to\infty} \sin\left(\frac{\pi}{2} - \frac{1}{n}\right) = \sin\frac{\pi}{2} = 1 \quad \wedge \quad \lim_{n\to\infty} \cos\left(\frac{\pi}{2} - \frac{1}{n}\right) = \cos\frac{\pi}{2} = 0$$

$$\wedge \quad \cos\left(\frac{\pi}{2} - \frac{1}{n}\right) > 0 \quad \Rightarrow \quad \lim_{n\to\infty} \tan\left(\frac{\pi}{2} - \frac{1}{n}\right) = \infty,$$

$$\lim_{n\to\infty} \sin\left(-\frac{\pi}{2} + \frac{1}{n}\right) = \sin\left(-\frac{\pi}{2}\right) = -1 \quad \wedge \quad \lim_{n\to\infty} \cos\left(-\frac{\pi}{2} + \frac{1}{n}\right) = \cos\left(-\frac{\pi}{2}\right) = 0$$

$$\wedge \quad \cos\left(-\frac{\pi}{2} + \frac{1}{n}\right) > 0 \quad \Rightarrow \quad \lim_{n\to\infty} \tan\left(-\frac{\pi}{2} + \frac{1}{n}\right) = -\infty,$$

$$\lim_{n\to\infty} \sin\frac{1}{n} = \sin 0 = 0 \quad \wedge \quad \lim_{n\to\infty} \cos\frac{1}{n} = \cos 0 = 1 \quad \wedge \quad \sin\frac{1}{n} > 0$$

$$\Rightarrow \quad \lim_{n\to\infty} \cot\frac{1}{n} = \infty,$$

$$\lim_{n\to\infty} \sin\left(\pi - \frac{1}{n}\right) = \sin\pi = 0 \quad \wedge \quad \lim_{n\to\infty} \cos\left(\pi - \frac{1}{n}\right) = \cos\pi = -1$$

$$\wedge \quad \sin\left(\pi - \frac{1}{n}\right) > 0 \quad \Rightarrow \quad \lim_{n\to\infty} \cot\left(\pi - \frac{1}{n}\right) = -\infty,$$

we obtain $\tan(\mathbb{R} \setminus \cos^{-1}\{0\}) = \cot(\mathbb{R} \setminus \sin^{-1}\{0\}) = \mathbb{R}$.

For each $k \in \mathbb{Z}$,

$$\text{tan is strictly increasing on} \quad \left]-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi\right[, \tag{8.50a}$$

$$\text{cot is strictly decreasing on} \quad \left]k\pi, (k+1)\pi\right[: \tag{8.50b}$$

On $]0, \frac{\pi}{2}[$, sin is strictly increasing and cos is strictly decreasing, i.e. tan is strictly increasing and cot is strictly decreasing. Since $\tan(-x) = \sin(-x)/\cos(-x) = -\tan(x)$, on $]-\frac{\pi}{2}, 0[$, tan is strictly increasing and cot is strictly decreasing. Taking into account the signs of tan and cot on the respective intervals and their π-periodicity according to (8.49) proves (8.50).

Definition and Remark 8.27. Since we have seen sin to be strictly increasing on $[-\frac{\pi}{2}, \frac{\pi}{2}]$ with range $[-1, 1]$, cos to be strictly decreasing one $[0, \pi]$ with range $[-1, 1]$, tan to be strictly increasing on $]-\frac{\pi}{2}, \frac{\pi}{2}[$ with range $\mathbb{R}$, and cot to be strictly decreasing one $]0, \pi[$ with range $\mathbb{R}$; and since all four functions are continuous, Th. 7.60 implies the existence of inverse functions, denoted by

$$\arcsin : [-1, 1] \longrightarrow [-\pi/2, \pi/2], \tag{8.51a}$$

$$\arccos : [-1, 1] \longrightarrow [0, \pi], \tag{8.51b}$$

$$\arctan : \mathbb{R} \longrightarrow]-\pi/2, \pi/2[, \tag{8.51c}$$

$$\text{arccot} : \mathbb{R} \longrightarrow]0, \pi[, \tag{8.51d}$$

respectively, where all four inverse functions are continuous, arcsin is strictly increasing, arccos is strictly decreasing, arctan is strictly increasing, and arccot is strictly decreasing.

Of course, using (8.45) and (8.50), respectively, one can also obtain the inverse functions on different intervals, and, in the literature, such inverse functions are, indeed, considered as well. Somewhat confusingly, it is common to denote all these different functions by the same symbols, namely the ones introduced in (8.51). Here, we will not need to pursue this any further, i.e. we will only consider the inverse functions precisely as defined in (8.51), which are also known as the *principle* inverse functions of sin, cos, tan, and cot, respectively.

8.5 Polar Form of Complex Numbers, Fundamental Theorem of Algebra

Theorem 8.28. *For each complex number $z \in \mathbb{C}$, there exist real numbers $r \geq 0$ and $\varphi \in \mathbb{R}$ such that*

$$z = r\, e^{i\varphi}. \tag{8.52}$$

Moreover, if (8.52) holds with $r \geq 0$ and $\varphi \in \mathbb{R}$, then r is the modulus of z and, for $z \neq 0$, φ is uniquely determined up to addition of an integer multiple of 2π, i.e.

$$\underset{z \in \mathbb{C} \setminus \{0\}}{\forall} \left(z = r\, e^{i\varphi_1} = r\, e^{i\varphi_2} \wedge r \geq 0 \quad \Rightarrow \quad r = |z| \wedge \underset{k \in \mathbb{Z}}{\exists}\ \varphi_1 - \varphi_2 = 2\pi k \right). \quad (8.53)$$

Proof. For $z = 0$, there is nothing to prove, so we assume $z \neq 0$ and set $r := |z|$. We write $z = x + iy$ with $x, y \in \mathbb{R}$, first assuming $y \geq 0$. Then

$$\frac{z}{r} = \xi + i\eta, \quad \text{where} \quad \xi = \frac{x}{r}, \quad \eta = \frac{y}{r} \geq 0, \quad \xi^2 + \eta^2 = 1. \quad (8.54)$$

In particular, $-1 \leq \xi \leq 1$. Thus, letting

$$\varphi := \arccos \xi,$$

we obtain $\varphi \in [0, \pi]$, $\xi = \cos \varphi$, and $\sin \varphi \geq 0$, yielding

$$\sin \varphi = \sqrt{1 - (\cos \varphi)^2} = \sqrt{1 - \xi^2} \overset{(8.54)}{=} \eta.$$

In consequence,

$$\frac{z}{r} = \xi + i\eta = \cos \varphi + i \sin \varphi \overset{(8.46a)}{=} e^{i\varphi},$$

as desired. If $y \leq 0$, then the above shows the existence of $\psi \in \mathbb{R}$ such that $\bar{z} = x - iy = re^{i\psi} = r \cos \psi + ir \sin \psi$. Letting $\varphi := -\psi$, we, once again, have $z = r \cos \psi - ir \sin \psi = re^{-i\psi} = re^{i\varphi}$, as desired, completing the existence proof for the representation (8.52). Now assume (8.52) holds with $r \geq 0$. Then

$$|z| = r|e^{i\varphi}| = r\sqrt{(\sin \varphi)^2 + (\cos \varphi)^2} = r.$$

Finally, if $r\, e^{i\varphi_1} = r\, e^{i\varphi_2}$ with $r > 0$, then $e^{i(\varphi_1 - \varphi_2)} = 1$, i.e. $i(\varphi_1 - \varphi_2) \in \{2k\pi i : k \in \mathbb{Z}\}$ by (8.47a). ∎

Definition and Remark 8.29. The representation of $z \in \mathbb{C}$ given by (8.52) is called its *polar form*, where (r, φ) are also called *polar coordinates* of z, φ is called an *argument* of z. For $z \neq 0$, one can fix the argument uniquely by the additional requirement $\varphi \in [0, 2\pi[$ (but one also finds other choices, for example $\varphi \in\,]-\pi, \pi]$, in the literature). The above terminology is consistent with the common use of calling (r, φ) *polar coordinates* of the vector $z = (x, y) \in \mathbb{R}^2 (= \mathbb{C})$ (in contrast to the *Cartesian coordiantes* (x, y)), where r constitutes the distance of the point $z = (x, y)$ from the origin $(0, 0)$ and φ is the angle between the vector $z = (x, y)$ and the x-axis (cf. the three introductory paragraphs of the previous Sec. 8.4). As promised, we can now better understand the geometric interpretation of complex multiplication already described in Rem. 5.10: If $z_1 = r_1 e^{i\varphi_1}$ and $z_2 = r_2 e^{i\varphi_2}$, then $z_1 z_2 = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}$, i.e. complex multiplication, indeed, means multiplying absolute values and adding arguments.

Corollary 8.30. *If $z \in \mathbb{C}$, then $|z| = 1$ holds if, and only if, there exists $\varphi \in \mathbb{R}$ such that $z = e^{i\varphi}$ – in other words, the map*

$$f : \mathbb{R} \longrightarrow \{z \in \mathbb{C} : |z| = 1\}, \quad f(\varphi) := e^{i\varphi}, \quad (8.55)$$

is surjective. Moreover $f(\varphi_1) = f(\varphi_2)$ holds if, and only if, $\varphi_1 - \varphi_2 = 2\pi k$ for some $k \in \mathbb{Z}$.

Proof. Everything is immediate from Th. 8.28. ■

Corollary 8.31 (Roots of Unity)**.** *For each $n \in \mathbb{N}$, the equation $z^n = 1$ has precisely n distinct solutions $\zeta_1, \ldots, \zeta_n \in \mathbb{C}$, where*

$$\underset{k=1,\ldots,n}{\forall} \quad \zeta_k := e^{k2\pi i/n} \overset{(8.46a)}{=} \cos \frac{k2\pi}{n} + i \sin \frac{k2\pi}{n} = \zeta_1^k. \tag{8.56}$$

The numbers $\zeta_1, \ldots, \zeta_n$ defined in (8.56) are called the nth roots of unity.

Proof. It is $\zeta_k^n = e^{k2\pi i} = 1$ for each $k \in \{1, \ldots, n\}$ and the $\zeta_1, \ldots, \zeta_n$ are all distinct by Cor. 8.30, since, for $k, l \in \{1, \ldots, n\}$ with $k \neq l$, $(k - l)/n \notin \mathbb{Z}$. As $\zeta_1, \ldots, \zeta_n$ are n distinct zeros of the polynomial $P : \mathbb{C} \longrightarrow \mathbb{C}$, $P(z) := z^n - 1$, and P has at most n zeros by Th. 6.6(a), $\zeta_1, \ldots, \zeta_n$ constitute all solutions to $z^n = 1$. ■

We are now in a position to prove one of the central results of analysis and algebra, namely the *fundamental theorem of algebra.* The following proof does not need any tools beyond the ones provided by this class – it is actually mainly founded on continuous functions attaining a min and a max on compact sets according to Th. 7.54 and the existence of nth roots of unity according to Cor. 8.31.

Theorem 8.32 (Fundamental Theorem of Algebra)**.** *Every polynomial $P : \mathbb{C} \longrightarrow \mathbb{C}$, $P(z) := \sum_{j=0}^{n} a_j z^j$, of degree $n \geq 1$ (i.e. $a_0, \ldots, a_n \in \mathbb{C}$ with $a_n \neq 0$) has at least one zero $z_0 \in \mathbb{C}$.*

Proof. Dividing the equation $P(z) = 0$ by $a_n \neq 0$, it suffices to consider the case $a_n = 1$. We therefore assume

$$\underset{z \in \mathbb{C}}{\forall} \quad P(z) = z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0.$$

Claim 1. The function $|P|$ attains its global min on $\mathbb{C}$, i.e. there exists $z_0 \in \mathbb{C}$ such that $|P|$ is minimal in z_0.

Proof. We first note

$$\underset{z \neq 0}{\forall} \quad P(z) = z^n \big(1 + r(z)\big), \quad \text{where} \quad r(z) := \frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n}.$$

Set $M := |a_0| + \cdots + |a_{n-1}|$ and $R := \max\{1, 2M\}$.

Then

$$\underset{|z| \geq R}{\forall} \quad |r(z)| \overset{|z| \geq 1}{\leq} \frac{M}{|z|} \overset{|z| \geq 2M}{\leq} \frac{1}{2}$$

and, thus,

$$\underset{|z| \geq R}{\forall} \quad |P(z)| = |z|^n \big|1 + r(z)\big| \geq \frac{|z|^n}{2} \geq M.$$

This estimate together with $|P(0)| = |a_0| \leq M$ shows that the min of $|P|$ on the compact disk $\overline{B}_R(0)$ (see Ex. 7.47(a)) (such a min $z_0 \in \overline{B}_R(0)$ exists due to Th. 7.54) must be the global min of $|P|$ on $\mathbb{C}$. ▲

Claim 2. If $|P|$ has a min in $z_0 \in \mathbb{C}$, then $P(z_0) = 0$.

Proof. Proceeding by contraposition, we assume $P(z_0) \neq 0$ and show that $|P|$ does not have a min in z_0. We need to construct $z_1 \in \mathbb{C}$ such that $|P(z_1)| < |P(z_0)|$. To this end, define

$$p : \mathbb{C} \longrightarrow \mathbb{C}, \quad p(z) := \frac{P(z_0 + z)}{P(z_0)}.$$

Then p is still a polynomial of degree n. Since $p(0) = 1$,

$$\underset{k \in \{1,\ldots,n\}}{\exists} \quad \underset{b_k,\ldots,b_n \in \mathbb{C}}{\exists} \quad \underset{z \in \mathbb{C}}{\forall} \quad p(z) = 1 + \sum_{j=k}^{n} b_j z^j, \quad b_k \neq 0.$$

Write $-b_k^{-1}$ in polar form, i.e. $-b_k^{-1} = r e^{i\varphi}$ with $r \in \mathbb{R}^+$ and $\varphi \in \mathbb{R}$. Define

$$\beta := \sqrt[k]{r}\, e^{i\varphi/k} \quad \left(\text{i.e. } \beta^k = r e^{i\varphi} = -b_k^{-1}\right)$$

and

$$q : \mathbb{C} \longrightarrow \mathbb{C}, \quad q(z) := p(\beta z) = 1 + b_k \beta^k z^k + \sum_{j=k+1}^{n} b_j \beta^j z^j = 1 - z^k + z^{k+1} S(z),$$

where S is the polynomial

$$S : \mathbb{C} \longrightarrow \mathbb{C}, \quad S(z) := \sum_{j=0}^{n-k-1} b_{k+1+j} \beta^{k+1+j} z^j \quad (S \equiv 0 \text{ in case } k = n).$$

Then, according to Th. 7.54,

$$\underset{C \in \mathbb{R}^+}{\exists} \quad \underset{z \in \overline{B}_1(0)}{\forall} \quad |S(z)| \leq C.$$

Letting

$$c := \min\{1, C^{-1}\},$$

one obtains

$$\underset{0 < |z| < c}{\forall} \quad \left|z^{k+1} S(z)\right| \leq C\, |z|^{k+1} < |z|^k$$

and, thus,

$$\underset{x \in]0,c[}{\forall} \quad |q(x)| \leq 1 - x^k + \left|x^{k+1} S(x)\right| < 1 - x^k + x^k = 1.$$

Thus, finally,

$$\underset{x \in]0,c[}{\forall} \quad \frac{|P(z_0 + \beta x)|}{|P(z_0)|} = |p(\beta x)| = |q(x)| < 1,$$

showing $|P|$ does not have a min in z_0.   ▲

Combining Claims 1 and 2 completes the proof of the theorem.   ∎

Corollary 8.33. (a) *For every polynomial $P : \mathbb{C} \longrightarrow \mathbb{C}$ of degree $n \geq 1$, there exist numbers $c, \zeta_1, \ldots, \zeta_n \in \mathbb{C}$ such that*

$$P(z) = c \prod_{j=1}^{n} (z - \zeta_j) = c(z - \zeta_1)(z - \zeta_2) \cdots (z - \zeta_n) \tag{8.57}$$

(the $\zeta_1, \ldots, \zeta_n$ are precisely all the zeros of P, some or all of which might be identical).

(b) *For every polynomial $P : \mathbb{R} \longrightarrow \mathbb{R}$ of degree $n \geq 1$, there exist numbers $n_1, n_2 \in \mathbb{N}_0$ and $c, \xi_1, \ldots, \xi_{n_1}, \alpha_1, \ldots, \alpha_{n_2}, \beta_1, \ldots, \beta_{n_2} \in \mathbb{R}$ such that*

$$n = n_1 + 2n_2, \tag{8.58a}$$

and

$$P(x) = c \prod_{j=1}^{n_1} (x - \xi_j) \prod_{j=1}^{n_2} (x^2 + \alpha_j x + \beta_j). \tag{8.58b}$$

Proof. For (a), one just combines Th. 8.32 with Rem. 6.7.

(b): If P has only real coefficients, then we can take complex conjugates to obtain

$$P(\zeta) = 0 \;\Rightarrow\; \overline{P(\zeta)} = P(\overline{\zeta}) = 0, \tag{8.59}$$

showing that the nonreal zeros of P (if any) must occur in conjugate pairs. Moreover,

$$(x - \zeta)(x - \overline{\zeta}) = x^2 - (\zeta + \overline{\zeta})x + \zeta\overline{\zeta} = x^2 - 2x \operatorname{Re} \zeta + |\zeta|^2, \tag{8.60}$$

showing that (8.57) implies (8.58). ∎

9 Differential Calculus

9.1 Definition of Differentiability and Rules

The basic idea of differential calculus is to locally approximate nonlinear functions f by linear functions. In our case, f will be defined on a subset M of $\mathbb{R}$ and, given $\xi \in M$ and $\mathbb{R}$-valued f, we will investigate the question if we can define a number $f'(\xi) \in \mathbb{R}$ that represents the slope of the graph of f at ξ such that the line through ξ with slope $f'(\xi)$ (called the tangent of f in ξ) can be considered as a local approximation of the graph of f.

If such a local approximation of f in ξ is at all reasonable, then, for $x \neq \xi$,

$$\frac{f(x) - f(\xi)}{x - \xi}$$

should provide "good" approximations of $f'(\xi)$ if x tends to ξ. This leads to the following Def. 9.1, where we also allow $\mathbb{C}$-valued functions (while the above-described geometric interpretation only works for $\mathbb{R}$-valued functions, it can be applied to both the real and the imaginary parts of a $\mathbb{C}$-valued function, cf. Rem. 9.2 below); but note that we do not consider differentiability of functions $f : \mathbb{C} \longrightarrow \mathbb{C}$, which would lead to the notion of *complex* differentiability or *holomorphicity*, which is studied in the field of *Complex Analysis* and is beyond the scope of this class.

Definition 9.1. Let $a < b$, $f :]a, b[\longrightarrow \mathbb{K}$ ($a = -\infty$, $b = \infty$ is admissible), and $\xi \in]a, b[$. Then f is said to be *differentiable* at ξ if, and only if, the following limit in (9.1) exists in the sense of Def. 8.17 (where $x \mapsto \frac{f(x)-f(\xi)}{x-\xi}$ plays the role of $x \mapsto f(x)$ in Def. 8.17). The limit is then called the *derivative* of f in ξ. Many symbols are used in the literature to denote derivatives, the following provides a selection:

$$f'(\xi) := \partial_x f(\xi) := \frac{df(\xi)}{dx} := \lim_{x \to \xi} \frac{f(x) - f(\xi)}{x - \xi} = \lim_{h \to 0} \frac{f(\xi + h) - f(\xi)}{h}. \tag{9.1}$$

Note both limits occurring in (9.1) are, indeed, identical, since the sequence $(x_k)_{k \in \mathbb{N}}$ in $]a, b[$ converges to ξ if, and only if, the sequence $(h_k)_{k \to \infty}$ with $h_k := x_k - \xi$ converges to 0. The number in (9.1) (if it exists) is also called a *differential quotient*, whereas $\frac{f(x)-f(\xi)}{x-\xi}$ is known as a *difference quotient*.

f is called *differentiable* if, and only if, it is differentiable at each $\xi \in]a, b[$. In that case, one calls the function

$$f' :]a, b[\longrightarrow \mathbb{K}, \quad x \mapsto f'(x), \tag{9.2}$$

the *derivative* of f.

Remark 9.2. In the situation of Def. 9.1, the complex-valued function $f :]a, b[\longrightarrow \mathbb{C}$ is differentiable at $\xi \in]a, b[$ if, and only if, both functions $\operatorname{Re} f, \operatorname{Im} f :]a, b[\longrightarrow \mathbb{R}$ are differentiable, and, in that case

$$f'(\xi) = (\operatorname{Re} f)'(\xi) + i (\operatorname{Im} f)'(\xi). \tag{9.3}$$

Indeed, we merely have to note

$$\forall_{\substack{x,\xi \in]a,b[, \\ x \neq \xi}} \frac{f(x) - f(\xi)}{x - \xi} = \frac{\operatorname{Re} f(x) - \operatorname{Re} f(\xi)}{x - \xi} + i \frac{\operatorname{Im} f(x) - \operatorname{Im} f(\xi)}{x - \xi} \tag{9.4}$$

and that, by (7.2) a sequence $(z_n)_{n \in \mathbb{N}}$ in $\mathbb{C}$ converges to $\zeta \in \mathbb{C}$ if, and only if, both $\lim_{n \to \infty} \operatorname{Re} z_n = \operatorname{Re} \zeta$ and $\lim_{n \to \infty} \operatorname{Im} z_n = \operatorname{Im} \zeta$ hold.

