Proofs in Elementary Analysis

Branko Ćurgus

June 10, 2025 00:00

Contents

CHAPTER 1

Introduction

1.1. Why Mathematics?

Mathematics has experienced millennia of glorious advancement. Its accomplishments are among the greatest achievements of humanity. Just over the last hundred years, we have witnessed the extensive mathematization of many aspects of life in general, and of all sciences in particular. No other human endeavor can claim such success. There is no reason to believe that this course will change in the millennia to come.

And all this has been achieved purely on the merit of its intellectual content, without resorting to marketing, propaganda, or persuasion so common in many other human activities.

The conclusion that inevitably follows from this success is that there are profound qualities in the broad field of mathematics. The single quality that must have contributed to its success is the intellectual rigor with which mathematics is developed.

Mathematics has something to offer for everyone. Everyone can benefit from the deliberate practice of rigorous thinking, from clarity and precision in the use of language, and from other intellectual experiences encountered in a well-structured mathematics course at any level.

1.2. The Branko Ćurgus Mathematical Experience

Mathematics has always been a personal experience for me. I want to create an environment where you can embrace mathematics as your own personal experience. How can this be achieved? Begin by acknowledging what you don't understand without fear. Discuss challenges openly with others. Open your mind, ask questions. Share your questions generously with others, treating them like precious gems. Questions are the gateway to deeper comprehension. Indeed, questions serve as a bridge from confusion to clarity.

I want the greatness of mathematics to permeate every class period with my students. I believe that every little piece of mathematics reflects some of its profound qualities, some of its essence. With my students and for my students, I want to make this engaging; I want to make it come alive.

1.3. Rigorous reasoning

The main benefit a student should take from a mathematics course is exposure to rigorous reasoning. In no other field is rigor as accessible to students as it is in a mathematics class. But what are the main features of rigorous reasoning? Below I offer five pillars of rigorous reasoning. Pillars do not provide a guide on how to construct a rigorous response to a presented problem. Instead, the pillars provide a framework to recognize whether a presented work is rigorous.

1.3.1. Five pillars of rigorous reasoning.

(I) Define all *relevant concepts* precisely, or provide references that rigorously present the definitions.
(II) State clearly which question or *problem* is under consideration.
(III) Clearly state the *assumptions* under which the proposed conclusion holds.
(IV) State the *background knowledge* you will use in your argument and provide references that rigorously present the background knowledge.
(V) Present *logical reasoning* that deduces the conclusion from the assumptions and the background knowledge. Present logical reasoning which demonstrates that the background knowledge is *applicable* as used in your reasoning.

1.3.2. How to ensure that your reasoning holds under critical reevaluation:

- ➢ Critically evaluate all aspects of your reasoning: Have all relevant concepts been included? Are all assumptions used? Are there some hidden assumptions that are used but not stated? Is your logic solid?
- ➢ Sometimes, it is necessary to be vague about background knowledge. Be honest and admit it in your presentation of the background knowledge.
- ➢ Be open to improvements. Look for hidden clues that suggest improvements.
- ➢ If the presentation is cumbersome, break it into smaller pieces to make it more transparent.
- ➢ Illustrate your reasoning with multiple specific examples. Do the examples reveal something that you might have overlooked in your work?
- ➢ Support your reasoning with analogies and similar settings to make it easier to understand.

1.4. Goals of this class

- ➢ To provide a systematic foundation for basic concepts encountered in calculus, especially those associated with the structure of the real numbers and in particular notions related to the Completeness Axiom.
- ➢ To introduce students to the structure and role of proofs in mathematics. Specifically, we assert that the only way to truly understand proofs is by constructing them on your own.
- ➢ To develop the ability to critically read and judge the correctness and completeness of mathematical reasoning.
- ➢ To develop a skill in presenting mathematical reasoning clearly and precisely.

1.5. Strategies

How to get started towards a solution of a problem?

- ➢ Illustrate the problem with several examples.
- ➢ Make sure that you understand the terminology used in the problem. Review all relevant definitions.
- ➢ Can you restate the problem as an implication? (Clearly identify the assumptions and the conclusion of the implication.)
- ➢ Identify problems done in class that are in some sense related to the problem that you are working on. Review proofs of those problems.
- ➢ Try to identify tools that can be used in the solution of the problem.
- ➢ If you can not solve the given problem, try to formulate a related simpler problem that you can solve. For example, try to solve a special case.
- ➢ Be flexible. Have in mind that there are many ways to approach each problem.
- ➢ Keep a detailed written record of your work.

How to avoid mistakes?

- ➢ Write your solution out carefully. Include justifications for all arguments that you use.
- ➢ Read your solution critically after a day or two. Is everything that you use in your proof justified?
- ➢ Imagine that a skeptic is reading your proof. Can you answer all sceptic's questions?

CHAPTER 2

A brief review of mathematical logic

Proofs in mathematics are based on logic, the science of rigorous reasoning. I believe the role of the formal logic is to make explicit what is already embedded in our intuitive thought processes. In the spirit of this belief, it is essential that you internalize logical rules not as abstractions to be memorized, but as guidelines that strengthen your natural reasoning abilities.

The fundamental building blocks of logic are propositions. The basic feature of a proposition is that it is either true or false.

2.1. Propositional calculus

2.1.1. Propositions. A *proposition*, also known as a *statement*, is a declarative sentence that is either true or false. For example, the sentence 'Two equals one plus one' is a true proposition. In mathematical notation, this proposition is written as $2 = 1 + 1$, and it is true. Conversely, an example of a false proposition is $0 = 1$.

2.1.2. Compound propositions. We now introduce several natural ways of combining propositions to form new propositions, known as *compound propositions*. Similar to how letters represent numbers in algebra and other mathematical disciplines, in logic, letters are used as variables to represent propositions. The most significant compound propositions in mathematics include the following: negation, conjunction, disjunction, implication (or conditional), and equivalence (or biconditional).

DEFINITION 2.1.1. The *negation* of a proposition p is the proposition "not p" which is false when p is true and which is true when p is false. This proposition is denoted by $\neg p$.

DEFINITION 2.1.2. The *conjunction* of propositions p and q is the proposition "p and q" which is true when both p and q are true and false otherwise. This proposition is denoted by $p \wedge q$. The conjunction of three propositions p, q and r is defined as $(p \wedge q) \wedge r$ which is true when all three propositions are true and false otherwise.

DEFINITION 2.1.3. The *disjunction* of propositions p and q is the proposition "p or q" which is false when both p and q are false and true otherwise. This proposition is denoted by $p \vee q$. The disjunction of three propositions p, q

and r is defined as $(p \vee q) \vee r$ which is false when all three propositions are false and true otherwise.

DEFINITION 2.1.4. The *implication* or *conditional* of propositions p and q is the proposition "If p, then q" which is false when p is true and q is false and true otherwise. This proposition is denoted by $p \Rightarrow q$.

DEFINITION 2.1.5. The *equivalence* or *biconditional* of propositions p and q is the proposition "p if and only if q" which is true when both $p \Rightarrow q$ and $q \Rightarrow p$ are true. This proposition is denoted by $p \Leftrightarrow q$.

The above definitions are summarized in the following *truth tables*. Since propositions can take only two truth values, *false* and *true*, we abbreviate them by F and T, respectively. In the tables below, the values F and T are always listed in alphabetical order. This consistent ordering provides a predictable structure that makes it easier to write, read and compare truth tables.

negation

p	$\neg p$
F	T
T	F

conjunction

p	q	$p \wedge q$
F	F	F
F	T	F
T	F	F
T	T	T

disjunction

p	q	$p \vee q$
F	F	F
F	T	T
T	F	T
T	T	T

implication

p	q	$p \Rightarrow q$
F	F	T
F	T	T
T	F	F
T	T	T

equivalence

p	q	$p \Leftrightarrow q$
F	F	T
F	T	F
T	F	F
T	T	T

2.1.3. Propositional calculus. Doing calculations with compound propositions is called *propositional calculus*. Here are the most important rules of propositional calculus.