Definition 9.3. If $f :]a, b[\longrightarrow \mathbb{R}$ as in Def. 9.1 is differentiable at $\xi \in]a, b[$, then the graph of the affine function

$$L : \mathbb{R} \longrightarrow \mathbb{R}, \quad L(x) := f(\xi) + f'(\xi)(x - \xi), \tag{9.5}$$

i.e. the line through $(\xi, f(\xi))$ with slope $f'(\xi)$ is called the *tangent* to the graph of f at ξ.

Theorem 9.4. *If $f :]a, b[\longrightarrow \mathbb{K}$ as in Def. 9.1 is differentiable at $\xi \in]a, b[$, then it is continuous at ξ. In particular, if f is everywhere differentiable, then it is everywhere continuous.*

Proof. Let $(x_k)_{k \in \mathbb{N}}$ be a sequence in $]a, b[\setminus \{\xi\}$ such that $\lim_{k \to \infty} x_k = \xi$. Then

$$\lim_{k \to \infty} \big(f(x_k) - f(\xi)\big) = \lim_{k \to \infty} \frac{(x_k - \xi)\big(f(x_k) - f(\xi)\big)}{x_k - \xi} = 0 \cdot f'(\xi) = 0, \qquad (9.6)$$

proving the continuity of f in ξ. ∎

Example 9.5. (a) For each $a, b \in \mathbb{K}$, the affine function $f : \mathbb{R} \longrightarrow \mathbb{K}$, $f(x) := ax + b$, is differentiable with $f'(x) = a$ for each $x \in \mathbb{R}$: If $x \in \mathbb{R}$ and $(h_k)_{k \in \mathbb{N}}$ is a sequence with $h_k \neq 0$ such that $\lim_{k \to \infty} h_k = 0$, then

$$\lim_{k \to \infty} \frac{f(x + h_k) - f(x)}{h_k} = \lim_{k \to \infty} \frac{a(x + h_k) + b - ax - b}{h_k} = \lim_{k \to \infty} \frac{a\, h_k}{h_k} = a. \qquad (9.7)$$

In particular, each constant function $f \equiv b$ has derivative $f' \equiv 0$.

(b) For each $c \in \mathbb{K}$, the function $f : \mathbb{R} \longrightarrow \mathbb{K}$, $f(x) := e^{cx}$, is differentiable with $f'(x) = c\, e^{cx}$ for each $x \in \mathbb{R}$ (in particular, $c = 1$ yields $f'(x) = e^x$ for $f(x) = e^x$, and $c = \ln a$ yields $f'(x) = (\ln a)\, a^x$ for $f(x) = a^x = e^{x \ln a}$, $a \in \mathbb{R}^+$): The case $c = 0$ was treated in (a). Thus, let $c \neq 0$. If $x \in \mathbb{R}$ and $(h_k)_{k \in \mathbb{N}}$ is a sequence with $h_k \neq 0$ such that $\lim_{k \to \infty} h_k = 0$, then

$$\lim_{k \to \infty} \frac{f(x + h_k) - f(x)}{h_k} = \lim_{k \to \infty} \frac{e^{cx + ch_k} - e^{cx}}{h_k} = c\, e^{cx} \lim_{k \to \infty} \frac{e^{ch_k} - 1}{ch_k} \overset{(8.32)}{=} c\, e^{cx}. \qquad (9.8)$$

(c) The sine and the cosine function $f, g : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) := \sin x$, $g(x) := \cos x$, are differentiable with $f'(x) = \cos x$ and $g'(x) = -\sin x$ for each $x \in \mathbb{R}$: If $x \in \mathbb{R}$ and $(h_k)_{k \in \mathbb{N}}$ is a sequence with $h_k \neq 0$ such that $\lim_{k \to \infty} h_k = 0$, then

$$
\begin{aligned}
\lim_{k \to \infty} \frac{f(x + h_k) - f(x)}{h_k} &= \lim_{k \to \infty} \frac{\sin(x + h_k) - \sin x}{h_k} \\
&\overset{(8.44c)}{=} \lim_{k \to \infty} \frac{\sin x \cos h_k + \cos x \sin h_k - \sin x}{h_k} \\
&= \sin x \lim_{k \to \infty} \frac{h_k(\cos h_k - 1)}{h_k^2} + \cos x \lim_{k \to \infty} \frac{\sin h_k}{h_k} \\
&\overset{(8.44j)}{=} (\sin x) \cdot 0 \cdot \left(-\frac{1}{2}\right) + (\cos x) \cdot 1 = \cos x. \qquad (9.9)
\end{aligned}
$$

The proof of $g'(x) = -\sin x$ is left as an exercise.

(d) The absolute value function $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) := |x|$, is *not* differentiable at $\xi = 0$:

$$\lim_{n \to \infty} \frac{f(0 + \frac{1}{n}) - f(0)}{\frac{1}{n}} = \lim_{n \to \infty} 1 = 1, \qquad (9.10a)$$

$$\lim_{n \to \infty} \frac{f(0 - \frac{1}{n}) - f(0)}{-\frac{1}{n}} = \lim_{n \to \infty} \frac{\frac{1}{n}}{-\frac{1}{n}} = -1, \qquad (9.10b)$$

showing that $\frac{f(0+h)-f(0)}{h}$ does *not* have a limit for $h \to 0$.

Remark 9.6. As just seen in Ex. 9.5(d), the absolute value function shows continuous functions do not need to be differentiable. Somewhat surprisingly, with a bit more effort, one can even construct continuous functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ that are not differentiable at *any* $x \in \mathbb{R}$ (see Appendix J.1).

Theorem 9.7. *Let* $a < b$, $f, g :]a, b[\longrightarrow \mathbb{K}$ *(*$a = -\infty$, $b = \infty$ *is admissible), and* $\xi \in]a, b[$. *Assume* f *and* g *are differentiable at* ξ.

(a) *For each* $\lambda \in \mathbb{K}$, λf *is differentiable at* ξ *and* $(\lambda f)'(\xi) = \lambda f'(\xi)$.

(b) $f + g$ *is differentiable at* ξ *and* $(f + g)'(\xi) = f'(\xi) + g'(\xi)$.

(c) Product Rule*:* fg *is differentiable at* ξ *and* $(fg)'(\xi) = f'(\xi)g(\xi) + f(\xi)g'(\xi)$.

(d) Quotient Rule*: If* $g(\xi) \neq 0$, *then* f/g *is differentiable at* ξ *and*

$$(f/g)'(\xi) = \frac{f'(\xi)g(\xi) - f(\xi)g'(\xi)}{(g(\xi))^2}, \quad \text{in particular} \quad (1/g)'(\xi) = -\frac{g'(\xi)}{(g(\xi))^2}.$$

Proof. Let $(h_k)_{k \in \mathbb{N}}$ be a sequence with $h_k \neq 0$ such that $\lim_{k \to \infty} h_k = 0$.

For (a), one computes

$$\lim_{k \to \infty} \frac{(\lambda f)(\xi + h_k) - (\lambda f)(\xi)}{h_k} = \lim_{k \to \infty} \frac{\lambda f(\xi + h_k) - \lambda f(\xi)}{h_k}$$

$$= \lambda \lim_{k \to \infty} \frac{f(\xi + h_k) - f(\xi)}{h_k} = \lambda f'(\xi).$$

For (b), one computes

$$\lim_{k \to \infty} \frac{(f + g)(\xi + h_k) - (f + g)(\xi)}{h_k} = \lim_{k \to \infty} \frac{f(\xi + h_k) - f(\xi) + g(\xi + h_k) - g(\xi)}{h_k}$$

$$= \lim_{k \to \infty} \frac{f(\xi + h_k) - f(\xi)}{h_k} + \lim_{k \to \infty} \frac{g(\xi + h_k) - g(\xi)}{h_k} = f'(\xi) + g'(\xi).$$

For (c), one computes

$$\lim_{k \to \infty} \frac{(fg)(\xi + h_k) - (fg)(\xi)}{h_k}$$

$$= \lim_{k \to \infty} \frac{f(\xi + h_k)g(\xi + h_k) - f(\xi)g(\xi + h_k) + f(\xi)g(\xi + h_k) - f(\xi)g(\xi)}{h_k}$$

$$= \lim_{k \to \infty} g(\xi + h_k) \lim_{k \to \infty} \frac{f(\xi + h_k) - f(\xi)}{h_k} + f(\xi) \lim_{k \to \infty} \frac{g(\xi + h_k) - g(\xi)}{h_k}$$

$$= f'(\xi)g(\xi) + f(\xi)g'(\xi),$$

where, in the last equality, we used the continuity of g in ξ according to Th. 9.4.

For (d), one first proves the special case $f \equiv 1$ by

$$\lim_{k\to\infty} \frac{(1/g)(\xi + h_k) - (1/g)(\xi)}{h_k} = \lim_{k\to\infty} \frac{g(\xi) - g(\xi + h_k)}{g(\xi + h_k)g(\xi)h_k} = -\frac{g'(\xi)}{(g(\xi))^2},$$

which implies the general case using (c):

$$(f/g)'(\xi) = \left(f \cdot \frac{1}{g}\right)'(\xi) = \frac{f'(\xi)}{g(\xi)} - \frac{f(\xi)g'(\xi)}{(g(\xi))^2} = \frac{f'(\xi)g(\xi) - f(\xi)g'(\xi)}{(g(\xi))^2},$$

completing the proof. ∎

Example 9.8. (a) Each polynomial is differentiable and the derivative is, again, a polynomial. More precisely,

$$P : \mathbb{R} \longrightarrow \mathbb{K}, \quad P(x) = \sum_{j=0}^{n} a_j x^j, \quad a_j \in \mathbb{K}$$

$$\Rightarrow \quad P' : \mathbb{R} \longrightarrow \mathbb{K}, \quad P'(x) = \sum_{j=1}^{n} j\, a_j x^{j-1} :$$

$$(9.11)$$

The cases $n = 0, 1$ are provided by Ex. 9.5(a). To complete the induction proof of (9.11), we carry out the induction step for each $n \in \mathbb{N}$: Writing $P(x) = \sum_{j=0}^{n} a_j x^j + a_{n+1} x\, x^n$ and applying the induction hypothesis as well as the rules of Th. 9.7 yields

$$P'(x) = \sum_{j=1}^{n} j\, a_j x^{j-1} + a_{n+1}(1 \cdot x^n + x \cdot n \cdot x^{n-1}) = \sum_{j=1}^{n+1} j\, a_j x^{j-1},$$

which establishes the case.

(b) Clearly, the derivatives of rational functions P/Q with polynomials P and Q can be computed from (9.11) and the quotient rule of Th. 9.7(d).

(c) The functions tan and cot as defined in (8.48) and restricted to $\mathbb{R} \setminus \cos^{-1}\{0\}$ and $\mathbb{R} \setminus \sin^{-1}\{0\}$, respectively, are differentiable and one obtains

$$\tan' : \underbrace{\mathbb{R} \setminus \cos^{-1}\{0\}}_{\mathbb{R}\setminus\{(2k+1)\frac{\pi}{2}\,:\,k\in\mathbb{Z}\}} \longrightarrow \mathbb{R}, \quad \tan' x = \frac{1}{(\cos x)^2} = 1 + (\tan x)^2, \tag{9.12a}$$

$$\cot' : \underbrace{\mathbb{R} \setminus \sin^{-1}\{0\}}_{\mathbb{R}\setminus\{k\pi\,:\,k\in\mathbb{Z}\}} \longrightarrow \mathbb{R}, \quad \cot' x = -\frac{1}{(\sin x)^2} = -(1 + (\cot x)^2) : \tag{9.12b}$$

One merely needs the derivatives of sin and cos from Ex. 9.5(c) and the quotient rule of Th. 9.7(d):

$$\tan' x = \frac{\cos x \cos x - \sin x(-\sin x)}{(\cos x)^2} \overset{(8.44e)}{=} \frac{1}{(\cos x)^2} \overset{(8.44e)}{=} 1 + (\tan x)^2,$$

$$\cot' x = \frac{-\sin x \sin x - \cos x \cos x}{(\sin x)^2} \overset{(8.44e)}{=} -\frac{1}{(\sin x)^2} \overset{(8.44e)}{=} -(1 + (\cot x)^2).$$

Theorem 9.9 (Derivative of Inverse Functions). *Let $a < b$, $I :=]a, b[$ ($a = -\infty$, $b = \infty$ is admissible). If $f : I \longrightarrow \mathbb{R}$ is differentiable and strictly increasing (resp. decreasing), then f has a continuous, strictly increasing (resp. decreasing) inverse function f^{-1} defined on the interval $J := f(I)$, i.e. $f^{-1} : J \longrightarrow I$, and, for each $\xi \in I$ with $f'(\xi) \neq 0$, f^{-1} is differentiable at $\eta := f(\xi)$ with*

$$(f^{-1})'(\eta) = \frac{1}{f'(\xi)} = \frac{1}{f'(f^{-1}(\eta))}. \tag{9.13}$$

Proof. As a differentiable function, f is continuous by Th. 9.4, i.e. Th. 7.60 provides all the present assertions, except differentiability at η and (9.13). Let $(y_k)_{k \in \mathbb{N}}$ be a sequence in $J \setminus \{\eta\}$ such that $\lim_{k \to \infty} y_k = \eta$. Then, as f^{-1} is bijective and continuous, $(f^{-1}(y_k))_{k \in \mathbb{N}}$ is a sequence in $I \setminus \{\xi\}$ such that $\lim_{k \to \infty} f^{-1}(y_k) = \xi$, and one obtains

$$\lim_{k \to \infty} \frac{f^{-1}(y_k) - f^{-1}(\eta)}{y_k - \eta} = \lim_{k \to \infty} \frac{f^{-1}(y_k) - f^{-1}(\eta)}{f(f^{-1}(y_k)) - f(f^{-1}(\eta))} = \frac{1}{f'(f^{-1}(\eta))}, \tag{9.14}$$

establishing the case. ■

Example 9.10. (a) The function $\ln : \mathbb{R}^+ \longrightarrow \mathbb{R}$ is differentiable and, for each $x \in \mathbb{R}^+$, $\ln' x = 1/x$: If $f(x) = e^x$, then $f'(x) = e^x \neq 0$ for each $x \in \mathbb{R}$, $\ln x = f^{-1}(x)$, and (9.13) yields

$$\ln' x = \frac{1}{f'(\ln x)} = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$

(b) The function $\arcsin\, :] -1, 1[\longrightarrow] -\pi/2, \pi/2[$ is differentiable and, for each $x \in\,] -1, 1[$, $\arcsin' x = 1/\sqrt{1 - x^2}$: If $f(x) = \sin x$, then $f'(x) = \cos x \neq 0$ for each $x \in\,] -\pi/2, \pi/2[$, $\arcsin x = f^{-1}(x)$, and (9.13) yields

$$\arcsin' x = \frac{1}{f'(\arcsin x)} = \frac{1}{\cos \arcsin x} \overset{(*)}{=} \frac{1}{\sqrt{1 - (\sin \arcsin x)^2}} = \frac{1}{\sqrt{1 - x^2}},$$

where, at $(*)$, it was used that $\cos^2 = 1 - \sin^2$ and $\cos t > 0$ for each $t \in\,] -\pi/2, \pi/2[$.

(c) The function $\arccos\, :] -1, 1[\longrightarrow]0, \pi[$ is differentiable and, for each $x \in\,] -1, 1[$, $\arccos' x = -1/\sqrt{1 - x^2}$: If $f(x) = \cos x$, then $f'(x) = -\sin x \neq 0$ for each $x \in\,]0, \pi[$, $\arccos x = f^{-1}(x)$, and (9.13) yields

$$\arccos' x = \frac{1}{f'(\arccos x)} = \frac{1}{-\sin \arccos x} \overset{(*)}{=} -\frac{1}{\sqrt{1 - (\cos \arccos x)^2}} = -\frac{1}{\sqrt{1 - x^2}},$$

where, at $(*)$, it was used that $\sin^2 = 1 - \cos^2$ and $\sin t > 0$ for each $t \in\,]0, \pi[$.

(d) The function $\arctan : \mathbb{R} \longrightarrow] -\pi/2, \pi/2[$ is differentiable and, for each $x \in \mathbb{R}$, $\arctan' x = 1/(1 + x^2)$: Apply Th. 9.9 with $f(x) = \tan x$ as an exercise.

(e) The function $\text{arccot} : \mathbb{R} \longrightarrow]0, \pi[$ is differentiable and, for each $x \in \mathbb{R}$, $\text{arccot}' x = -1/(1 + x^2)$: Apply Th. 9.9 with $f(x) = \cot x$ as an exercise.

Theorem 9.11 (Chain Rule). *Let $a < b$, $c < d$, $f :]a, b[\longrightarrow \mathbb{R}$, $g :]c, d[\longrightarrow \mathbb{K}$, $f(]a, b[) \subseteq]c, d[$ ($a, c = -\infty$; $b, d = \infty$ is admissible). If f is differentiable in $\xi \in]a, b[$ and g is differentiable in $f(\xi) \in]c, d[$, then $g \circ f :]a, b[\longrightarrow \mathbb{K}$ is differentiable in ξ and*

$$(g \circ f)'(\xi) = f'(\xi)g'(f(\xi)). \tag{9.15}$$

Proof. Let $\eta := f(\xi)$ and define the auxiliary function

$$\tilde{g} :]c, d[\longrightarrow \mathbb{K}, \quad \tilde{g}(x) := \begin{cases} \frac{g(x)-g(\eta)}{x-\eta} & \text{for } x \neq \eta, \\ g'(x) & \text{for } x = \eta. \end{cases} \tag{9.16}$$

Then

$$\underset{x \in]c,d[}{\forall} \quad g(x) - g(\eta) = \tilde{g}(x)(x - \eta). \tag{9.17}$$

Let $(x_k)_{k \in \mathbb{N}}$ be a sequence in $]a, b[\backslash\{\xi\}$ such that $\lim_{k \to \infty} x_k = \xi$. One obtains

$$\begin{aligned}
\lim_{k \to \infty} \frac{g(f(x_k)) - g(f(\xi))}{x_k - \xi} & \overset{(9.17)}{=} \lim_{k \to \infty} \frac{\tilde{g}(f(x_k))(f(x_k) - f(\xi))}{x_k - \xi} \\
& = \lim_{k \to \infty} \tilde{g}(f(x_k)) \lim_{k \to \infty} \frac{f(x_k) - f(\xi)}{x_k - \xi} \\
& = f'(\xi)g'(f(\xi)), \tag{9.18}
\end{aligned}$$

establishing the case. ∎

Example 9.12. **(a)** According to the chain rule of Th. 9.11, the function $h : \mathbb{R} \longrightarrow \mathbb{R}$, $h(x) := \sin(-x^3)$ is differentiable and, for each $x \in \mathbb{R}$, $h'(x) = -3x^2 \cos(-x^3)$.

(b) According to the chain rule of Th. 9.11, each power function $h : \mathbb{R}^+ \longrightarrow \mathbb{K}$, $h(x) := x^\alpha = e^{\alpha \ln x}$, $\alpha \in \mathbb{K}$, is differentiable and, for each $x \in \mathbb{R}^+$, $h'(x) = \frac{\alpha}{x} e^{\alpha \ln x} = \alpha \, x^{\alpha-1}$. Indeed, $h = g \circ f$, where $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $f(x) := \ln x$ with $f' : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $f'(x) := \frac{1}{x}$, according to Ex. 9.5(b), and $g : \mathbb{R} \longrightarrow \mathbb{K}$, $g(x) := e^{\alpha x}$, with $g' : \mathbb{R} \longrightarrow \mathbb{K}$, $g'(x) := \alpha \, e^{\alpha x}$ according to Ex. 9.10(a).

9.2 Higher Order Derivatives and the Sets C^k

Definition 9.13. Let $a < b$, $I :=]a, b[$, $f : I \longrightarrow \mathbb{K}$ ($a = -\infty$, $b = \infty$ is admissible). If f is differentiable, then f' might or might not itself be differentiable. If f' is differentiable, then its derivative is denoted by f'' and is called the second derivative of f. Clearly, this process can be iterated, leading to the following general recursive definition of higher-order derivatives:

Let $f^{(0)} := f$. For $k \in \mathbb{N}_0$ assume the kth derivative of f, denoted by $f^{(k)}$ exists on I. Then f is said to have a derivative of order $k + 1$ at $\xi \in I$ if, and only if, $f^{(k)}$ is differentiable at ξ. In that case, define

$$f^{(k+1)}(\xi) := (f^{(k)})'(\xi). \tag{9.19}$$

If $f^{(k+1)}(\xi)$ exists for all $\xi \in I$, then f is said to be $(k+1)$-times differentiable and the function $f^{(k+1)} : I \longrightarrow \mathbb{K}$, $x \mapsto f^{(k+1)}(\xi)$, is called the $(k+1)$st derivative of f. It is common to write $f' := f^{(1)}$, $f'' := f^{(2)}$, $f''' := f^{(3)}$, but $f^{(k)}$ if $k \geq 4$.

If $f^{(k)}$ exists, it might or might not be continuous (cf. Ex. 9.14(c) below). One defines

$$\underset{k \in \mathbb{N}_0}{\forall} \quad C^k(I, \mathbb{K}) := \left\{ f \in \mathcal{F}(I, \mathbb{K}) : f^{(k)} \text{ exists and is continuous on } I \right\}, \qquad (9.20)$$

$$C^\infty(I, \mathbb{K}) := \bigcap_{k \in \mathbb{N}_0} C^k(I, \mathbb{K}) \qquad (9.21)$$

(note $C^0(I, \mathbb{K}) = C(I, \mathbb{K})$ and $C(I, \mathbb{K}) \supseteq C^1(I, \mathbb{K}) \supseteq C^2(I, \mathbb{K}) \supseteq \dots$). Finally, we define the notation $C^k(I) := C^k(I, \mathbb{R})$ for $k \in \mathbb{N}_0 \cup \{\infty\}$.

Example 9.14. (a) One has $\sin \in C^\infty(\mathbb{R})$ with $\sin' = \cos$, $\sin'' = -\sin$, $\sin''' = -\cos$, $\sin^{(4)} = \sin$, $\dots$

(b) A simple induction shows, for each polynomial $P : \mathbb{R} \longrightarrow \mathbb{K}$, $P(x) = \sum_{j=0}^{n} a_j x^j$, $a_j \in \mathbb{K}$, $n \in \mathbb{N}_0$, that $P^{(n)}(x) = n!\, a_n$. In particular, $P \in C^\infty(\mathbb{R}, \mathbb{K})$.

(c) It is an exercise to show the following function f is differentiable, but f' is not continuous, i.e. $f \notin C^1(\mathbb{R})$:

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x) := \begin{cases} x^2 \cos\left(\frac{1}{x}\right) & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

9.3   Mean Value Theorem, Monotonicity, and Extrema

Theorem 9.15. *Let $a < b$. If $f :]a, b[\longrightarrow \mathbb{R}$ is differentiable in $\xi \in]a, b[$ and f has a local min or max in ξ, then $f'(\xi) = 0$.*

Proof. Suppose f has a local max at ξ. Then there exists $\epsilon > 0$ such that $|h| < \epsilon$ implies $f(\xi + h) - f(\xi) \leq 0$. Now let $(h_k)_{k \in \mathbb{N}}$ be a sequence in $]0, \epsilon[$ with $\lim_{k \to \infty} h_k = 0$. Then $f(\xi \pm h_k) - f(\xi) \leq 0$ for all $k \in \mathbb{N}$ implies

$$f'(\xi) = \lim_{k \to \infty} \frac{f(\xi + h_k) - f(\xi)}{h_k} \leq 0, \quad f'(\xi) = \lim_{k \to \infty} \frac{f(\xi - h_k) - f(\xi)}{-h_k} \geq 0, \qquad (9.22)$$

showing $f'(\xi) = 0$. Now, if f has a local min at ξ, then $-f$ has a local max at ξ, and $f'(\xi) = -(-f)'(\xi) = 0$ establishes the case. ∎

Remark 9.16. For $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) := x^3$, it is $f'(0) = 0$, but f does not have a local min or max at 0, showing that, while being *necessary* for an differentiable function f to have a local extremum at ξ, $f'(\xi) = 0$ is not a *sufficient* condition for such an extremum at ξ. Points ξ with $f'(\xi) = 0$ are sometimes called *stationary* or *critical* points of f.

———

Now, we first prove an important special case of the mean value theorem:

Theorem 9.17 (Rolle's Theorem). *Let $a < b$. If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous on the compact interval $[a, b]$, differentiable on the open interval $]a, b[$, and $f(a) = f(b)$, then there exists $\xi \in]a, b[$ such that $f'(\xi) = 0$.*

Proof. If f is constant, then $f'(\xi) = 0$ holds for each $\xi \in]a, b[$. If f is nonconstant, then there exists $x \in]a, b[$ with $f(x) \neq f(a)$. If $f(x) > f(a)$, then Th. 7.54 implies the existence of $\xi \in]a, b[$ such that f attains its (global and, thus, local) max in ξ. Then Th. 9.15 yields $f'(\xi) = 0$. The case $f(x) < f(a)$ is treated analogously. ∎

Theorem 9.18 (Mean Value Theorem). *Let $a < b$. If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous on the compact interval $[a, b]$ and differentiable on the open interval $]a, b[$, then there exists $\xi \in]a, b[$ such that*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi). \tag{9.23}$$

Proof. One applies Rolle's Th. 9.17 to the auxiliary function

$$\phi : [a, b] \longrightarrow \mathbb{R}, \quad \phi(x) := f(x) - \alpha\,(x - a), \quad \text{where} \quad \alpha := \frac{f(b) - f(a)}{b - a}. \tag{9.24}$$

Since f is continuous on $[a, b]$ and differentiable on $]a, b[$, so is ϕ. Moreover $\phi(a) = f(a) = \phi(b)$, i.e. Rolle's Th. 9.17 applies and yields $\xi \in]a, b[$ satisfying $0 = \phi'(\xi) = f'(\xi) - \alpha$, proving (9.23). ∎

Corollary 9.19. *Let $c < d$ and $f :]c, d[\longrightarrow \mathbb{R}$ be differentiable ($c = -\infty$, $d = \infty$ is admissible).*

(a) *If $f' \geq 0$ (resp. $f' \leq 0$), then f is increasing (resp. decreasing). Moreover, if the inequalities are strict, then the monotonicity of f is strict as well.*

(b) *If $f' \equiv 0$, then f is constant.*

Proof. If $c < a < b < d$ and $f' \geq 0$ (resp. $f' \leq 0$, resp. $f' \equiv 0$), then (9.23) implies $f(b) \geq f(a)$ (resp. $f(b) \leq f(a)$, resp. $f(b) = f(a)$). Moreover, strict inequalities for f' yield strict inequality between $f(b)$ and $f(a)$. ∎

Lemma 9.20. *Let $a < b$, $f :]a, b[\longrightarrow \mathbb{R}$, $\xi \in]a, b[$, and assume f is differentiable at ξ. If $f'(\xi) > 0$ (resp. $f'(\xi) < 0$), then there exists $\epsilon > 0$ such that $]\xi - \epsilon, \xi + \epsilon[\subseteq]a, b[$ and*

$$\underset{a_1 \in]\xi-\epsilon,\xi[}{\forall} \quad \underset{b_1 \in]\xi,\xi+\epsilon[}{\forall} \quad f(a_1) < f(\xi) < f(b_1) \quad \big(resp. \quad f(a_1) > f(\xi) > f(b_1)\big).$$

Proof. If there does not exist $\epsilon > 0$ such that $f(a_1) < f(\xi) < f(b_1)$ for each $a_1 \in]\xi - \epsilon, \xi[$ and each $b_1 \in]\xi, \xi + \epsilon[$, then then there exists a sequence $(x_k)_{k \in \mathbb{N}}$ in $]a, b[\setminus \{\xi\}$ such that $\lim_{k \to \infty} x_k = \xi$ and

$$\underset{k \in \mathbb{N}}{\forall} \quad \frac{f(x_k) - f(\xi)}{x_k - \xi} \leq 0,$$

showing $f'(\xi) \leq 0$. Analogously, one obtains that $f'(\xi) \geq 0$ provided there does not exist $\epsilon > 0$ such that $f(a_1) > f(\xi) > f(b_1)$ for each $a_1 \in]\xi - \epsilon, \xi[$ and each $b_1 \in]\xi, \xi + \epsilon[$. ∎

Caveat 9.21. The hypotheses of Lem. 9.20 are not sufficient for f to be increasing or decreasing in any neighborhood of ξ: It is an exercise to find a counterexample.

Theorem 9.22 (Sufficient Conditions for Extrema). *Let $c < d$, let $f :]c, d[\longrightarrow \mathbb{R}$ be differentiable, and assume $f'(\xi) = 0$ for some $\xi \in]c, d[$.*

(a) *If $f'(x) > 0$ for each $x \in]c, \xi[$ and $f'(x) < 0$ for each $x \in]\xi, d[$, then f has a strict max at ξ. Likewise, if $f''(\xi)$ exists and is negative, then f has a strict max at ξ.*

(b) *If $f'(x) < 0$ for each $x \in]c, \xi[$ and $f'(x) > 0$ for each $x \in]\xi, d[$, then f has a strict min at ξ. Likewise, if $f''(\xi)$ exists and is positive, then f has a strict min at ξ.*

Proof. We just present the proof for (a); (b) is proved analogously. If $f'(x) > 0$ for each $x \in]c, \xi[$, then (9.23) shows $f(\xi) - f(a) > 0$ for each $c < a < \xi$; analogously, if $f'(x) < 0$ for each $x \in]\xi, d[$, then (9.23) shows $f(\xi) - f(b) > 0$ for each $\xi < b < d$. Altogether, we have shown f to have a strict max at ξ. If $f''(\xi)$ exists and is negative, then Lem. 9.20 yields the existence of $\epsilon > 0$ such that f' is positive on $]\xi - \epsilon, \xi[$ and negative on $]\xi, \xi + \epsilon[$. Applying what we have already proved with $c := \xi - \epsilon$ and $d := \xi + \epsilon$ establishes the case. ∎

Example 9.23. One obtains

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \qquad\qquad f(x) := x\,e^x, \qquad\qquad (9.25a)$$

$$f' : \mathbb{R} \longrightarrow \mathbb{R}, \qquad\qquad f'(x) = e^x + x\,e^x = (1+x)\,e^x, \qquad (9.25b)$$

$$f'' : \mathbb{R} \longrightarrow \mathbb{R}, \qquad\qquad f''(x) = 2e^x + x\,e^x = (2+x)\,e^x. \qquad (9.25c)$$

From Th. 9.15, we know that f can have at most one extremum, namely at $\xi = -1$, where $f'(\xi) = 0$. Since $f''(\xi) = e^{-x} > 0$, Th. 9.22(b) implies that f has a strict min at -1.

—

The following Th. 9.24, the intermediate value theorem for derivatives, is another application of Lem. 9.20. Even though Ex. 9.14(c) shows that derivatives do not have to be continuous, not every function can occur as a derivative: The following Th. 9.24 shows that derivatives always satisfy an intermediate value property, even if they are not continuous (that also means that discontinuities in derivatives are always due to oscillations rather than jumps).

Theorem 9.24 (Intermediate Value Theorem for Derivatives). *Let $a, b, c, d \in \mathbb{R}$ with $a < c \leq d < b$. If $f :]a, b[\longrightarrow \mathbb{R}$ is differentiable, then f' assumes every value between $f'(c)$ and $f'(d)$, i.e.*

$$\Big[\min\{f'(c), f'(d)\}, \max\{f'(c), f'(d)\}\Big] \subseteq f'([c, d]). \qquad (9.26)$$

Proof. Exercise (hint: use a suitable auxiliary function and apply Lem. 9.20 together with Th. 7.54 and Th. 9.15). ∎

9.4 L'Hôpital's Rule

We need a slight generalization of the mean value Th. 9.18:

Theorem 9.25. *Let $a < b$. If $f, g : [a, b] \longrightarrow \mathbb{R}$ are continuous on the compact interval $[a, b]$, differentiable on the open interval $]a, b[$, and $g'(x) \neq 0$ for each $x \in]a, b[$, then there exists $\xi \in]a, b[$ such that*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}. \tag{9.27}$$

Proof. First note that the mean value Th. 9.18 and $g' \neq 0$ imply $g(b) - g(a) \neq 0$. Define the auxiliary function

$$h : [a, b] \longrightarrow \mathbb{R}, \quad h(x) := f(x) - \big(g(x) - g(a)\big) \frac{f(b) - f(a)}{g(b) - g(a)}. \tag{9.28}$$

Then h is continuous on $[a, b]$ and differentiable on $]a, b[$. Moreover, $h(a) = f(a) = h(b)$. Applying Th. 9.18 to h, yields the existence of some $\xi \in]a, b[$ satisfying $h'(\xi) = 0$. However, (9.28) implies $h'(\xi) = 0$ is equivalent to (9.27). ∎

L'Hôpital's rule is a result that can help to determine (function) limits (cf. Def. 8.17).

Theorem 9.26 (L'Hôpital's Rule). *Let $\xi \in \mathbb{R}$ and either $I =]a, \xi[$ with $a < \xi$ or $I :=]\xi, b[$ with $\xi < b$. Moreover, assume $f, g : I \longrightarrow \mathbb{R}$ are differentiable, $g'(x) \neq 0$ for each $x \in I$, and one of the following two conditions* (a), (b) *is satisfied:*

(a) $\lim_{x \to \xi} f(x) = \lim_{x \to \xi} g(x) = 0$.

(b) $\lim_{x \to \xi} g(x) = \infty$ *or* $\lim_{x \to \xi} g(x) = -\infty$, *where Def. 8.17 is extended to the case $\eta \in \{-\infty, \infty\}$ in the obvious way.*

Then

$$\lim_{x \to \xi} \frac{f'(x)}{g'(x)} = \eta \quad \Rightarrow \quad \lim_{x \to \xi} \frac{f(x)}{g(x)} = \eta. \tag{9.29}$$

The above statement also holds for $\xi \in \{-\infty, \infty\}$ and/or $\eta \in \{-\infty, \infty\}$ if, as in (b), *one extends Def. 8.17 to these cases in the obvious way.*

Proof. First, we assume (a). Consider the case $\xi \in \mathbb{R}$. Since f and g are continuous, (a) implies f and g remain continuous, if we extend them to ξ by letting $f(\xi) := g(\xi) = 0$. This extension will now allow us to apply Th. 9.25 to f and g. To prove (9.29), let $(x_k)_{k \in \mathbb{N}}$ be a sequence in I with $\lim_{k \to \infty} x_k = \xi$. Then (9.27) yields, for each $k \in \mathbb{N}$, some $\xi_k \in]x_k, \xi[$ if $x_k < \xi$ and some $\xi_k \in]\xi, x_k[$ if $\xi < x_k$, satisfying

$$\frac{f(x_k)}{g(x_k)} = \frac{f(x_k) - f(\xi)}{g(x_k) - g(\xi)} = \frac{f'(\xi_k)}{g'(\xi_k)}. \tag{9.30}$$

From the Sandwich Th. 7.16, we obtain $\lim_{k \to \infty} \xi_k = \xi$, i.e. (9.30) and $\lim_{x \to \xi} \frac{f'(x)}{g'(x)} = \eta$ imply $\lim_{x \to \xi} \frac{f(x)}{g(x)} = \eta$ (also for $\eta \in \{-\infty, \infty\}$). Now consider the case $\xi \in \{-\infty, \infty\}$ and let $(x_k)_{k \in \mathbb{N}}$ be as before. If $\xi = \infty$, then choose $1 \leq c \in I$ and set $\tilde{I} :=\,]0, c^{-1}[$; if $\xi = -\infty$, then choose $-1 \geq c \in I$ and set $\tilde{I} :=\,]c^{-1}, 0[$. We apply what we have already proved above to the auxiliary functions

$$\tilde{f} : \tilde{I} \longrightarrow \mathbb{R}, \quad \tilde{f}(x) := f(1/x), \quad \tilde{g} : \tilde{I} \longrightarrow \mathbb{R}, \quad \tilde{g}(x) := g(1/x)$$

at $\tilde{\xi} := 0$. From the chain rule (9.15), we know $\tilde{f}'(x) = -\frac{f'(1/x)}{x^2}$ and $\tilde{g}'(x) = -\frac{g'(1/x)}{x^2}$ for each $x \in \tilde{I}$. Thus, $\lim_{x \to \xi} \frac{f'(x)}{g'(x)} = \eta$ implies,

$$\eta = \lim_{k \to \infty} \frac{f'(x_k)}{g'(x_k)} = \lim_{k \to \infty} \frac{-x_k^2 f'(x_k)}{-x_k^2 g'(x_k)} = \lim_{k \to \infty} \frac{\tilde{f}'(1/x_k)}{\tilde{g}'(1/x_k)} = \lim_{k \to \infty} \frac{\tilde{f}(1/x_k)}{\tilde{g}(1/x_k)} = \lim_{k \to \infty} \frac{f(x_k)}{g(x_k)},$$

proving $\lim_{x \to \xi} \frac{f(x)}{g(x)} = \eta$.

We now assume (b), still letting $(x_k)_{k \in \mathbb{N}}$ be as before. Note that $g' \neq 0$ implies g is injective by Rolle's Th. 9.17. Then the intermediate value theorem implies g is either strictly increasing or strictly decreasing. We proceed with the proof for the case $I =\,]a, \xi[$, the proof for $I =\,]\xi, b[$ can be done completely analogous. We first consider the case where g is strictly increasing, i.e. $\lim_{x \to \xi} g(x) = \infty$. Assume $\eta \in \mathbb{R}$ and $\epsilon > 0$. Then $\lim_{x \to \xi} \frac{f'(x)}{g'(x)} = \eta$ and $\lim_{x \to \xi} g(x) = \infty$ imply

$$\mathop{\exists}_{c \in]a, \xi[} \quad \mathop{\forall}_{x \in]c, \xi[} \quad \left(g(x) > 0 \quad \wedge \quad \eta - \frac{\epsilon}{2} < \frac{f'(x)}{g'(x)} < \eta + \frac{\epsilon}{2} \right).$$

Since $\lim_{k \to \infty} x_k = \xi$, there exists $N_0 \in \mathbb{N}$ such that, for each $k > N_0$, $c < x_k < \xi$. Next, according to Th. 9.25,

$$\mathop{\forall}_{k > N_0} \quad \mathop{\exists}_{\xi_k \in]c, x_k[} \quad \eta - \frac{\epsilon}{2} < \frac{f(x_k) - f(c)}{g(x_k) - g(c)} = \frac{f'(\xi_k)}{g'(\xi_k)} < \eta + \frac{\epsilon}{2}.$$

In consequence, using $g(x_k) > g(c)$, as g is strictly increasing,

$$\mathop{\forall}_{k > N_0} \quad \left(\eta - \frac{\epsilon}{2} \right) (g(x_k) - g(c)) < f(x_k) - f(c) < \left(\eta + \frac{\epsilon}{2} \right) (g(x_k) - g(c))$$

and

$$\mathop{\forall}_{k > N_0} \quad \left(\eta - \frac{\epsilon}{2} \right) + \frac{f(c) - \left(\eta - \frac{\epsilon}{2} \right) g(c)}{g(x_k)} < \frac{f(x_k)}{g(x_k)} < \left(\eta + \frac{\epsilon}{2} \right) + \frac{f(c) - \left(\eta + \frac{\epsilon}{2} \right) g(c)}{g(x_k)}.$$

Since $\lim_{k \to \infty} g(x_k) = \infty$,

$$\mathop{\exists}_{N \geq N_0} \quad \mathop{\forall}_{k > N} \quad \left(\left| \frac{f(c) - \left(\eta - \frac{\epsilon}{2} \right) g(c)}{g(x_k)} \right| < \frac{\epsilon}{2} \quad \wedge \quad \left| \frac{f(c) - \left(\eta + \frac{\epsilon}{2} \right) g(c)}{g(x_k)} \right| < \frac{\epsilon}{2} \right),$$

that means

$$\underset{k>N}{\forall} \quad \eta - \epsilon < \frac{f(x_k)}{g(x_k)} < \eta + \epsilon,$$

proving $\lim_{x\to\xi} \frac{f(x)}{g(x)} = \eta$. For $\eta = \infty$ and given $n \in \mathbb{N}$, the argument is similar: $\lim_{x\to\xi} \frac{f'(x)}{g'(x)} = \eta$ and $\lim_{x\to\xi} g(x) = \infty$ imply

$$\underset{c\in]a,\xi[}{\exists} \quad \underset{x\in]c,\xi[}{\forall} \quad \left(g(x) > 0 \quad \wedge \quad n < \frac{f'(x)}{g'(x)} \right).$$

As before, since $\lim_{k\to\infty} x_k = \xi$, there exists $N_0 \in \mathbb{N}$ such that, for each $k > N_0$, $c < x_k < \xi$. Again, according to Th. 9.25,

$$\underset{k>N_0}{\forall} \quad \underset{\xi_k\in]c,x_k[}{\exists} \quad n < \frac{f(x_k) - f(c)}{g(x_k) - g(c)} = \frac{f'(\xi_k)}{g'(\xi_k)}.$$

In consequence, using $g(x_k) > g(c)$, as g is strictly increasing,

$$\underset{k>N_0}{\forall} \quad n\left(g(x_k) - g(c)\right) < f(x_k) - f(c)$$

and

$$\underset{k>N_0}{\forall} \quad n + \frac{f(c) - n\, g(c)}{g(x_k)} < \frac{f(x_k)}{g(x_k)}.$$

Since $\lim_{k\to\infty} g(x_k) = \infty$,

$$\underset{N\geq N_0}{\exists} \quad \underset{k>N}{\forall} \quad \left| \frac{f(c) - n\, g(c)}{g(x_k)} \right| < 1,$$

that means

$$\underset{k>N}{\forall} \quad n - 1 < \frac{f(x_k)}{g(x_k)},$$

proving $\lim_{x\to\xi} \frac{f(x)}{g(x)} = \eta$. If $\eta = -\infty$, then, using what we have already shown,

$$\lim_{x\to\xi} \frac{f'(x)}{g'(x)} = \eta \quad \Rightarrow \quad \lim_{x\to\xi} \frac{-f'(x)}{g'(x)} = \infty = \lim_{x\to\xi} \frac{-f(x)}{g(x)} \quad \Rightarrow \quad \lim_{x\to\xi} \frac{f(x)}{g(x)} = \eta.$$

Finally, if g strictly decreasing, then $-g$ is strictly increasing and we obtain

$$\lim_{x\to\xi} \frac{f'(x)}{g'(x)} = \eta \quad \Rightarrow \quad \lim_{x\to\xi} \frac{f'(x)}{-g'(x)} = -\eta = \lim_{x\to\xi} \frac{f(x)}{-g(x)} \quad \Rightarrow \quad \lim_{x\to\xi} \frac{f(x)}{g(x)} = \eta,$$

concluding the proof.   ∎

Example 9.27. (a) Applying L'Hôpital's rule to $f :]-\pi/2, \pi/2[\longrightarrow \mathbb{R}$, $f(x) := \tan x$, $g :]-\pi/2, \pi/2[\longrightarrow \mathbb{R}$, $g(x) := e^x - 1$, with $\xi = 0$ yields

$$\lim_{x\to 0} \frac{\tan x}{e^x - 1} = \lim_{x\to 0} \frac{1 + \tan^2 x}{e^x} = \frac{1}{1} = 1 \qquad (9.31)$$

(note $g'(x) = e^x \neq 0$ for each $x \in]-\pi/2, \pi/2[$).