The double negation rule: *The negation of the negation of a proposition is logically equivalent to the original proposition.* In symbols: $\neg(\neg p) \Leftrightarrow p$. To prove this equivalence, we construct the truth table:

p	$\neg p$	$\neg(\neg p)$	p
F	T	F	F
T	F	T	T

De Morgan's Laws: *The negation of a conjunction is logically equivalent to the disjunction of the corresponding negations.* In symbols: $\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$. To prove this equivalence, we construct the truth table:

p	q	$p \wedge q$	$\neg(p \wedge q)$	$\neg p \vee \neg q$	$\neg p$	$\neg q$
F	F	F	T	T	T	T
F	T	F	T	T	T	F
T	F	F	T	T	F	T
T	T	T	F	F	F	F

The negation of a disjunction is logically equivalent to the conjunction of the corresponding negations. In symbols: $\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q$. To prove this equivalence, we construct the truth table:

p	q	$p \vee q$	$\neg(p \vee q)$	$\neg p \wedge \neg q$	$\neg p$	$\neg q$
F	F	F	T	T	T	T
F	T	T	F	F	T	F
T	F	T	F	F	F	T
T	T	T	F	F	F	F

The negation of an implication: *The negation of an implication $p \Rightarrow q$ is logically equivalent to the conjunction of p and the negation of q.* In symbols: $\neg(p \Rightarrow q) \Leftrightarrow p \wedge \neg q$. To prove this equivalence, we construct the truth table:

p	q	$p \Rightarrow q$	$\neg(p \Rightarrow q)$	$p \wedge (\neg q)$	p	$\neg q$
F	F	T	F	F	F	T
F	T	T	F	F	F	F
T	F	F	T	T	T	T
T	T	T	F	F	T	F

REMARK 2.1.6. None of the four rules of propositional calculus stated above should be memorized as formal rules devoid of common-sense content. These rules are here to help us organize our intuitive thought process.

To illustrate this, consider a natural occurrence of the negation of an implication in a parent-child interaction. Suppose a parent says to their child: “If you finish your spinach, then I will serve ice cream.” The child later sees that they were not served ice cream and immediately thinks, or complains: “I was not served ice cream, even though I finished my spinach! The parent broke their promise! Instead of a promised implication, I am being served its negation.”

Here, the child naturally understands the negation of the original implication: “If p, then q” was promised, but, instead, p occurred and $\neg q$ occurred. The child’s intuitive reasoning explicitly invokes the negation of the implication: $p \wedge \neg q$. ◁

2.2. Implications

Since the most important statements in mathematics are formulated as implications, implications deserve a section in this brief summary.

2.2.1. Implication in the English Language. Probably because of the importance of implications, the English language offers many ways to express the implication $p \Rightarrow q$. Below are twelve common formulations:

(I) If p, then q. (II) If p, q. (III) q if p. (IV) q when p.

(V) p is sufficient for q. (VI) q is necessary for p.

(VII) A sufficient condition for q is p. (VIII) A necessary condition for p is q.

(IX) p implies q. (X) p only if q.

(XI) q whenever p. (XII) q follows from p.

DEFINITION 2.2.1. Let $p \Rightarrow q$ be an implication. The proposition p is called the *hypothesis* of the implication $p \Rightarrow q$. The proposition q is called the *conclusion*.

2.2.2. Three related implications. For a given implication, there are three named related implications.

DEFINITION 2.2.2. Let $p \Rightarrow q$ be an implication. The *contrapositive* of $p \Rightarrow q$ is the implication $\neg q \Rightarrow \neg p$. The *converse* of $p \Rightarrow q$ is the implication $q \Rightarrow p$. The *inverse* of $p \Rightarrow q$ is the implication $\neg p \Rightarrow \neg q$.