(b) It can happen that a single application of L'Hôpital's rule does not, yet, yield a useful result, but that a repeated application does. An example is provided by considering $\alpha > 0$, $n \in \mathbb{N}$, and $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $f(x) := e^{\alpha x}$, $g : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $g(x) := x^n$, $\xi := \infty$. Applying L'Hôpital's rule n times yields

$$\underset{\alpha \in \mathbb{R}^+}{\forall} \quad \underset{n \in \mathbb{N}}{\forall} \quad \lim_{x \to \infty} \frac{e^{\alpha x}}{x^n} = \lim_{x \to \infty} \frac{\alpha^n e^{\alpha x}}{n!} = \infty \tag{9.32}$$

(note $g^{(k)}(x) = n(n-1)\cdots(n-k+1)x^{n-k} \neq 0$ for each $k \in \{1, \ldots, n\}$ and each $x \in \mathbb{R}^+$).

(c) It can also happen that even repeated applications of L'Hôpital's rule do not help at all, even though $\lim_{x \to \xi} \frac{f(x)}{g(x)}$ does exist and the hypotheses of Th. 9.26 are all satisfied. A simple example is given by $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) := e^x$, $g : \mathbb{R} \longrightarrow \mathbb{R}$, $g(x) := 2e^x$, and $\xi = -\infty$. Even though $\lim_{x \to -\infty} \frac{f(x)}{g(x)} = \frac{1}{2}$, one has $\lim_{x \to -\infty} f^{(n)}(x) = \lim_{x \to -\infty} g^{(n)}(x) = 0$ for every $n \in \mathbb{N}$.

9.5 Convex Functions

In the present section, we provide an introduction to (one-dimensional) convex functions. They have many important applications, some of which we well need in Analysis II, when studying so-called norms on $\mathbb{K}^n$.

The idea is to call a function $f : I \longrightarrow \mathbb{R}$ (where $I \subseteq \mathbb{R}$ is an interval) convex if, and only if, each line segment connecting two points on the graph of f lies above this graph, and to call f concave if, and only if, each such line segment lies below the graph of f. Noting that, for $x_1 < x_2$, the line through the two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ is represented by the equation

$$L(x) = \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2), \tag{9.33}$$

this leads to the following definition:

Definition 9.28. Let $I \subseteq \mathbb{R}$ be an interval (I can be open, closed, or half-open, it can be for finite or of infinite length) and $f : I \longrightarrow \mathbb{R}$. Then f is called *convex* if, and only if, for each $x_1, x, x_2 \in I$ such that $x_1 < x < x_2$, one has

$$f(x) \leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2); \tag{9.34a}$$

f is called *concave* if, and only if, for each $x_1, x, x_2 \in I$ such that $x_1 < x < x_2$, one has

$$f(x) \geq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2). \tag{9.34b}$$

Moreover, f is called strictly convex (resp. strictly concave) if, and only if, (9.34a) (resp. (9.34b)) always holds with strict inequality.

Lemma 9.29. *Let $I \subseteq \mathbb{R}$ be an interval. Then $f : I \longrightarrow \mathbb{R}$ is (strictly) concave if, and only if, $-f$ is (strictly) convex.*

Proof. Merely multiply (9.34b) by (-1) and compare with (9.34a). ∎

The following Prop. 9.30 provides equivalences for convexity. One can easily obtain the corresponding equivalences for concavity by combining Prop. 9.30 with Lem. 9.29.

Proposition 9.30. *Let $I \subseteq \mathbb{R}$ be an interval and $f : I \longrightarrow \mathbb{R}$. Then the following statements are equivalent:*

  (i) *f is (strictly) convex.*

  (ii) *For each $a, b \in I$ such that $a \neq b$ and each $\lambda \in]0, 1[$, the following estimate holds (with strict inequality):*

$$f\big(\lambda a + (1 - \lambda)b\big) \leq \lambda f(a) + (1 - \lambda)f(b). \tag{9.35}$$

  (iii) *For each $x_1, x, x_2 \in I$ such that $x_1 < x < x_2$, one has (with strict inequality)*

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}. \tag{9.36}$$

  (iv) *For each $x_1, x, x_2 \in I$ such that $x_1 < x < x_2$, one has (with strict inequality)*

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}. \tag{9.37}$$

Proof. We leave the proof as an exercise. Suggestion: Establish the following implications: (i) $\Leftrightarrow$ (ii), (i) $\Leftrightarrow$ (iii), (iv) $\Rightarrow$ (iii), (i) $\Rightarrow$ (iv). ∎

Example 9.31. Since $|\lambda x + (1 - \lambda)y| \leq \lambda|x| + (1 - \lambda)|y|$ for each $0 < \lambda < 1$ and each $x, y \in \mathbb{R}$, the absolute value function is convex. This example also shows that a convex function does not need to be differentiable.

—

For differentiable functions, one can formulate convexity criteria in terms of the derivative:

Proposition 9.32. *Let $a < b$, and suppose that $f : [a, b] \longrightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on $]a, b[$. Then f is (strictly) convex (resp. (strictly) concave) on $[a, b]$ if, and only if, the derivative f' is (strictly) increasing (resp. (strictly) decreasing) on $]a, b[$.*

Proof. Since $(-f)' = -f'$ and $-f'$ is (strictly) increasing if, and only if, f' is (strictly) decreasing, it suffices to consider the (strictly) convex case. So assume that f is (strictly) convex. Then for each $x_1, x, x_0, y, x_2 \in]a, b[$ such that $x_1 < x < x_0 < y < x_2$, applying Prop. 9.30(iv), one has (with strict inequalities),

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_0) - f(x_1)}{x_0 - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(y)}{x_2 - y}. \tag{9.38}$$

Thus,

$$f'(x_1) = \lim_{x \downarrow x_1} \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_0) - f(x_1)}{x_0 - x_1} \overset{(*)}{\leq} \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

$$\leq \lim_{x \uparrow x_2} \frac{f(x_2) - f(x)}{x_2 - x} = f'(x_2) \tag{9.39}$$

(where the inequality at $(*)$ is strict if it is strict in (9.38)), showing that f' is (strictly) increasing on $]a, b[$. On the other hand, if f' is (strictly) increasing on $]a, b[$, then for each $x_1, x, x_2 \in [a, b]$ such that $x_1 < x < x_2$, Th. 9.18 yields $\xi_1 \in]x_1, x[$ and $\xi_2 \in]x, x_2[$ such that

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \quad \text{and} \quad \frac{f(x_2) - f(x)}{x_2 - x} = f'(\xi_2). \tag{9.40}$$

As $\xi_1 < \xi_2$ and f' is (strictly) increasing, (9.40) implies (9.36) (with strict inequality) and, thus, the (strict) convexity of f. ∎

Proposition 9.33. *Let $a < b$, and suppose that $f : [a, b] \longrightarrow \mathbb{R}$ is continuous on $[a, b]$ and twice differentiable on $]a, b[$.*

(a) *f is convex (resp. concave) on $[a, b]$ if, and only if, $f'' \geq 0$ (resp. $f'' \leq 0$) on $]a, b[$.*

(b) *If $f'' > 0$ (resp. $f'' < 0$) on $]a, b[$, then f is strictly convex (resp. strictly concave) (as a caveat we remark that, here, the converse does not hold – for example $x \mapsto x^4$ is strictly convex, but its second derivative $x \mapsto 12x^2$ is 0 at $x = 0$).*

Proof. Since $-f'' \geq 0$ if, and only if $f'' \leq 0$; and $-f'' > 0$ if, and only if $f'' < 0$, if suffices to consider the convex cases. Moreover, for (a), one merely has to combine Prop. 9.32 with the fact that f' is increasing on $]a, b[$ if, and only if, $f'' \geq 0$ on $]a, b[$. The proof of (b) is left as an exercise. ∎

Example 9.34. (a) Since for $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = e^x$, it is $f''(x) = e^x > 0$, the exponential function is strictly convex on $\mathbb{R}$.

(b) Since for $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $f(x) = \ln x$, it is $f''(x) = -1/x^2 < 0$, the natural logarithm is strictly concave on $\mathbb{R}^+$.

Theorem 9.35 (Jensen's inequality). *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \longrightarrow \mathbb{R}$ be convex. If $n \in \mathbb{N}$ and $\lambda_1, \ldots, \lambda_n > 0$ such that $\lambda_1 + \cdots + \lambda_n = 1$, then*

$$\underset{x_1,\ldots,x_n \in I}{\forall} \quad f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n). \tag{9.41a}$$

If f is concave, then

$$\underset{x_1,\ldots,x_n\in I}{\forall}\quad f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \geq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n). \tag{9.41b}$$

If f is strictly convex or strictly concave, then equality in the above inequalities can only hold if $x_1 = \cdots = x_n$.

Proof. If one lets $a := \min\{x_1,\ldots,x_n\}$, $b := \max\{x_1,\ldots,x_n\}$, and $\bar{x} := \lambda_1 x_1 + \cdots + \lambda_n x_n$, then

$$a = \sum_{j=1}^{n}\lambda_j\, a \leq \bar{x} \leq \sum_{j=1}^{n}\lambda_j\, b = b \quad \Rightarrow \quad \bar{x} \in I. \tag{9.42}$$

Since f is (strictly) concave if, and only if, $-f$ is (strictly) convex, it suffices to consider the cases where f is convex and where f is strictly convex. Thus, we assume that f is convex and prove (9.41a) by induction. For $n = 1$, one has $\lambda_1 = 1$ and there is nothing to prove. For $n = 2$, (9.41a) reduces to (9.35), which holds due to the convexity of f. Finally, let $n > 2$ and assume that (9.41a) already holds for each $1 \leq l \leq n - 1$. Set

$$\lambda := \lambda_1 + \cdots + \lambda_{n-1}, \quad x := \frac{\lambda_1}{\lambda}x_1 + \cdots + \frac{\lambda_{n-1}}{\lambda}x_{n-1}. \tag{9.43}$$

Then $x \in I$ follows as in (9.42). One computes

$$
\begin{aligned}
f(\lambda_1 x_1 + \cdots + \lambda_n x_n) &= f\left(\sum_{j=1}^{n-1}\lambda_j x_j + \lambda_n x_n\right) = f(\lambda x + \lambda_n x_n) \\
&\overset{l=2}{\leq} \lambda f(x) + \lambda_n f(x_n) \overset{l=n-1}{\leq} \lambda \sum_{j=1}^{n-1}\frac{\lambda_j}{\lambda}f(x_j) + \lambda_n f(x_n) \\
&= \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n),
\end{aligned}
\tag{9.44}
$$

thereby completing the induction, and, thus, the proof of (9.41a). If f is strictly convex and (9.41a) holds with equality, then one can also proceed by induction to prove the equality of the x_j. Again, if $n = 1$, then there is nothing to prove. If $n = 2$, and $x_1 \neq x_2$, then strict convexity requires (9.41a) to hold with strict inequality. Thus $x_1 = x_2$. Now let $n > 2$. It is noted that (9.44) still holds. By hypothesis, the first and last term in (9.44) are now equal, implying that all terms in (9.44) must be equal. Using the induction hypothesis for $l = 2$ and the corresponding equality in (9.44), we conclude that $x = x_n$. Using the induction hypothesis for $l = n-1$ and the corresponding equality in (9.44), we conclude that $x_1 = \cdots = x_{n-1}$. Finally, $x = x_n$ and $x_1 = \cdots = x_{n-1}$ are combined using (9.43) to get $x_1 = x_n$, finishing the proof of the theorem. ∎

Theorem 9.36 (Inequality Between the Weighted Arithmetic Mean and the Weighted Geometric Mean). *If $n \in \mathbb{N}$, $x_1,\ldots,x_n \geq 0$ and $\lambda_1,\ldots,\lambda_n > 0$ such that $\lambda_1+\cdots+\lambda_n = 1$, then*

$$x_1^{\lambda_1}\cdots x_n^{\lambda_n} \leq \lambda_1 x_1 + \cdots + \lambda_n x_n, \tag{9.45}$$

where equality occurs if, and only if, $x_1 = \cdots = x_n$. In particular, for $\lambda_1 = \cdots = \lambda_n = \frac{1}{n}$, one recovers the inequality between the arithmetic and the geometric mean without weights, known from Th. 7.63.

Proof. If at least one of the x_j is 0, then (9.45) becomes the true statement $0 \leq \sum_{j=1}^{n} \lambda_j x_j$ with strict inequality if, and only if, at least one $x_j > 0$. Thus, it remains to consider the case $x_1, \ldots, x_n > 0$. As we noted in Ex. 9.34(b), the natural logarithm $\ln : \mathbb{R}^+ \longrightarrow \mathbb{R}$ is concave and even strictly concave. Employing Jensen's inequality (9.41b) yields

$$\ln(\lambda_1 x_1 + \cdots + \lambda_n x_n) \geq \lambda_1 \ln x_1 + \cdots + \lambda_n \ln x_n = \ln(x_1^{\lambda_1} \cdots x_n^{\lambda_n}). \tag{9.46}$$

Applying the exponential function to both sides of (9.46), one obtains (9.45). Since (9.46) is equivalent to (9.45), the strict concavity of $\ln$ yields that equality in (9.46) implies $x_1 = \cdots = x_n$. ■

10   The Riemann Integral on Intervals in $\mathbb{R}$

10.1   Definition and Simple Properties

Given a nonnegative function $f : M \longrightarrow \mathbb{R}_0^+$, $M \subseteq \mathbb{R}$, we aim to compute the area $\int_M f$ of the set "under the graph" of f, i.e. of the set

$$\big\{ (x,y) \in \mathbb{R}^2 : x \in M \text{ and } 0 \leq y \leq f(x) \big\}. \tag{10.1}$$

This area $\int_M f$ (if it exists) will be called the *integral* of f over M. Moreover, for functions $f : M \longrightarrow \mathbb{R}$ that are not necessarily nonnegative, we would like to count areas of sets of the form (10.1) (which are below the graph of f and above the set $M \cong \big\{ (x,0) \in \mathbb{R}^2 : x \in M \big\} \subseteq \mathbb{R}^2$) with a positive sign, and whereas we would like to count areas of sets above the graph of f and below the set M with a negative sign. In other words, making use of the positive and negative parts f^+ and f^- of $f = f^+ - f^-$ as defined in (6.1i) and (6.1j), respectively, we would like our integral to satisfy

$$\int_M f = \int_M f^+ - \int_M f^-. \tag{10.2}$$

Difficulties arise from the fact that both the function f and the set M can be extremely complicated. To avoid dealing with complicated sets M, we restrict ourselves to the situation of integrals over compact intervals, i.e. to integrals over sets of the form $M = [a,b]$. Moreover, we will also restrict ourselves to bounded functions f, which we now define:

Definition 10.1. Let $\emptyset \neq M \subseteq \mathbb{R}$ and $f : M \longrightarrow \mathbb{R}$. Then f is called *bounded* if, and only if, the set $\{|f(x)| : x \in M\} \subseteq \mathbb{R}_0^+$ is bounded, i.e. if, and only if,

$$\|f\|_{\sup} := \sup\{|f(x)| : x \in M\} \in \mathbb{R}_0^+. \tag{10.3}$$

The basic idea for the definition of the Riemann integral $\int_M f$ is rather simple: Decompose the set M into small pieces $I_1, \ldots, I_N$ and approximate $\int_M f$ by the finite sum

$\sum_{j=1}^{N} f(x_j)|I_j|$, where $x_j \in I_j$ and $|I_j|$ denotes the size of the set I_j. Define $\int_M f$ as the limit of such sums as the size of the I_j tends to zero (if the limit exists). However, to carry out this idea precisely and rigorously does require some work.

As stated before, we will assume that M is a closed bounded interval, and we will choose the I_j to be closed bounded intervals as well. To emphasize we are dealing with intervals, in the following, we will prefer to use the symbol I instead of M.

Definition 10.2. If $a, b \in \mathbb{R}$, $a \le b$, and $I := [a, b]$, then we call

$$|I| := b - a = |a - b|, \tag{10.4}$$

the *length* or the (1-dimensional) *size*, *volume*, or *measure* of I.

Definition 10.3. Given a real interval $I := [a, b] \subseteq \mathbb{R}$, $a, b \in \mathbb{R}$, $a < b$, the $(N+1)$-tuple $\Delta := (x_0, \dots, x_N) \in \mathbb{R}^{N+1}$, $N \in \mathbb{N}$, is called a *partition* of I if, and only if, $a = x_0 < x_1 < \cdots < x_N = b$. We call $x_0, \dots, x_N$ the *nodes* of Δ, and let $\nu(\Delta) := \{x_0, \dots, x_N\}$ be the set of all nodes. A *tagged partition* of I is a partition together with an N-tuple $(t_1, \dots, t_N) \in \mathbb{R}^N$ such that $t_j \in [x_{j-1}, x_j]$ for each $j \in \{1, \dots, N\}$. Given a partition Δ (with or without tags) of I as above and letting $I_j := [x_{j-1}, x_j]$, the number

$$|\Delta| := \max\{|I_j| : j \in \{1, \dots, N\}\}, \tag{10.5}$$

is called the *mesh size* of Δ. It is sometimes convenient, if we extend our definitions to trivial intervals, consisting of just one point: For $a = b$, we have $I = [a, a] = \{a\}$. We then define $\Delta = x_0 = a$ to be a partition of I, $\nu(\Delta) = \{x_0\}$, and a is then the only tag that makes Δ into a tagged partition. We also set $I_0 := I = \{a\}$, and the mesh size in this case is $|\Delta| := 0$.

Definition 10.4. Let Δ be a partition of $I = [a, b] \subseteq \mathbb{R}$, $a \le b$, as in Def. 10.3. Given a function $f : I \longrightarrow \mathbb{R}$ that is bounded according to Def. 10.1, define

$$m_j := m_j(f) := \inf\{f(x) : x \in I_j\}, \quad M_j := M_j(f) := \sup\{f(x) : x \in I_j\}, \tag{10.6}$$

and

$$r(\Delta, f) := \sum_{j=1}^{N} m_j |I_j| = \sum_{j=1}^{N} m_j(x_j - x_{j-1}), \tag{10.7a}$$

$$R(\Delta, f) := \sum_{j=1}^{N} M_j |I_j| = \sum_{j=1}^{N} M_j(x_j - x_{j-1}), \tag{10.7b}$$

where $r(\Delta, f)$ is called the *lower Riemann sum* and $R(\Delta, f)$ is called the *upper Riemann sum* associated with Δ and f. If Δ is tagged by $\tau := (t_1, \dots, t_N)$, then we also define the *intermediate Riemann sum*

$$\rho(\Delta, f) := \sum_{j=1}^{N} f(t_j) |I_j| = \sum_{j=1}^{N} f(t_j)(x_j - x_{j-1}). \tag{10.7c}$$

Note that, for $a = b$, all the above sums are empty and we have $r(\Delta, f) = R(\Delta, f) = \rho(\Delta, f) = 0$.

Definition 10.5. Let $I = [a, b] \subseteq \mathbb{R}$ be an interval, $a \leq b$, and suppose $f : I \longrightarrow \mathbb{R}$ is bounded.

(a) Define

$$J_*(f, I) := \sup \{r(\Delta, f) : \Delta \text{ is a partition of } I\}, \qquad (10.8a)$$
$$J^*(f, I) := \inf \{R(\Delta, f) : \Delta \text{ is a partition of } I\}. \qquad (10.8b)$$

We call $J_*(f, I)$ the *lower Riemann integral* of f over I and $J^*(f, I)$ the *upper Riemann integral* of f over I.

(b) The function f is called *Riemann integrable* over I if, and only if, $J_*(f, I) = J^*(f, I)$. If f is Riemann integrable over I, then

$$\int_a^b f(x)\, dx := \int_I f(x)\, dx := \int_a^b f := \int_I f := J_*(f, I) = J^*(f, I) \qquad (10.9)$$

is called the *Riemann integral* of f over I. The set of all functions $f : I \longrightarrow \mathbb{R}$ that are Riemann integrable over I is denoted by $\mathcal{R}(I, \mathbb{R})$ or just by $\mathcal{R}(I)$.

(c) The function $g : I \longrightarrow \mathbb{C}$ is called *Riemann integrable* over I if, and only if, both $\operatorname{Re} g$ and $\operatorname{Im} g$ are Riemann integrable. The set of all Riemann integrable functions $g : I \longrightarrow \mathbb{C}$ is denoted by $\mathcal{R}(I, \mathbb{C})$. If $g \in \mathcal{R}(I, \mathbb{C})$, then

$$\int_I g := \left(\int_I \operatorname{Re} g, \int_I \operatorname{Im} g \right) = \int_I \operatorname{Re} g + i \int_I \operatorname{Im} g \in \mathbb{C} \qquad (10.10)$$

is called the Riemann integral of g over I.

Remark 10.6. If $I = [a, b] \subseteq \mathbb{R}$, Δ is a partition of I, and $f : I \longrightarrow \mathbb{R}$ is bounded, then (10.6) implies

$$m_j(f) \overset{(4.6c)}{=} -M_j(-f) \qquad \text{and} \qquad m_j(-f) \overset{(4.6d)}{=} -M_j(f), \qquad (10.11a)$$

(10.7) implies

$$r(\Delta, f) = -R(\Delta, -f) \qquad \text{and} \qquad r(\Delta, -f) = -R(\Delta, f), \qquad (10.11b)$$

and (10.8) implies

$$J_*(f, I) = -J^*(-f, I) \qquad \text{and} \qquad J_*(-f, I) = -J^*(f, I). \qquad (10.11c)$$

Example 10.7. (a) If $I = [a, b] \subseteq \mathbb{R}$ as before and $f : I \longrightarrow \mathbb{R}$ is constant, i.e. $f \equiv c$ with $c \in \mathbb{R}$, then $f \in \mathcal{R}(I)$ and

$$\int_a^b f = c\,(b - a) = c\,|I| : \qquad (10.12)$$

We have, for each partition Δ of I,

$$r(\Delta, f) = \sum_{j=1}^{N} m_j \, |I_j| = c \sum_{j=1}^{N} |I_j| = c \, |I| = c \, (b - a) = \sum_{j=1}^{N} M_j \, |I_j| = R(\Delta, f),$$

$$(10.13)$$

proving $J_*(f, I) = c \, (b - a) = J^*(f, I)$.

(b) An example of a function that is not Riemann integrable for $a < b$ is given by the *Dirichlet function*

$$f : [a, b] \longrightarrow \mathbb{R}, \quad f(x) := \begin{cases} 0 & \text{for } x \text{ irrational,} \\ 1 & \text{for } x \text{ rational,} \end{cases} \quad a < b. \quad (10.14)$$

Since $r(\Delta, f) = 0$ and $R(\Delta, f) = \sum_{j=1}^{N} |I_j| = b - a$ for every partition Δ of I, one obtains $J_*(f, I) = 0 \neq (b - a) = J^*(f, I)$, showing that $f \notin \mathcal{R}(I)$.

Definition 10.8. (a) If Δ is a partition of $[a, b] \subseteq \mathbb{R}$ as in Def. 10.3, then another partition Δ' of $[a, b]$ is called a *refinement* of Δ if, and only if, $\nu(\Delta) \subseteq \nu(\Delta')$, i.e. if, and only if, the nodes of Δ' include all the nodes of Δ.

(b) If Δ and Δ' are partitions of $[a, b] \subseteq \mathbb{R}$, then the *superposition* of Δ and Δ', denoted $\Delta + \Delta'$, is the unique partition of $[a, b]$ having $\nu(\Delta) \cup \nu(\Delta')$ as its set of nodes. Note that the superposition of Δ and Δ' is always a common refinement of Δ and Δ'.

Lemma 10.9. *Let $a, b \in \mathbb{R}$, $a < b$, $I := [a, b]$, and suppose $f : I \longrightarrow \mathbb{R}$ is bounded with $M := \|f\|_{\sup} \in \mathbb{R}_0^+$. Let Δ' be a partition of I and assume*

$$\alpha := \#\Big(\nu(\Delta') \setminus \{a, b\}\Big) \geq 1 \quad (10.15)$$

is the number of interior nodes that occur in Δ'. Then, for each partition Δ of I, the following holds:

$$r(\Delta, f) \leq r(\Delta + \Delta', f) \leq r(\Delta, f) + 2 \, \alpha \, M \, |\Delta|, \quad (10.16a)$$

$$R(\Delta, f) \geq R(\Delta + \Delta', f) \geq R(\Delta, f) - 2 \, \alpha \, M \, |\Delta|. \quad (10.16b)$$

Proof. We carry out the proof of (10.16a) – the proof of (10.16b) can be conducted completely analogous. Consider the case $\alpha = 1$ and let ξ be the single element of $\nu(\Delta') \setminus \{a, b\}$. If $\xi \in \nu(\Delta)$, then $\Delta + \Delta' = \Delta$, and (10.16a) is trivially true. If $\xi \notin \nu(\Delta)$, then $x_{k-1} < \xi < x_k$ for a suitable $k \in \{1, \ldots, N\}$. Define

$$I' := [x_{k-1}, \xi], \quad I'' := [\xi, x_k] \quad (10.17)$$

and

$$m' := \inf\{f(x) : x \in I'\}, \quad m'' := \inf\{f(x) : x \in I''\}. \quad (10.18)$$

Then we obtain

$$r(\Delta + \Delta', f) - r(\Delta, f) = m' \, |I'| + m'' \, |I''| - m_k \, |I_k| = (m' - m_k) \, |I'| + (m'' - m_k) \, |I''|. \quad (10.19)$$

Together with the observation

$$0 \le m' - m_k \le 2M, \quad 0 \le m'' - m_k \le 2M, \tag{10.20}$$

(10.19) implies

$$0 \le r(\Delta + \Delta', f) - r(\Delta, f) \le 2M \left(|I'| + |I''| \right) \le 2M \, |\Delta|. \tag{10.21}$$

The general form of (10.16a) follows by an induction on α. ∎

Theorem 10.10. *Let* $a, b \in \mathbb{R}$, $a \le b$, $I := [a, b]$, *and let* $f : I \longrightarrow \mathbb{R}$ *be bounded.*

(a) *Suppose* Δ *and* Δ' *are partitions of* I *such that* Δ' *is a refinement of* Δ. *Then*

$$r(\Delta, f) \le r(\Delta', f), \quad R(\Delta, f) \ge R(\Delta', f). \tag{10.22}$$

(b) *For arbitrary partitions* Δ *and* Δ', *the following holds:*

$$r(\Delta, f) \le R(\Delta', f). \tag{10.23}$$

(c) $J_*(f, I) \le J^*(f, I)$.

(d) *For each sequence of partitions* $(\Delta_n)_{n \in \mathbb{N}}$ *of* I *such that* $\lim_{n \to \infty} |\Delta_n| = 0$, *one has*

$$\lim_{n \to \infty} r(\Delta_n, f) = J_*(f, I), \quad \lim_{n \to \infty} R(\Delta_n, f) = J^*(f, I). \tag{10.24}$$

In particular, if $f \in \mathcal{R}(I)$, *then*

$$\lim_{n \to \infty} r(\Delta_n, f) = \lim_{n \to \infty} R(\Delta_n, f) = \int_I f, \tag{10.25a}$$

and if $f \in \mathcal{R}(I)$ *and the* Δ_n *are tagged, then also*

$$\lim_{n \to \infty} \rho(\Delta_n, f) = \int_I f. \tag{10.25b}$$

Proof. (a): If Δ' is a refinement of Δ, then $\Delta' = \Delta + \Delta'$. Thus, (10.22) is immediate from (10.16).

(b): This also follows from (10.16):

$$r(\Delta, f) \overset{(10.16a)}{\le} r(\Delta + \Delta', f) \overset{(10.7)}{\le} R(\Delta + \Delta', f) \overset{(10.16b)}{\le} R(\Delta', f). \tag{10.26}$$

(c): One just combines (10.8) with (b).

(d): For $a = b$, there is nothing to show. For $a < b$, let $(\Delta_n)_{n \in \mathbb{N}}$ be a sequence of partitions of I such that $\lim_{n \to \infty} |\Delta_n| = 0$, and let Δ' be an arbitrary partition of I with numbers α and M defined as in Lem. 10.9. Then, according to (10.16a):

$$r(\Delta_n, f) \le r(\Delta_n + \Delta', f) \le r(\Delta_n, f) + 2 \alpha M \, |\Delta_n| \quad \text{for each } n \in \mathbb{N}. \tag{10.27}$$

From (b), we conclude the sequence $\big(r(\Delta_n, f)\big)_{n\in\mathbb{N}}$ is bounded. According to the Bolza-no-Weierstrass Th. 7.27, if we can show that the sequence has $J_*(f, I)$ as its only cluster point, then the first equality of (10.24) must hold. Thus, according to Prop. 7.26, it suffices to show that every converging subsequence of $(r(\Delta_n, f))_{n\in\mathbb{N}}$ converges to $J_*(f, I)$. To this end, suppose $(r(\Delta_{n_k}, f))_{k\in\mathbb{N}}$ is a converging subsequence of $(r(\Delta_n, f))_{n\in\mathbb{N}}$ with $\beta := \lim_{k\to\infty} r(\Delta_{n_k}, f)$. First note $\beta \leq J_*(f, I)$ due to the definition of $J_*(f, I)$. Moreover, (10.27) implies $\lim_{k\to\infty} r(\Delta_{n_k} + \Delta', f) = \beta$. Since $r(\Delta', f) \leq r(\Delta_{n_k} + \Delta', f)$ and Δ' is arbitrary, we obtain $J_*(f, I) \leq \beta$, i.e. $J_*(f, I) = \beta$. Thus, we have shown that, indeed, every subsequence of $(r(\Delta_n, f))_{n\in\mathbb{N}}$ converges to $\beta = J_*(f, I)$. In the same manner, one conducts the proof of $J^*(f, I) = \lim_{n\to\infty} R(\Delta_n, f)$. Then (10.25a) is immediate from the definition of Riemann integrability, and (10.25b) follows from (10.25a), since (10.7) implies $r(\Delta, f) \leq \rho(\Delta, f) \leq R(\Delta, f)$ for each tagged partition Δ of I. ∎

Theorem 10.11. *Let $a, b \in \mathbb{R}$, $a \leq b$, $I := [a, b]$.*

(a) *The integral is linear: More precisely, if $f, g \in \mathcal{R}(I, \mathbb{K})$ and $\lambda, \mu \in \mathbb{K}$, then $\lambda f + \mu g \in \mathcal{R}(I, \mathbb{K})$ and*

$$\int_I (\lambda f + \mu g) = \lambda \int_I f + \mu \int_I g. \tag{10.28}$$

(b) *Let $\tilde{\Delta} = (y_0, \ldots, y_M)$, $M \in \mathbb{N}$, be a partition of I, $J_k := [y_{k-1}, y_k]$. Then $f \in \mathcal{R}(I, \mathbb{K})$ if, and only if, $f \in \mathcal{R}(J_k, \mathbb{K})$ for each $k \in \{1, \ldots, M\}$. If $f \in \mathcal{R}(I, \mathbb{K})$, then*

$$\int_a^b f = \int_I f = \sum_{k=1}^M \int_{J_k} f = \sum_{k=1}^M \int_{y_{k-1}}^{y_k} f. \tag{10.29}$$

(c) *Monotonicity of the Integral: If $f, g : I \longrightarrow \mathbb{R}$ are bounded and $f \leq g$ (i.e. $f(x) \leq g(x)$ for each $x \in I$), then $J_*(f, I) \leq J_*(g, I)$ and $J^*(f, I) \leq J^*(g, I)$. In particular, if $f, g \in \mathcal{R}(I)$ and $f \leq g$, then*

$$\int_I f \leq \int_I g. \tag{10.30}$$

(d) *Triangle Inequality: For each $f \in \mathcal{R}(I, \mathbb{C})$, one has*

$$\left| \int_I f \right| \leq \int_I |f|. \tag{10.31}$$

Proof. (a): First, consider $\mathbb{K} = \mathbb{R}$, i.e. $f, g : I \longrightarrow \mathbb{R}$ and $\lambda, \mu \in \mathbb{R}$. For $a = b$, there is nothing to prove, so let $a < b$. Let $(\Delta_n)_{n\in\mathbb{N}}$ be a sequence of partitions of I, $\Delta_n = (x_{n,0}, \ldots, x_{n,N_n})$, $I_{n,j} := [x_{n,j-1}, x_{n,j}]$, satisfying $\lim_{n\to\infty} |\Delta_n| = 0$. Note that, for

each $n \in \mathbb{N}$ and each $j \in \{1, \ldots, N_n\}$,

$$
\begin{aligned}
m_{n,j}(f+g) \;&=\; \inf\{f(x)+g(x) : x \in I_{n,j}\} \\
&\overset{(4.6b)}{=}\; \inf\{f(x) : x \in I_{n,j}\} + \inf\{g(x) : x \in I_{n,j}\} \\
&=\; m_{n,j}(f) + m_{n,j}(g), \hspace{4cm} (10.32a) \\
M_{n,j}(f+g) \;&=\; \sup\{f(x)+g(x) : x \in I_{n,j}\} \\
&\overset{(4.6a)}{=}\; \sup\{f(x) : x \in I_{n,j}\} + \inf\{g(x) : x \in I_{n,j}\} \\
&=\; M_{n,j}(f) + M_{n,j}(g), \hspace{3.5cm} (10.32b)
\end{aligned}
$$

$$
\underset{\lambda \in \mathbb{R}}{\forall} \quad m_{n,j}(\lambda f) = \inf\{\lambda f(x) : x \in I_{n,j}\}
$$

$$
\overset{(4.6d)}{=}
\begin{cases}
\lambda \inf\{f(x) : x \in I_{n,j}\} = \lambda\, m_{n,j}(f) & \text{for } \lambda \geq 0, \\
\lambda \sup\{f(x) : x \in I_{n,j}\} = \lambda\, M_{n,j}(f) & \text{for } \lambda < 0,
\end{cases} \quad (10.32c)
$$

$$
\underset{\lambda \in \mathbb{R}}{\forall} \quad M_{n,j}(\lambda f) = \sup\{\lambda f(x) : x \in I_{n,j}\}
$$

$$
\overset{(4.6c)}{=}
\begin{cases}
\lambda \sup\{f(x) : x \in I_{n,j}\} = \lambda\, M_{n,j}(f) & \text{for } \lambda \geq 0, \\
\lambda \inf\{f(x) : x \in I_{n,j}\} = \lambda\, m_{n,j}(f) & \text{for } \lambda < 0.
\end{cases} \quad (10.32d)
$$

Thus,

$$
J_*(f+g, I) \overset{(10.24)}{=} \lim_{n \to \infty} r(\Delta_n, f+g) \overset{(10.7a)}{=} \lim_{n \to \infty} \sum_{j=1}^{N_n} m_{n,j}(f+g)\,|I_{n,j}|
$$

$$
\overset{(10.32a)}{=} \lim_{n \to \infty} \big(r(\Delta_n, f) + r(\Delta_n, g)\big) = J_*(f, I) + J_*(g, I), \quad (10.33a)
$$

$$
J^*(f+g, I) \overset{(10.24)}{=} \lim_{n \to \infty} R(\Delta_n, f+g) \overset{(10.7b)}{=} \lim_{n \to \infty} \sum_{j=1}^{N_n} M_{n,j}(f+g)\,|I_{n,j}|
$$

$$
\overset{(10.32b)}{=} \lim_{n \to \infty} \big(R(\Delta_n, f) + R(\Delta_n, g)\big) = J^*(f, I) + J^*(g, I), (10.33b)
$$

$$
\underset{\lambda \in \mathbb{R}}{\forall} \quad J_*(\lambda f, I) \overset{(10.24)}{=} \lim_{n \to \infty} r(\Delta_n, \lambda f) \overset{(10.7a)}{=} \lim_{n \to \infty} \sum_{j=1}^{N_n} m_{n,j}(\lambda f)\,|I_{n,j}|
$$

$$
\overset{(10.32c)}{=}
\begin{cases}
\lambda \lim_{n \to \infty} r(\Delta_n, f) = \lambda\, J_*(f, I) & \text{for } \lambda \geq 0, \\
\lambda \lim_{n \to \infty} R(\Delta_n, f) = \lambda\, J^*(f, I) & \text{for } \lambda < 0,
\end{cases} \quad (10.33c)
$$

$$
\underset{\lambda \in \mathbb{R}}{\forall} \quad J^*(\lambda f, I) \overset{(10.24)}{=} \lim_{n \to \infty} R(\Delta_n, \lambda f) \overset{(10.7b)}{=} \lim_{n \to \infty} \sum_{j=1}^{N_n} M_{n,j}(\lambda f)\,|I_{n,j}|
$$

$$
\overset{(10.32d)}{=}
\begin{cases}
\lambda \lim_{n \to \infty} R(\Delta_n, f) = \lambda\, J^*(f, I) & \text{for } \lambda \geq 0, \\
\lambda \lim_{n \to \infty} r(\Delta_n, f) = \lambda\, J_*(f, I) & \text{for } \lambda < 0.
\end{cases} \quad (10.33d)
$$

Thus, if f and g are both Riemann integrable over I, then we obtain $J_*(f+g, I) = J_*(f, I) + J_*(g, I) = J^*(f, I) + J^*(g, I) = J^*(f+g, I)$, i.e. $(f+g) \in \mathcal{R}(I)$; and $J_*(\lambda f, I) =$

$\lambda J_*(f, I) = \lambda J^*(f, I)$ for $\lambda \geq 0$, $J_*(\lambda f, I) = \lambda J^*(f, I) = \lambda J_*(f, I)$ for $\lambda < 0$, i.e. $(\lambda f) \in \mathcal{R}(I)$ in each case. In particular, for each $\lambda, \mu \in \mathbb{R}$,

$$\int_I (\lambda f + \mu g) = J_*(\lambda f + \mu g, I) = \lambda J_*(f, I) + \mu J_*(g, I) = \lambda \int_I f + \mu \int_I g,$$

proving (10.28) for $\mathbb{K} = \mathbb{R}$. It remains to consider $f, g \in \mathcal{R}(I, \mathbb{C})$ and $\lambda, \mu \in \mathbb{C}$. One computes, using the real-valued case,

$$\int_I (\lambda f) = \left(\int_I (\operatorname{Re}\lambda \operatorname{Re} f - \operatorname{Im}\lambda \operatorname{Im} f), \int_I (\operatorname{Re}\lambda \operatorname{Im} f + \operatorname{Im}\lambda \operatorname{Re} f) \right)$$

$$= \left(\operatorname{Re}\lambda \int_I \operatorname{Re} f - \operatorname{Im}\lambda \int_I \operatorname{Im} f, \ \operatorname{Re}\lambda \int_I \operatorname{Im} f + \operatorname{Im}\lambda \int_I \operatorname{Re} f \right)$$

$$= \lambda \int_I f$$

and

$$\int_I (f + g) = \left(\int_I \operatorname{Re}(f + g), \int_I \operatorname{Im}(f + g) \right) = \left(\int_I \operatorname{Re} f + \int_I \operatorname{Re} g, \int_I \operatorname{Im} g + \int_I \operatorname{Im} g \right)$$

$$= \left(\int_I \operatorname{Re} f, \int_I \operatorname{Im} f \right) + \left(\int_I \operatorname{Re} g, \int_I \operatorname{Im} g \right) = \int_I f + \int_I g.$$

(b): Once again, consider $\mathbb{K} = \mathbb{R}$ first. For $a = b$, there is nothing to prove, so let $a < b$. For $M = 1$, there is still nothing to prove. For $M = 2$, we have $a = y_0 < y_1 < y_2 = b$. Consider a sequence $(\Delta_n)_{n \in \mathbb{N}}$ of partitions of I, $\Delta_n = (x_{n,0}, \ldots, x_{n,N_n})$, such that $\lim_{n \to \infty} |\Delta_n| = 0$ and $y_1 \in \nu(\Delta_n)$ for each $n \in \mathbb{N}$. Define $\Delta'_n := (x_{n,0}, \ldots, y_1)$, $\Delta''_n := (y_1, \ldots, x_{n,N_n})$. Then Δ'_n and Δ''_n are partitions of J_1 and J_2, respectively, and $\lim_{n \to \infty} |\Delta'_n| = \lim_{n \to \infty} |\Delta''_n| = 0$. Moreover,

$$\underset{n \in \mathbb{N}}{\forall} \ \left(r(\Delta_n, f) = r(\Delta'_n, f) + r(\Delta''_n, f), \quad R(\Delta_n, f) = R(\Delta'_n, f) + R(\Delta''_n, f) \right),$$

implying $J_*(f, I) = J_*(f, J_1) + J_*(f, J_2)$ and $J^*(f, I) = J^*(f, J_1) + J^*(f, J_2)$. This proves $\int_I f = \int_{J_1} f + \int_{J_2} f$ provided $f \in \mathcal{R}(I) \cap \mathcal{R}(J_1) \cap \mathcal{R}(J_2)$. So it just remains to show the claimed equivalence between $f \in \mathcal{R}(I)$ and $f \in \mathcal{R}(J_1) \cap \mathcal{R}(J_2)$. If $f \in \mathcal{R}(J_1) \cap \mathcal{R}(J_2)$, then $J_*(f, I) = J_*(f, J_1) + J_*(f, J_2) = J^*(f, J_1) + J^*(f, J_2) = J^*(f, I)$, showing $f \in \mathcal{R}(I)$. Conversely, $J_*(f, I) = J^*(f, I)$ implies $J_*(f, J_1) = J^*(f, J_1) + J^*(f, J_2) - J_*(f, J_2) \geq J^*(f, J_1)$, showing $J_*(f, J_1) = J^*(f, J_1)$ and $f \in \mathcal{R}(J_1)$; $f \in \mathcal{R}(J_2)$ follows completely analogous. The general case now follows by induction on M. If, $f \in \mathcal{R}(I, \mathbb{C})$, then one computes, using the real-valued case,

$$\int_I f = \left(\int_I \operatorname{Re} f, \int_I \operatorname{Im} f \right) = \left(\sum_{k=1}^M \int_{J_k} \operatorname{Re} f, \sum_{k=1}^M \int_{J_k} \operatorname{Im} f \right) = \sum_{k=1}^M \int_{J_k} f.$$

(c): If $f, g : I \longrightarrow \mathbb{R}$ are bounded and $f \leq g$, then, for each partition Δ of I, $r(\Delta, f) \leq r(\Delta, g)$ and $R(\Delta, f) \leq R(\Delta, g)$ are immediate from (10.7). As these inequalities are preserved when taking the sup and the inf, respectively, all claims of (c) are established.

(d): We will see in Th. 10.18(c) below, that $f \in \mathcal{R}(I, \mathbb{K})$ implies $|f| \in \mathcal{R}(I)$. Let Δ be an arbitrary partition of I, tagged by $(t_1, \ldots, t_N)$. Then, using the same notation as in Def. 10.3 and Def. 10.4,

$$\left| \left(\rho(\Delta, \mathrm{Re}\, f),\, \rho(\Delta, \mathrm{Im}\, f) \right) \right| := \left| \left(\sum_{j=1}^{N} \mathrm{Re}\, f(t_j)\, |I_j|,\ \sum_{j=1}^{N} \mathrm{Im}\, f(t_j)\, |I_j| \right) \right|$$

$$\leq \sum_{j=1}^{N} \left| \left(\mathrm{Re}\, f(t_j),\ \mathrm{Im}\, f(t_j) \right) \right| |I_j|$$

$$= \sum_{j=1}^{N} |f(t_j)|\, |I_j| =: \rho(\Delta, |f|). \tag{10.34}$$

Since the intermediate Riemann sums in (10.34) converge to the respective integrals by (10.25b), one obtains

$$\left| \int_I f \right| = \lim_{|\Delta| \to 0} \left| \left(\rho(\Delta, \mathrm{Re}\, f),\, \rho(\Delta, \mathrm{Im}\, f) \right) \right| \overset{(10.34)}{\leq} \lim_{|\Delta| \to 0} \rho(\Delta, |f|) = \int_I |f|,$$

proving (10.31). ∎

Theorem 10.12 (Mean Value Theorem). *Let $a, b \in \mathbb{R}$, $a \leq b$, $I := [a, b]$. If $f, p \in \mathcal{R}(I)$ and $p \geq 0$, then, for each $m, M \in \mathbb{R}$ with $m \leq f \leq M$:*

$$m \int_I p \leq \int_I fp \leq M \int_I p. \tag{10.35a}$$

In particular, if f is continuous, then

$$\underset{\xi \in I}{\exists}\ \int_I fp = f(\xi) \int_I p. \tag{10.35b}$$

Returning to a general $f \in \mathcal{R}(I)$, if $p \equiv 1$, then we obtain the theorem's classical form:

$$m\,(b - a) = m\,|I| \leq \int_a^b f = \int_I f \leq M\,|I| = M\,(b - a). \tag{10.35c}$$

The theorem's name comes from the fact that, for $a < b$, $|I|^{-1} \int_I f$ is sometimes referred to as the mean value *of f on I.*

Proof. Since $mp \leq fp \leq Mp$, we compute

$$m \int_I p \overset{\text{Th. 10.11(c)}}{\leq} \int_I fp \overset{\text{Th. 10.11(c)}}{\leq} M \int_I p. \tag{10.36}$$

If $\int_I p = 0$, then (10.35a) implies $\int_I fp = 0$ and (10.35b) is immediate. It remains to consider that $\int_I p > 0$. If f is continuous, then we can let m, M be such that $f(I) = [m, M]$ by Th. 7.54 and the intermediate value Th. 7.57. Then (10.35a) shows there is $\xi \in I$ such that $f(\xi) = \frac{\int_I fp}{\int_I p}$, i.e. (10.35b) holds. ∎

Theorem 10.13 (Riemann's Integrability Criterion)**.** *Let* $I = [a,b] \subseteq \mathbb{R}$ *and suppose* $f : I \longrightarrow \mathbb{R}$ *is bounded. Then* f *is Riemann integrable over* I *if, and only if, for each* $\epsilon > 0$, *there exists a partition* Δ *of* I *such that*

$$R(\Delta, f) - r(\Delta, f) < \epsilon. \tag{10.37}$$

Proof. Suppose, for each $\epsilon > 0$, there exists a partition Δ of I such that (10.37) is satisfied. Then

$$J^*(f, I) - J_*(f, I) \leq R(\Delta, f) - r(\Delta, f) < \epsilon, \tag{10.38}$$

showing $J^*(f, I) \leq J_*(f, I)$. As the opposite inequality always holds, we have $J^*(f, I) = J_*(f, I)$, i.e. $f \in \mathcal{R}(I)$ as claimed. Conversely, if $f \in \mathcal{R}(I)$ and $(\Delta_n)_{n \in \mathbb{N}}$ is a sequence of partitions of I with $\lim_{n \to \infty} |\Delta_n| = 0$, then (10.25a) implies that, for each $\epsilon > 0$, there is $N \in \mathbb{N}$ such that $R(\Delta_n, f) - r(\Delta_n, f) < \epsilon$ for each $n > N$. $\blacksquare$

The previous theorem will allow us to prove that every continuous function on $[a, b]$ is Riemann integrable. However, we will also need to make use of the following result:

Proposition 10.14. *Let* $I = [a, b] \subseteq \mathbb{R}$, $a \leq b$, $f : I \longrightarrow \mathbb{R}$. *If* f *is continuous, then* f *is even* uniformly continuous, *i.e.*

$$\underset{\epsilon \in \mathbb{R}^+}{\forall} \; \underset{\delta \in \mathbb{R}^+}{\exists} \; \underset{x,y \in I}{\forall} \; \left(|x - y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \epsilon \right). \tag{10.39}$$

Proof. Arguing by contraposition, we assume f not to be uniformly continuous on I. Then the negation of (10.39) must hold, i.e.

$$\underset{\epsilon_0 \in \mathbb{R}^+}{\exists} \; \underset{\delta \in \mathbb{R}^+}{\forall} \; \underset{x,y \in I}{\exists} \; \left(|x - y| < \delta \quad \wedge \quad |f(x) - f(y)| \geq \epsilon_0 \right). \tag{10.40}$$

In particular, for each $n \in \mathbb{N}$, there exist $x_n, y_n \in I$ such that

$$|x_n - y_n| < \delta_n := 1/n \tag{10.41}$$

and $|f(x_n) - f(y_n)| \geq \epsilon_0$. Then the sequence $(x_n)_{n \in \mathbb{N}}$ is bounded and the Bolzano-Weierstrass Th. 7.27 provides a convergent subsequence $(x_{\phi(n)})_{n \in \mathbb{N}}$, i.e. there is $\xi \in \mathbb{R}$ with $\lim_{n \to \infty} x_{\phi(n)} = \xi$. Clearly, $\xi \in [a, b]$ and (10.41) implies $\lim_{n \to \infty} y_{\phi(n)} = \xi$ as well. However, due to $|f(x_{\phi(n)}) - f(y_{\phi(n)})| \geq \epsilon_0 > 0$, the sequences $\left(f(x_{\phi(n)}) \right)_{n \in \mathbb{N}}$ and $\left(f(y_{\phi(n)}) \right)_{n \in \mathbb{N}}$ can not both converge to $f(\xi)$, showing that f can not be continuous. $\blacksquare$

Caveat 10.15. It is important in Prop. 10.14 that f is defined on a *compact* interval I. The examples $f :]0, 1] \longrightarrow \mathbb{R}$, $f(x) := 1/x$, and $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) := x^2$ are examples of continuous functions that are *not* uniformly continuous.

Theorem 10.16. *Let* $I = [a, b] \subseteq \mathbb{R}$, $a \leq b$.

(a) *If* $f : I \longrightarrow \mathbb{C}$ *is continuous, then* f *is Riemann integrable over* I.

(b) *If* $f : I \longrightarrow \mathbb{R}$ *is increasing or decreasing, then* f *is Riemann integrable over* I.

Proof. (a): As f is continuous if, and only if, $\operatorname{Re} f$ and $\operatorname{Im} f$ are both continuous, it suffices to consider a real-valued continuous f. For $a = b$, there is nothing to prove, so let $a < b$. First note that, if f is continuous on $I = [a, b]$, then f is bounded by Th. 7.54. Moreover, f is uniformly continuous due to Prop. 10.14. Thus, given $\epsilon > 0$, there is $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon/|I|$ for each $x, y \in I$. Then, for each partition Δ of I satisfying $|\Delta| < \delta$, we obtain

$$R(\Delta, f) - r(\Delta, f) = \sum_{j=1}^{N}(M_j - m_j)|I_j| \leq \frac{\epsilon}{|I|}\sum_{j=1}^{N}|I_j| = \epsilon, \qquad (10.42)$$

as $|\Delta| < \delta$ implies $|x - y| < \delta$ for each $x, y \in I_j$ and each $j \in \{1, \ldots, N\}$. Finally, (10.42) implies $f \in \mathcal{R}(I)$ due to Riemann's integrability criterion of Th. 10.13.

(b): Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is increasing. Then f is bounded, as $f(a) \leq f(x) \leq f(b)$ for each $x \in [a, b]$. If $f(a) = f(b)$, then f is constant. Thus, assume $f(a) < f(b)$. Moreover, if $\Delta = (x_0, \ldots, x_N)$ is a partition of I as in Def. 10.3, then

$$R(\Delta, f) - r(\Delta, f) = \sum_{j=1}^{N}(M_j - m_j)|I_j| = \sum_{j=1}^{N}\big(f(x_j) - f(x_{j-1})\big)|I_j| \leq |\Delta|\big(f(b) - f(a)\big).$$
$$(10.43)$$

Thus, given $\epsilon > 0$, we have $R(\Delta, f) - r(\Delta, f) < \epsilon$ for each partition Δ of I satisfying $|\Delta| < \epsilon/(f(b) - f(a))$. In consequence, $f \in \mathcal{R}(I)$, once again due to Riemann's integrability criterion of Th. 10.13. If f is decreasing, then $-f$ is increasing, and Th. 10.11(a) establishes the case. ∎

Definition and Remark 10.17. Let $M \subseteq \mathbb{C}$. A function $f : M \longrightarrow \mathbb{C}$ f is called *Lipschitz continuous* in M with *Lipschitz constant L* if, and only if,

$$\underset{L \in \mathbb{R}_0^+}{\exists} \quad \underset{x,y \in M}{\forall} \quad |f(x) - f(y)| \leq L\,|x - y|. \qquad (10.44)$$

Every Lipschitz continuous function is, indeed, continuous, since, if $\xi \in M$ and $(y_n)_{n \in \mathbb{N}}$ is a sequence in M with $\lim_{n \to \infty} y_n = \xi$, then (10.44) implies

$$\underset{n \in \mathbb{N}}{\forall} \quad |f(\xi) - f(y_n)| \leq L\,|\xi - y_n|, \qquad (10.45)$$

proving $\lim_{n \to \infty} f(y_n) = f(\xi)$. Moreover, it is not too much harder to prove Lipschitz continuous functions are even uniformly continuous, but we will not pursue this right now. On the other hand, $f : \mathbb{R}_0^+ \longrightarrow \mathbb{R}$, $f(x) := \sqrt{x}$, is an example of a continuous function (actually, even uniformly continuous) that is *not* Lipschitz continuous.

Theorem 10.18. *Let $a, b \in \mathbb{R}$, $a \leq b$, $I := [a, b]$.*

(a) *If $f \in \mathcal{R}(I, \mathbb{R})$ and $\phi : f(I) \longrightarrow \mathbb{C}$ is Lipschitz continuous, then $\phi \circ f \in \mathcal{R}(I, \mathbb{C})$.*

(b) *If $f \in \mathcal{R}(I, \mathbb{C})$ and $\phi : f(I) \longrightarrow \mathbb{R}$ is Lipschitz continuous, then $\phi \circ f \in \mathcal{R}(I, \mathbb{R})$.*

(c) *If* $f \in \mathcal{R}(I)$*, then* $|f|, f^2, f^+, f^- \in \mathcal{R}(I)$*. In particular, we, indeed, have* (10.2) *from the introduction (with M replaced by I). If, in addition, there exists* $\delta > 0$ *such that* $f(x) \geq \delta$ *for each* $x \in I$*, then* $1/f \in \mathcal{R}(I)$*. Moreover,* $|f| \in \mathcal{R}(I)$ *also holds for* $f \in \mathcal{R}(I, \mathbb{C})$*.*

(d) *If* $f, g \in \mathcal{R}(I)$*, then* $\max(f, g), \min(f, g) \in \mathcal{R}(I)$*. If* $f, g \in \mathcal{R}(I, \mathbb{K})$*, then* $\bar{f}, fg \in \mathcal{R}(I, \mathbb{K})$*. If, in addition, there exists* $\delta > 0$ *such that* $|g(x)| \geq \delta$ *for each* $x \in I$*, then* $f/g \in \mathcal{R}(I, \mathbb{K})$*.*

Proof. (a),(b): We just carry out the proof for the case $f \in \mathcal{R}(I)$, $\phi : f(I) \longrightarrow \mathbb{R}$, and leave the extension to the case that f or ϕ is $\mathbb{C}$-valued as an exercise. Thus, let $f \in \mathcal{R}(I)$ and let $\phi : f(I) \longrightarrow \mathbb{R}$ be Lipschitz continuous. Then there exists $L \geq 0$ such that

$$\left|\phi(x) - \phi(y)\right| \leq L\,|x - y| \quad \text{for each } x, y \in f(I). \tag{10.46}$$

As $f \in \mathcal{R}(I)$, given $\epsilon > 0$, Th. 10.13 provides a partition Δ of I such that $R(\Delta, f) - r(\Delta, f) < \epsilon/L$, and we obtain

$$
\begin{aligned}
R(\Delta, \phi \circ f) - r(\Delta, \phi \circ f) &= \sum_{j=1}^{N} \left(M_j(\phi \circ f) - m_j(\phi \circ f)\right)|I_j| \\
&\leq \sum_{j=1}^{N} L\left(M_j(f) - m_j(f)\right)|I_j| \\
&= L\left(R(\Delta, f) - r(\Delta, f)\right) < \epsilon.
\end{aligned}
\tag{10.47}
$$

Thus, $\phi \circ f \in \mathcal{R}(I)$ by another application of Th. 10.13.

(c): $|f|, f^2, f^+, f^- \in \mathcal{R}(I)$ follows from (b) (for $|f|$ also for $f \in \mathcal{R}(I, \mathbb{C})$), since each of the maps $x \mapsto |x|$, $x \mapsto x^2$, $x \mapsto \max\{x, 0\}$, $x \mapsto -\min\{x, 0\}$ is Lipschitz continuous on the bounded set $f(I)$ (recall that $f \in \mathcal{R}(I)$ implies that f is bounded). Since $f = f^+ - f^-$, (10.2) is implied by (10.28). Finally, if $f(x) \geq \delta > 0$, then $x \mapsto x^{-1}$ is Lipschitz continuous on the bounded set $f(I)$, and $f^{-1} \in \mathcal{R}(I)$ follows from (b).

(d): For $f, g \in \mathcal{R}(I)$, we note that, due to

$$fg = \frac{1}{4}\left((f + g)^2 - (f - g)^2\right), \tag{10.48a}$$

$$\max(f, g) = f + (g - f)^+, \tag{10.48b}$$

$$\min(f, g) = g - (f - g)^-, \tag{10.48c}$$

everything is a consequence of (c). For $f, g \in \mathcal{R}(I, \mathbb{C})$, due to

$$\bar{f} = (\operatorname{Re} f, -\operatorname{Im} f), \tag{10.48d}$$

$$fg = (\operatorname{Re} f \operatorname{Re} g - \operatorname{Im} f \operatorname{Im} g,\ \operatorname{Re} f \operatorname{Im} g + \operatorname{Im} f \operatorname{Re} g), \tag{10.48e}$$

$$1/g = (\operatorname{Re} g/|g|^2,\ -\operatorname{Im} g/|g|^2), \tag{10.48f}$$

everything follows from the real-valued case together with (c) and Th. 10.11(a), where $|g| \geq \delta > 0$ guarantees $|g|^2 \geq \delta^2 > 0$). $\blacksquare$

10.2   Important Theorems

This section compiles a number of important theorems on Riemann integrals, which, in particular, provide powerful tools to actually evaluate such integrals.

10.2.1   Fundamental Theorem of Calculus

We provide two variants of the fundamental theorem with slightly different flavors: In the first variant, Th. 10.20(a), we start with a function f, obtain another function F by means of integrating f, and recover f by taking the derivative of F. In the second variant, Th. 10.20(b), one first differentiates the given function F, obtaining $f := F'$, followed by integrating f, recovering F *up to an additive constant.*

Notation 10.19. If $a, b \in \mathbb{R}$, $a \leq b$, $I := [a, b]$, $f : I \longrightarrow \mathbb{R}$, then denote

$$\int_a^b f := \int_I f, \qquad\qquad \int_b^a f := -\int_a^b f, \qquad\qquad (10.49\text{a})$$

$$[f(t)]_a^b := [f]_a^b := f(b) - f(a), \qquad [f(t)]_b^a := [f]_b^a := f(a) - f(b), \qquad (10.49\text{b})$$

where $f \in \mathcal{R}(I)$ for (10.49a).

Theorem 10.20. *Let* $a, b \in \mathbb{R}$, $a < b$, $I := [a, b]$.

(a) *If* $f \in \mathcal{R}(I, \mathbb{K})$ *is continuous in* $\xi \in I$, *then, for each* $c \in I$, *the function*

$$F_c : I \longrightarrow \mathbb{K}, \quad F_c(x) := \int_c^x f(t)\,\mathrm{d}t\,, \qquad\qquad (10.50)$$

is differentiable in ξ *with* $F'(\xi) = f(\xi)$. *In particular, if* $f \in C(I, \mathbb{K})$, *then* $F \in C^1(I, \mathbb{K})$ *and* $F'(x) = f(x)$ *for each* $x \in I$.

(b) *If* $F \in C^1(I, \mathbb{K})$ *or, alternatively,* $F : I \longrightarrow \mathbb{K}$ *is differentiable with integrable derivative* $F' \in \mathcal{R}(I, \mathbb{K})$, *then*

$$F(b) - F(a) = [F(t)]_a^b = \int_a^b F'(t)\,\mathrm{d}t\,, \qquad\qquad (10.51\text{a})$$

and

$$F(x) = F(c) + \int_c^x F'(t)\,\mathrm{d}t \quad \text{*for each* } c, x \in I. \qquad\qquad (10.51\text{b})$$

Proof. It suffices to prove the case $\mathbb{K} = \mathbb{R}$, since the case $\mathbb{K} = \mathbb{C}$ then follows by applying the case $\mathbb{K} = \mathbb{R}$ to $\operatorname{Re} F_c$ and $\operatorname{Im} F_c$ (for (a)) and to $\operatorname{Re} F$ and $\operatorname{Im} F$ (for (b)). Thus, for the rest of the proof, we assume $\mathbb{K} = \mathbb{R}$.

(a): We need to show that

$$\lim_{h \to 0} A(h) = 0, \quad \text{where} \quad A(h) := \frac{F(\xi + h) - F(\xi)}{h} - f(\xi). \qquad\qquad (10.52)$$

One computes

$$A(h) = \frac{1}{h} \int_\xi^{\xi+h} f(t)\,\mathrm{d}t - \frac{1}{h}f(\xi) \int_\xi^{\xi+h} \mathrm{d}t = \frac{1}{h} \int_\xi^{\xi+h} \big(f(t) - f(\xi)\big)\,\mathrm{d}t. \qquad (10.53)$$

Now, given $\epsilon > 0$, the continuity of f in ξ allows us to find $\delta > 0$ such that $|(f(t) - f(\xi)| < \epsilon/2$ for each t with $|t - \xi| < \delta$. Thus, for each h with $|h| < \delta$, we obtain

$$|A(h)| \le \frac{1}{h} \int_\xi^{\xi+h} \big|f(t) - f(\xi)\big|\,\mathrm{d}t \le \frac{\epsilon h}{2h} < \epsilon, \qquad (10.54)$$

thereby proving $\lim_{h\to 0} A(h) = 0$, i.e. $f(\xi) = F'(\xi)$.

(b): First assume $F \in C^1(I)$. Then F' is continuous on I, and we can apply (a) to the function

$$G : I \longrightarrow \mathbb{R}, \quad G(x) := \int_a^x F'(t)\,\mathrm{d}t, \qquad (10.55)$$

to obtain $G' = F'$. Thus, for $H := F - G$, we obtain $H' \equiv 0$, showing that H must be constant on I, i.e. $H(x) = H(a) = F(a) - G(a) = F(a)$ for each $x \in I$. Evaluating at $x = b$ yields

$$F(a) = H(b) = F(b) - \int_a^b F'(t)\,\mathrm{d}t, \qquad (10.56)$$

thereby establishing the case.

Now we consider the remaining case of a differentiable F with integrable derivative $F' \in \mathcal{R}(I)$. Consider a partition $\Delta = (x_0, \ldots, x_N)$ of I as in Def. 10.3. Then, for each $j \in \{1, \ldots, N\}$, the mean value theorem provides $\xi_j \in]x_{j-1}, x_j[$ such that $F(x_j) - F(x_{j-1}) = (x_j - x_{j-1})\,F'(\xi_j)$. Thus,

$$F(b) - F(a) = \sum_{j=1}^N \big(F(x_j) - F(x_{j-1})\big) = \sum_{j=1}^N (x_j - x_{j-1})\,F'(\xi_j) = \rho(\Delta, F'). \qquad (10.57)$$

If we choose a sequence of partitions Δ of I such that $|\Delta| \to 0$, then the integrability of f implies that the right-hand side of (10.57) converges to $\int_a^b F'$, once again establishing the case. ∎

Definition 10.21. If $I \subseteq \mathbb{R}$, $f : I \longrightarrow \mathbb{K}$, and $F : I \longrightarrow \mathbb{K}$ is a differentiable function with $F' = f$, then F is called a *primitive* or *antiderivative* of f.

Example 10.22. (a) Due to the fundamental theorem, if we know a function's antiderivative, we can easily compute its integral over a given interval. Here are three simple examples:

$$\int_0^1 (x^5 - 3x)\,\mathrm{d}x = \left[\frac{x^6}{6} - \frac{3x^2}{2}\right]_0^1 = \frac{1}{6} - \frac{3}{2} = -\frac{4}{3}, \qquad (10.58a)$$

$$\int_1^e \frac{1}{x}\,\mathrm{d}x = [\ln x]_1^e = \ln e - \ln 1 = 1, \qquad (10.58b)$$

$$\int_0^\pi \sin x\,\mathrm{d}x = [-\cos x]_0^\pi = 2. \qquad (10.58c)$$

(b) As a more complicated example, consider a rational function $x \mapsto r(x) = p(x)/q(x)$ with polynomials p and q. To find an antiderivative of r, one first writes r in the form $r = s + \tilde{p}/\tilde{q}$ with polynomials $s, \tilde{p}, \tilde{q}$ such that $\deg(\tilde{p}) < \deg(\tilde{q})$ (this can be done using so-called *polynomial long division*). One then applies the *partial fraction decomposition* of Sec. G of the Appendix to $\tilde{p}/\tilde{q}$: As a concrete example, consider

$$r : \mathbb{R} \setminus \{-1, 2\} \longrightarrow \mathbb{R}, \quad r(x) := \frac{3x^8 - 9x^6 - 6x^5 + x^3 + 5x^2 - 1}{x^3 - 3x - 2}.$$

One then finds

$$r(x) = 3x^5 + 1 + \frac{5x^2 + 3x + 1}{(x+1)^2(x-2)},$$

and the partial fraction decomposition according to (G.5) is

$$r(x) = 3x^5 + 1 + \frac{2}{x+1} - \frac{1}{(x+1)^2} + \frac{3}{x-2}.$$

One can now easily provide an antiderivative for each summand, and putting everything together yields the antiderivative

$$R : \mathbb{R} \setminus \{-1, 2\} \longrightarrow \mathbb{R}, \quad R(x) := \frac{1}{2}x^6 + x + 2\ln|x+1| + \frac{1}{x+1} + 3\ln|x-2|,$$

for r. We note that, to apply partial fraction decomposition, one always needs the zeros of the denominator (i.e. of $\tilde{q}$). For polynomials of high degree, it will usually not be possible to determine theses zeros exactly, but merely approximitely.

10.2.2   Integration by Parts Formula

Theorem 10.23. *Let $a, b \in \mathbb{R}$, $a < b$, $I := [a, b]$. If $f, g \in C^1(I, \mathbb{K})$, then the following integration by parts formula holds:*

$$\int_a^b fg' = [fg]_a^b - \int_a^b f'g. \tag{10.59}$$

Proof. If $f, g \in C^1(I, \mathbb{K})$, then, according to the product rule, $fg \in C^1(I, \mathbb{K})$ with $(fg)' = f'g + fg'$. Applying (10.51a), we obtain

$$[fg]_a^b = \int_a^b (fg)' = \int_a^b f'g + \int_a^b fg', \tag{10.60}$$

which is precisely (10.59). ∎

Example 10.24. We compute the integral $\int_0^{2\pi} \sin^2 t \, \mathrm{d}t$:

$$\int_0^{2\pi} \sin^2 t \, \mathrm{d}t = [-\sin t \cos t]_0^{2\pi} + \int_0^{2\pi} \cos^2 t \, \mathrm{d}t = \int_0^{2\pi} \cos^2 t \, \mathrm{d}t. \tag{10.61}$$

Adding $\int_0^{2\pi} \sin^2 t \, \mathrm{d}t$ on both sides of (10.61) and using $\sin^2 + \cos^2 \equiv 1$ yields

$$2 \int_0^{2\pi} \sin^2 t \, \mathrm{d}t = \int_0^{2\pi} 1 \, \mathrm{d}t = 2\pi, \tag{10.62}$$

i.e. $\int_0^{2\pi} \sin^2 t \, \mathrm{d}t = \pi$.

10.2.3 Change of Variables

Theorem 10.25. *Let* $I, J \subseteq \mathbb{R}$ *be intervals,* $\phi \in C^1(I)$ *and* $f \in C(J, \mathbb{K})$. *If* $\phi(I) \subseteq J$, *then the following* change of variables *formula holds for each* $a, b \in I$:

$$\int_{\phi(a)}^{\phi(b)} f = \int_{\phi(a)}^{\phi(b)} f(x) \, dx = \int_a^b f(\phi(t)) \, \phi'(t) \, dt = \int_a^b (f \circ \phi) \, \phi'. \tag{10.63}$$

Proof. We first consider $\mathbb{K} = \mathbb{R}$. Let

$$F : \ J \longrightarrow \mathbb{R}, \quad F(x) := \int_{\phi(a)}^x f(t) \, dt \,. \tag{10.64}$$

According to Th. 10.20(a) and the chain rule of Th. 9.11, we obtain

$$(F \circ \phi)' : \ I \longrightarrow \mathbb{R}, \quad (F \circ \phi)'(x) = \phi'(x) f(\phi(x)). \tag{10.65}$$

Thus, we can apply (10.51a), which yields

$$\int_{\phi(a)}^{\phi(b)} f = F(\phi(b)) - F(\phi(a)) = \int_a^b (f \circ \phi) \, \phi', \tag{10.66}$$

proving (10.63). For the case $\mathbb{K} = \mathbb{C}$, we compute

$$\int_{\phi(a)}^{\phi(b)} f = \left(\int_{\phi(a)}^{\phi(b)} \operatorname{Re} f, \ \int_{\phi(a)}^{\phi(b)} \operatorname{Im} f \right)$$

$$= \left(\int_a^b (\operatorname{Re} f \circ \phi) \, \phi', \ \int_a^b (\operatorname{Im} f \circ \phi) \, \phi' \right) = \int_a^b (f \circ \phi) \, \phi',$$

thereby completing the proof. ∎

Example 10.26. We compute the integral $\int_0^1 t^2 \sqrt{1 - t} \, dt$ using the change of variables $x := \phi(t) := 1 - t$, $\phi'(t) = -1$:

$$\int_0^1 t^2 \sqrt{1 - t} \, dt = - \int_1^0 (1 - x)^2 \sqrt{x} \, dx = \int_0^1 (\sqrt{x} - 2x\sqrt{x} + x^2 \sqrt{x}) \, dx$$

$$= \left[\frac{2x^{\frac{3}{2}}}{3} - \frac{4x^{\frac{5}{2}}}{5} + \frac{2x^{\frac{7}{2}}}{7} \right]_0^1 = \frac{16}{105}. \tag{10.67}$$

10.3 Application: Taylor's Theorem

Theorem 10.27 (Taylor's Theorem). *Let* $I \subseteq \mathbb{R}$ *be an open interval and* $a, x \in I$, $x \neq a$. *If* $m \in \mathbb{N}_0$ *and* $f \in C^{m+1}(I, \mathbb{K})$, *then*

$$f(x) = T_m(x, a) + R_m(x, a), \tag{10.68}$$

where

$$T_m(x,a) := \sum_{k=0}^{m} \frac{f^{(k)}(a)}{k!}(x-a)^k = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(m)}(a)}{m!}(x-a)^m$$

(10.69)

is the mth Taylor polynomial *and*

$$R_m(x,a) := \int_a^x \frac{(x-t)^m}{m!} f^{(m+1)}(t)\,\mathrm{d}t$$

(10.70)

is the integral form *of the* remainder term. *For $\mathbb{K} = \mathbb{R}$, one can also write the remainder term in* Lagrange form:

$$R_m(x,a) = \frac{f^{(m+1)}(\theta)}{(m+1)!}(x-a)^{m+1} \quad \text{with some suitable } \theta \in]x,a[.$$

(10.71)

Proof. The integral form (10.70) of the remainder term we prove by using induction on m: For $m = 0$, the assertion is

$$f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t\,,$$

(10.72)

which holds according to the fundamental theorem of calculus in the form Th. 10.20(b). For the induction step, we assume (10.68) holds for fixed $m \in \mathbb{N}_0$ with $R_m(x,a)$ in integral form (10.70) and consider $f \in C^{m+2}(I, \mathbb{K})$. For fixed $x \in I$, we define the function

$$g : I \longrightarrow \mathbb{K}, \quad g(t) := \frac{(x-t)^{m+1}}{(m+1)!} f^{(m+1)}(t).$$

(10.73)

Using the product rule, its derivative is

$$g' : I \longrightarrow \mathbb{K}, \quad g'(t) = \frac{(x-t)^{m+1}}{(m+1)!} f^{(m+2)}(t) - \frac{(x-t)^m}{m!} f^{(m+1)}(t).$$

(10.74)

Applying the fundamental theorem to g then yields

$$-g(a) = g(x) - g(a) = \int_a^x g'(t)\,\mathrm{d}t \overset{(10.74)}{=} R_{m+1}(x,a) - R_m(x,a),$$

(10.75)

with $R_m(x,a)$ and $R_{m+1}(x,a)$ defined according to (10.70). Thus,

$$T_{m+1}(x,a) + R_{m+1}(x,a) \overset{(10.75)}{=} T_m(x,a) + \frac{f^{(m+1)}(a)}{(m+1)!}(x-a)^{m+1} + R_m(x,a) - g(a)$$

$$= T_m(x,a) + R_m(x,a) \overset{\text{ind. hyp.}}{=} f(x),$$

(10.76)

thereby completing the induction and the proof of (10.70).

It remains to prove the Lagrange form (10.71) of the remainder term for $\mathbb{K} = \mathbb{R}$. Since $(x-t)^m > 0$ for $x > t$ or $m = 0$, and $(x-t)^m < 0$ for $x < t$ and $m > 0$, (10.70) and (10.35b) imply the existence of some $\theta \in [x,a]$ such that

$$R_m(x,a) = f^{(m+1)}(\theta) \int_a^x \frac{(x-t)^m}{m!}\,\mathrm{d}t = \frac{f^{(m+1)}(\theta)}{(m+1)!}(x-a)^{m+1},$$

(10.77)

proving the Lagrange form, where it remains an exercise to show one can always choose $\theta \in]x, a[$.   ∎

Remark 10.28. The importance of Taylor's Th. 10.27 does not lie in the decomposition $f = T_m + R_m$, which can be accomplished simply by defining $R_m := f - T_m$. The importance lies rather in the specific formulas for the remainder term.

Example 10.29. Depending on $f \in C^\infty(I)$ and $x, a \in I$, the *Taylor series*

$$\lim_{m \to \infty} T_m(x, a) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k$$

can converge to $f(x)$ (see (a) below), diverge (see (b) below), or converge to $\eta \neq f(x)$ (see (c) below):

(a) For $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = e^x$, $x \in \mathbb{R}$, $a = 0$, we have $f^{(k)}(a) = e^0 = 1$ for each $k \in \mathbb{N}_0$, recovering the power series for the exponential function, which we already know to converge from Def. and Rem. 8.14:

$$\lim_{m \to \infty} T_m(x, 0) = \lim_{m \to \infty} \sum_{k=0}^{m} \frac{1}{k!} x^k = e^x.$$

(b) For $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $f(x) = \ln x$, we have $f^{(m)}(x) = (-1)^{m-1}(m-1)!\, x^{-m}$ for each $m \in \mathbb{N}_0$ and each $x \in \mathbb{R}^+$. Thus, for $x = \frac{3}{2}$ and $a = \frac{1}{2}$, using $f^{(k)}(a) = (-1)^{k-1}(k-1)!\, 2^k$, we have

$$T_m(x, a) = \sum_{k=0}^{m} \frac{f^{(k)}(a)}{k!} (x - a)^k = \sum_{k=0}^{m} \frac{(-1)^{k-1}\, 2^k}{k},$$

which diverges for $m \to \infty$, since the summands do not converge to 0.

(c) For

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x) := \begin{cases} e^{-1/x^2} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases}$$

it is an exercise to show $f \in C^\infty(\mathbb{R})$ with $f^k(0) = 0$ for each $k \in \mathbb{N}_0$ (hint: for $x \neq 0$, one obtains $f^k(x) = P_k(x^{-1})e^{-1/x^2}$, where P_k is a polynomial of degree $3k$). Thus, $T_m(x, 0) = 0$ for each $x \in \mathbb{R}$ and each $m \in \mathbb{N}_0$, implying

$$\lim_{m \to \infty} T_m(x, 0) = 0 \neq f(x) \quad \text{for each } x \in \mathbb{R} \setminus \{0\}.$$

10.4   Improper Integrals

For our definition of the Riemann integral in Def. 10.5, it was important that we considered *bounded* functions on *compact* intervals (where the boundedness of the intervals

was more important than the closedness) – for unbounded functions and/or unbounded intervals, even Def. 10.4 of lower and upper Riemann sums no longer makes sense.

Still, for sufficiently benign functions, it is possible to extend the notion of a definite Riemann integral to both unbounded intervals and unbounded functions, and in such situations we will speak of *improper* integrals (cf. Def. 10.33 below).

Definition 10.30. Let $\emptyset \neq I \subseteq \mathbb{R}$ be an interval. We call $f : \mathbb{R} \longrightarrow \mathbb{R}$ to be *locally Riemmann integrable* if, and only if, $f \in \mathcal{R}(J)$ for each compact interval $J \subseteq I$. Let $\mathcal{R}_{\mathrm{loc}}(I)$ denote the set of all locally Riemmann integrable functions on I.

Remark 10.31. In particular, locally Riemann integrable functions are *bounded* on every compact interval. Moreover, $\mathcal{R}_{\mathrm{loc}}(I) = \mathcal{R}(I)$ if, and only if, I is a compact interval. For example, for each $a, b \in \mathbb{R}$ with $a < b$, the function given by the assignment rule

$$f(x) := \frac{1}{(x-a)(x-b)}$$

is clearly locally Riemmann integrable, but not bounded on each of the intervals $]-\infty, a[$, $]a, b[$, and $]b, \infty[$.

———

Before we can define improper Riemann integral, we define, in partial extension of Def. 8.17:

Definition 10.32. Let $M \subseteq \mathbb{R}$. If M is unbounded from above (resp. below, then $f : M \longrightarrow \mathbb{K}$ is said to tend to $\eta \in \mathbb{K}$ (or to have the *limit* $\eta \in \mathbb{K}$) for $x \to \infty$ (resp., for $x \to -\infty$) (denoted by $\lim_{x \to \pm\infty} f(x) = \eta$) if, and only if, for each sequence $(\xi_k)_{k \in \mathbb{N}}$ in M with $\lim_{k \to \infty} \xi_k = \infty$ (resp. with $\lim_{k \to \infty} \xi_k = -\infty$), the sequence $(f(\xi_k))_{k \in \mathbb{N}}$ converges to $\eta \in \mathbb{K}$, i.e.

$$\lim_{x \to \pm\infty} f(x) = \eta \quad \Leftrightarrow \quad \underset{(\xi_k)_{k \in \mathbb{N}} \text{ in } M}{\forall} \left(\lim_{k \to \infty} \xi_k = \pm\infty \quad \Rightarrow \quad \lim_{k \to \infty} f(\xi_k) = \eta \right). \quad (10.78)$$

Definition 10.33. Let $a < c < b$ ($a = -\infty$, $b = \infty$ is admissible).

(a) Let $I := [c, b[$, $f \in \mathcal{R}_{\mathrm{loc}}(I)$, and assume $b = \infty$ or f is unbounded. Consider the function

$$F : I \longrightarrow \mathbb{R}, \quad F(x) := \int_c^x f.$$

If the limit

$$\lim_{x \to b} F(x) = \lim_{x \to b} \int_c^x f \tag{10.79}$$

exists in $\mathbb{R}$, then we define

$$\int_I f := \int_c^b f(t) \, \mathrm{d}t := \int_c^b f := \lim_{x \to b} \int_c^x f.$$

(b) Let $I :=]a, c]$, $f \in \mathcal{R}_{\mathrm{loc}}(I)$, and assume $a = -\infty$ or f is unbounded. Consider the function

$$F : I \longrightarrow \mathbb{R}, \quad F(x) := \int_x^c f.$$

If the limit

$$\lim_{x \to a} F(x) = \lim_{x \to a} \int_x^c f \tag{10.80}$$

exists in $\mathbb{R}$, then we define

$$\int_I f := \int_a^c f(t)\,\mathrm{d}t := \int_a^c f := \lim_{x \to a} \int_x^c f.$$

(c) Let $I =]a, b[$, $f \in \mathcal{R}_{\mathrm{loc}}(I)$. If the conditions of both (a) and (b) hold, i.e. (i) – (iv), where

(i) $b = \infty$ or f is unbounded on $[c, b[$,

(ii) $\lim_{x \to b} \int_c^x f$ exists in $\mathbb{R}$,

(iii) $a = -\infty$ or f is unbounded on $]a, c]$,

(iv) $\lim_{x \to a} \int_x^c f$ exists in $\mathbb{R}$,

then we define

$$\int_I f := \int_a^b f(t)\,\mathrm{d}t := \int_a^b f := \int_a^c f + \int_c^b f.$$

All the above limits of Riemann integrals (if they exist) are called *improper* Riemann integrals. In each case, if the limit exists, we call f *improperly Riemann integrable* and write $f \in \mathcal{R}(I)$.

Remark 10.34. (a) The definitions in Def. 10.33 are consistent with what occurs if the limits are *proper* Riemann integrals: Let $a, c, b \in \mathbb{R}$, $a < c < b$, and $f \in \mathcal{R}[a, b]$. Then

$$\lim_{x \to b} \int_c^x f = \int_c^b f \quad \text{and} \quad \lim_{x \to a} \int_x^c f = \int_a^c f. \tag{10.81}$$

Indeed, since $f \in \mathcal{R}[a, b]$, $|f|$ is bounded by some $M \in \mathbb{R}^+$; and if $(x_k)_{k \in \mathbb{N}}$ is a sequence in $[a, b[$ such that $\lim_{k \to \infty} x_k = b$, then

$$\left| \int_{x_k}^b f \right| \leq \int_{x_k}^b |f| \leq M\,(b - x_k) \to 0 \quad \text{for} \quad k \to \infty,$$

implying

$$\lim_{k \to \infty} \int_c^{x_k} f \overset{\text{Th. 10.11(b)}}{=} \lim_{k \to \infty} \left(\int_c^b f - \int_{x_k}^b f \right) = \int_c^b f - 0 = \int_c^b f.$$

An analogous argument shows the remaining equality in (10.81).

(b) In Def. 10.33(c), it can occur that $\int_{-\infty}^{\infty} f$ does *not* exist, even though the limit $\lim_{x \to \infty} \int_{-x}^{x} f$ exists: For example, if $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x$, then $f \in \mathcal{R}_{\mathrm{loc}}(\mathbb{R})$, and, for each sequence $(x_k)_{k \in \mathbb{N}}$ in $\mathbb{R}$ such that $\lim_{k \to \infty} x_k = \infty$ and each $c \in \mathbb{R}$, one has

$$\lim_{k \to \infty} \int_{-x_k}^{x_k} t \, dt = \lim_{k \to \infty} \left[\frac{t^2}{2} \right]_{-x_k}^{x_k} = \lim_{k \to \infty} \frac{x_k^2 - x_k^2}{2} = 0,$$

$$\lim_{k \to \infty} \int_{c}^{x_k} t \, dt = \lim_{k \to \infty} \left[\frac{t^2}{2} \right]_{c}^{x_k} = \lim_{k \to \infty} \frac{x_k^2 - c^2}{2} = \infty,$$

$$\lim_{k \to \infty} \int_{-x_k}^{c} t \, dt = \lim_{k \to \infty} \left[\frac{t^2}{2} \right]_{-x_k}^{c} = \lim_{k \to \infty} \frac{c^2 - x_k^2}{2} = -\infty,$$

i.e. $\lim_{x \to \infty} \int_{-x}^{x} t \, dt = 0$, but neither $\lim_{x \to \infty} \int_{c}^{x} t \, dt$ nor $\lim_{x \to -\infty} \int_{x}^{c} t \, dt$ exists in $\mathbb{R}$.

(c) Let $a < c_1 < c_2 < b$ ($a = -\infty$, $b = \infty$ is admissible). If $I := [c_1, b[$, $f \in \mathcal{R}_{\mathrm{loc}}(I)$, and $b = \infty$ or f is unbounded, then $\int_{c_1}^{b} f$ exists if, and only if, $\int_{c_2}^{b} f$ exists. Moreover, if the integrals exist, then

$$\int_{c_1}^{b} f = \int_{c_1}^{c_2} f + \int_{c_2}^{b} f. \tag{10.82a}$$

Indeed, if $(x_k)_{k \to \infty}$ is a sequence in $[c_1, b[$ such that $\lim_{k \to \infty} x_k = b$ and if $\int_{c_1}^{b} f$ exists, then

$$\lim_{k \to \infty} \int_{c_2}^{x_k} f \overset{\text{Th. 10.11(b)}}{=} \lim_{k \to \infty} \left(\int_{c_1}^{x_k} f - \int_{c_1}^{c_2} f \right) = \int_{c_1}^{b} f - \int_{c_1}^{c_2} f,$$

proving $\int_{c_2}^{b} f$ exists and (10.82a) holds. Conversely, if $\int_{c_2}^{b} f$ exists, then

$$\lim_{k \to \infty} \int_{c_1}^{x_k} f \overset{\text{Th. 10.11(b)}}{=} \lim_{k \to \infty} \left(\int_{c_2}^{x_k} f + \int_{c_1}^{c_2} f \right) = \int_{c_2}^{b} f + \int_{c_1}^{c_2} f,$$

proving $\int_{c_1}^{b} f$ exists and (10.82a) holds. Analogously, one shows that if $I :=]a, c_2]$, $f \in \mathcal{R}_{\mathrm{loc}}(I)$, and $a = -\infty$ or f is unbounded, then $\int_{a}^{c_1} f$ exists if, and only if, $\int_{a}^{c_2} f$ exists, where, if the integrals exist, then

$$\int_{a}^{c_2} f = \int_{c_1}^{c_2} f + \int_{a}^{c_1} f. \tag{10.82b}$$

In particular, we see that neither the existence nor the value of the improper integral in Def. 10.33(c) depends on the choice of c.

Example 10.35. **(a)** Let $0 < \alpha < 1$. We claim that

$$\int_{0}^{1} \frac{1}{t^\alpha} \, dt = \frac{1}{1 - \alpha} \quad \left(\alpha = \frac{1}{2} \text{ yields } \int_{0}^{1} \frac{1}{\sqrt{t}} \, dt = 2 \right). \tag{10.83}$$

Indeed, if $(x_k)_{k \in \mathbb{N}}$ is a sequence in $]0, 1]$ such that $\lim_{k \to \infty} x_k = 0$, then

$$\lim_{k \to \infty} \int_{x_k}^{1} \frac{1}{t^\alpha} \, dt = \lim_{k \to \infty} \left[\frac{t^{1-\alpha}}{1 - \alpha} \right]_{x_k}^{1} = \lim_{k \to \infty} \frac{1 - x_k^{1-\alpha}}{1 - \alpha} = \frac{1}{1 - \alpha}.$$

(b) If $(x_k)_{k\in\mathbb{N}}$ is a sequence in $]0,1]$ such that $\lim_{k\to\infty} x_k = 0$, then

$$\lim_{k\to\infty} \int_{x_k}^1 \frac{1}{t}\,\mathrm{d}t = \lim_{k\to\infty} \Big[\ln t\Big]_{x_k}^1 = \lim_{k\to\infty}\Big(0 - \ln x_k\Big) = \infty,$$

showing the limit does not exist in $\mathbb{R}$, but diverges to ∞. Sometimes, this is stated in the form

$$\int_0^1 \frac{1}{t}\,\mathrm{d}t = \infty. \tag{10.84}$$

(c) We claim that

$$\int_{-\infty}^0 e^t\,\mathrm{d}t = 1. \tag{10.85}$$

Indeed, if $(x_k)_{k\in\mathbb{N}}$ is a sequence in $\mathbb{R}_0^-$ such that $\lim_{k\to\infty} x_k = -\infty$, then

$$\lim_{k\to\infty} \int_{x_k}^0 e^t\,\mathrm{d}t = \lim_{k\to\infty}\Big[e^t\Big]_{x_k}^0 = \lim_{k\to\infty}(1 - e^{x_k}) = 1.$$

(d) Consider the function

$$f : \mathbb{R}_0^+ \longrightarrow \mathbb{R}, \quad f(t) := \begin{cases} n & \text{for } n \leq t \leq n + \frac{1}{n2^n},\ n\in\mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\lim_{t\to\infty} f(t)$ does not exist and f is not even bounded. However $f \in \mathcal{R}(\mathbb{R}_0^+)$ and

$$\int_0^\infty f = \sum_{n=1}^\infty \int_n^{n+1/(n2^n)} n\,\mathrm{d}t = \sum_{n=1}^\infty 2^{-n} = \frac{1}{1-\frac{1}{2}} - 1 = 1.$$

Lemma 10.36. *Let $a < c < b$ ($a = -\infty$, $b = \infty$ is admissible). Let $I \subseteq]a,b[$ be one of the three kinds of intervals occurring in Def. 10.33 (i.e. $I = [c,b[$, $I =]a,c]$, or $I =]a,b[$), and assume $f,g : I \longrightarrow \mathbb{R}$ to be improperly Riemann integrable over I.*

(a) Linearity: *For each $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g$ is improperly Riemann integrable over I and*

$$\int_I (\lambda f + \mu g) = \lambda \int_I f + \mu \int_I g.$$

(b) Monotonicity: *If $f \leq g$, then*

$$\int_I f \leq \int_I g.$$

Proof. We conduct the proof for the case $I = [c,b[$ – the case $I =]a,b]$ can be shown analogously, and the case $I =]a,b[$ then also follows. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in I such that $\lim_{k\to\infty} x_k = b$.

(a): One computes

$$\lim_{k\to\infty} \int_c^{x_k} (\lambda f + \mu g) \overset{\text{Th. 10.11(a)}}{=} \lim_{k\to\infty} \left(\lambda \int_c^{x_k} f + \mu \int_c^{x_k} g \right) = \lambda \int_c^b f + \mu \int_c^b g,$$

showing $(\lambda f + \mu g) \in \mathcal{R}(I)$ and proving (a).

(b): One estimates

$$\int_c^b f = \lim_{k\to\infty} \int_c^{x_k} f \overset{\text{Th. 10.11(c)}}{\leq} \lim_{k\to\infty} \int_c^{x_k} g = \int_c^b g,$$

proving (b).   ∎

Definition 10.37. Let $a < c < b$ ($a = -\infty$, $b = \infty$ is admissible). Let $I \subseteq]a,b[$ be one of the three kinds of intervals occurring in Def. 10.33 (i.e. $I = [c,b[$, $I =]a,c]$, or $I =]a,b[$), and assume $f \in \mathcal{R}_{\text{loc}}(I)$. Then, by Th. 10.18(c), $|f| \in \mathcal{R}_{\text{loc}}(I)$. If $\int_I |f|$ exists as an improper integral, then we call the improper integral $\int_I f$ *absolutely convergent*.

—

Before we can proceed to Prop. 10.39 about convergence criteria for improper integrals, we need to prove the analogon of Th. 7.19 for limits of functions.

Proposition 10.38. *Let $\emptyset \neq M \subseteq \mathbb{R}$, $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$, and assume*

$$a = \begin{cases} \inf(M \setminus \{a\}) & \text{if } M \text{ is bounded from below,} \\ -\infty & \text{if } M \text{ is unbounded from below,} \end{cases} \tag{10.86a}$$

$$b = \begin{cases} \sup(M \setminus \{a\}) & \text{if } M \text{ is bounded from above,} \\ \infty & \text{if } M \text{ is unbounded from above.} \end{cases} \tag{10.86b}$$

Let $f : M \longrightarrow \mathbb{R}$ be monotone (increasing or decreasing). Defining $A := f(M) = \{f(x) : x \in M\}$, the following holds:

$$\lim_{x\to b} f(x) = \begin{cases} \sup A & \text{if } f \text{ is increasing and } A \text{ is bounded from above,} \\ \infty & \text{if } f \text{ is increasing and } A \text{ is not bounded from above,} \\ \inf A & \text{if } f \text{ is decreasing and } A \text{ is bounded from below,} \\ -\infty & \text{if } f \text{ is decreasing and } A \text{ is not bounded from below,} \end{cases} \tag{10.87a}$$

$$\lim_{x\to a} f(x) = \begin{cases} \sup A & \text{if } f \text{ is decreasing and } A \text{ is bounded from above,} \\ \infty & \text{if } f \text{ is decreasing and } A \text{ is not bounded from above,} \\ \inf A & \text{if } f \text{ is increasing and } A \text{ is bounded from below,} \\ -\infty & \text{if } f \text{ is increasing and } A \text{ is not bounded from below.} \end{cases} \tag{10.87b}$$

Proof. We prove (10.87a) for the case, where f is increasing – the remaining case of (10.87a) as well as (10.87b) can be proved completely analogous. Let $(x_k)_{k\in\mathbb{N}}$ be a

sequence in $M \setminus \{b\}$ such that $\lim_{k \to \infty} x_k = b$. We have to show that $\lim_{k \to \infty} f(x_k) = \eta$, where $\eta := \sup A$ for A bounded from above and $\eta := \infty$ for A not bounded from above. Seeking a contradiction, assume $\lim_{k \to \infty} f(x_k) = \eta$ does not hold. Due to the choice of b, there then must be $\epsilon > 0$ and a subsequence $(y_k)_{k \in \mathbb{N}}$ of $(x_k)_{k \in \mathbb{N}}$ such that $(y_k)_{k \in \mathbb{N}}$ is strictly increasing and

$$\underset{k \in \mathbb{N}}{\forall} \quad f(y_k) \leq \begin{cases} \eta - \epsilon & \text{if } \eta = \sup A, \\ \epsilon & \text{if } \eta = \infty. \end{cases}$$

Since $\lim_{k \to \infty} y_k = b$ and f is increasing, this means $\sup A \leq \eta - \epsilon$ or $\sup A = \epsilon$, which means a contradiction in each case. Thus, $\lim_{k \to \infty} f(x_k) = \eta$ must hold and the proof is complete. ∎

Proposition 10.39. *Let $a < c < b$ ($a = -\infty$, $b = \infty$ is admissible). Let $I \subseteq]a, b[$ be one of the three kinds of intervals occurring in Def. 10.33 (i.e. $I = [c, b[$, $I =]a, c]$, or $I =]a, b[$), and assume $f \in \mathcal{R}_{\mathrm{loc}}(I)$.*

(a) *If $g \in \mathcal{R}_{\mathrm{loc}}(I)$, $0 \leq f \leq g$, and $\int_I g$ exists, then $\int_I f$ exists as well. Conversely, if $0 \leq g \leq f$ and $\int_I g$ diverges, then $\int_I f$ diverges as well.*

(b) *If $\int_I f$ is an improper integral that is absolutely convergent, then it is also convergent.*

Proof. (a): We consider the case $I = [c, b[$ – the proof for the case $I =]a, c]$ is completely analogous, and the case $I =]a, b[$ then also follows. First, suppose $0 \leq f \leq g$, and $\int_I g$ exists. Since $0 \leq f$, the function

$$F : [c, b[\longrightarrow \mathbb{R}_0^+, \quad F(x) := \int_c^x f,$$

is increasing. Due to

$$\underset{x \in [c,b[}{\forall} \quad F(x) = \int_c^x f \leq \int_c^x g \leq \int_c^b g \in \mathbb{R}_0^+,$$

F is also bounded from above (in the sense that $\{F(x) : x \in [c, b[\}$ is bounded from above), i.e. Prop. 10.38 yields that $\lim_{x \to b} F(x) = \lim_{x \to b} \int_c^x f$ exists in $\mathbb{R}$ as claimed.

Now suppose $0 \leq g \leq f$ and $\int_I g$ diverges. As the function F above, the function

$$G : [c, b[\longrightarrow \mathbb{R}_0^+, \quad G(x) := \int_c^x g,$$

is increasing. Since we assume that $\lim_{x \to b} G(x)$ does not exist in $\mathbb{R}$, Prop. 10.38 implies $\lim_{x \to b} G(x) = \infty$. As a consequence, if $(x_k)_{k \in \mathbb{N}}$ is a sequence in $[c, b[$ such that $\lim_{k \to \infty} x_k = b$, then

$$\lim_{k \to \infty} F(x_k) = \lim_{k \to \infty} \int_c^{x_k} f = \lim_{k \to \infty} \int_c^{x_k} g = \infty,$$

showing that $\int_I f$ diverges as well.

(b): We assume $\int_I f$ to converge absolutely, i.e. $\int_I |f|$ must exist in $\mathbb{R}$. Since $0 \leq f^+ \leq |f|$ and $0 \leq f^- \leq |f|$, (a) then implies the existence of $\int_I f^+$ and of $\int_I f^-$. Thus, according to Lem. 10.36(a), $\int_I f = \int_I f^+ - \int_I f^-$ must also exist.   ∎

Example 10.40. (a) We will use Prop. 10.39(a) to show that the improper integral

$$\int_0^\infty e^{-t^2}\, dt$$

exists. Indeed,

$$\underset{t \in \mathbb{R}}{\forall} \left((t-1)^2 = t^2 - 2t + 1 \geq 0 \quad \Rightarrow \quad -t^2 \leq -2t + 1 \quad \Rightarrow \quad 0 \leq e^{-t^2} \leq e^{-2t+1} \right),$$

and, since

$$\int_0^\infty e^{-2t+1}\, dt = \lim_{x \to \infty} \int_0^x e^{-2t+1}\, dt = \lim_{x \to \infty} \left[-\frac{e^{-2t+1}}{2} \right]_0^x = \lim_{x \to \infty} \frac{e - e^{-2x+1}}{2} = \frac{e}{2},$$

Prop. 10.39(a) implies that $\int_0^\infty e^{-t^2}\, dt$ exists in $\mathbb{R}$.

(b) We will use Prop. 10.39(a) to show that

$$\int_0^\infty e^{t^2}\, dt$$

diverges. Indeed,

$$\underset{t \in \mathbb{R}}{\forall} \left(t^2 \geq 0 \quad \Rightarrow \quad e^{t^2} \geq 1 \right),$$

and, since

$$\lim_{x \to \infty} \int_0^x 1\, dt = \lim_{x \to \infty} x = \infty,$$

Prop. 10.39(a) implies that $\int_0^\infty e^{t^2}\, dt = \infty$.

(c) We provide an example that shows an improper integral can converge without converging absolutely: Consider the function

$$f : [0, \infty[\longrightarrow \mathbb{R}, \quad f(t) := \begin{cases} (-1)^{n+1} & \text{for } n \leq t \leq n + \frac{1}{n},\ n \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases} \tag{10.88}$$

Then

$$\int_0^\infty |f| = \lim_{k \to \infty} \sum_{k=1}^n \int_k^{k+\frac{1}{k}} 1\, dt = \lim_{k \to \infty} \sum_{k=1}^n \frac{1}{k} \overset{(7.72)}{=} \infty, \tag{10.89}$$

showing $\int_0^\infty f$ does not converge absolutely. However, we will show

$$\int_0^\infty f = \sum_{j=1}^\infty \frac{(-1)^{j+1}}{j} =: \alpha > 0. \tag{10.90}$$

We know $\alpha > 0$ from Ex. 7.86(a) and Th. 7.85. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}_0^+$ such that $\lim_{k\to\infty} x_k = \infty$. Given $\epsilon > 0$, choose $K \in \mathbb{N}$ such that $\frac{1}{K} < \frac{\epsilon}{2}$ and $N \in \mathbb{N}$ such that

$$\underset{k>N}{\forall} \quad x_k > K. \tag{10.91}$$

Then, for each $k > N$, there exists $K_1 \in \mathbb{N}$ such that $K < K_1 \le x_k < K_1 + 1$. Thus

$$\int_0^{x_k} f(t)\, \mathrm{d}t = \min\left\{ x_k - K_1, \frac{1}{K_1} \right\} + \sum_{j=1}^{K_1-1} \frac{(-1)^{j+1}}{j} \tag{10.92}$$

and

$$\left| \alpha - \int_0^{x_k} f(t)\, \mathrm{d}t \right| = \left| \sum_{j=K_1}^{\infty} \frac{(-1)^{j+1}}{j} - \min\left\{ x_k - K_1, \frac{1}{K_1} \right\} \right| \overset{(7.79)}{<} \frac{1}{K_1} + \frac{1}{K_1}$$

$$< \frac{2}{K} < 2 \cdot \frac{\epsilon}{2} = \epsilon, \tag{10.93}$$

thereby proving (10.90).

References

[EFT07]   H.-D. EBBINGHAUS, J. FLUM, and W. THOMAS. *Einführung in die mathematische Logik*, 5th ed. Spektrum Akademischer Verlag, Heidelberg, 2007 (German).

[EHH+95]   H.-D. EBBINGHAUS, H. HERMES, F. HIRZEBRUCH, M. KOECHER, K. MAINZER, J. NEUKIRCH, A. PRESTEL, and R. REMMERT. *Numbers.* Graduate Texts in Mathematics, Vol. 123, Springer-Verlag, New York, 1995, corrected 3rd printing.

[Jec73]   T. JECH. *The Axiom of Choice.* North-Holland, Amsterdam, 1973.

[Kun12]   KENNETH KUNEN. *The Foundations of Mathematics.* Studies in Logic, Vol. 19, College Publications, London, 2012.

[Wal02]   WOLFGANG WALTER. *Analysis 2*, 5th ed. Springer-Verlag, Berlin, 2002 (German).