Although we are interested in mathematical propositions, it is often useful to explore real life propositions. My favorite real-life example relates to Red Square, the main square with a fountain on the WWU campus.

Name	Claim
p	It rains on Red Square.
q	Red Square is wet.
$\neg p$	It does not rain on Red Square.
$\neg q$	Red Square is not wet.

When dealing with real-life we assume that the claims that we use as propositions truly are either true or false. We do not accept that there are ambiguous situation in which we can not decide whether it is raining or not.

Let us explore the implication $p \Rightarrow q$ and three related implications which involve the above real-life propositions p and q.

	Claim (below RS stands for Red Square)	Name
$p \Rightarrow q$	If it rains on RS, then RS is wet.	Implication
$q \Rightarrow p$	If RS is wet, then it rains on RS.	Converse
$\neg q \Rightarrow \neg p$	If RS is not wet, then it does not rain on RS.	Contrapositive
$\neg p \Rightarrow \neg q$	If it does not rain on RS, then RS is not wet.	Inverse

In my experience the implication $p \Rightarrow q$ in the preceding table is true. I have never witnessed that Red Square is not wet when it is raining on Red Square. In my experience the converse $q \Rightarrow p$ is not always true. On few occasions, I have witnessed wet Red Square, on a sunny day. For example, the fountain flooded once and Red Square was truly wet on a sunny, that is a definitely nonrainy, day. From this we conclude that the implication and its converse are not related. It is possible for an implication to be true, while the converse is not true. In fact, whenever you see an implication proved in mathematics, a good question to ask: Is the converse true?

2.2.3. The contrapositive. The following is the most important fact in this brief review: *A proposition is equivalent to its contrapositive.* The easiest way to see this is to look at the negations. The negation of the implication $p \Rightarrow q$ is $p \wedge (\neg q)$. The negation of the contrapositive $(\neg q) \Rightarrow (\neg p)$ is $(\neg q) \wedge \big(\neg(\neg p)\big)$. By the double negation rule the proposition $\neg(\neg p)$ is equivalent to p. Hence the negation of the contrapositive is equivalent to $(\neg q) \wedge p$. But, $p \wedge (\neg q)$ and $(\neg q) \wedge p$ are clearly equivalent. So, the negations of the implication and its contrapositive have the identical truth values. Therefore the implication and its contrapositive have the identical truth values.

The fact that *a proposition is equivalent to its contrapositive* can also be seen by forming the following truth table:

p	q	$p \Rightarrow q$	$\neg q \Rightarrow \neg p$	$\neg q$	$\neg p$
F	F	T	T	T	T
F	T	T	T	F	T
T	F	F	F	T	F
T	T	T	T	F	F

Many mathematical statements are formulated as implications. Since an implication and its contrapositive are logically equivalent, it is good practice, whenever we encounter an implication, to write down and internalize the meaning of its contrapositive.

In the spirit of formal logic as a framework for our intuitive thought processes, consider a literary example from the Sherlock Holmes story *Silver Blaze* by Arthur Conan Doyle. Holmes reasons: "If the dog did not bark, then the midnight visitor was someone the dog knew well." This implication fits the storys context, but it is the contrapositive of the more explicit logical form: "If the midnight visitor was someone the dog did not know, then the dog would have barked." Holmes never states the logical structure outright; he applies it intuitively. Humans often reason in this way, even without formalizing the logic.

2.2.4. Implication Combined with Conjunction. I will continue with the Sherlock Holmes story *Silver Blaze*. The setting is that Silver Blaze, a famous racing horse, was abducted from its stable by a midnight visitor while the dog was in the stable. So, Holmess implication could have read: "If the midnight visitor was in the stable and the dog did not bark, then the visitor was someone the dog knew well." Here we have a combination of a conjunction and an implication. The structure is of the form $p \wedge q \Rightarrow r$.

What is an equivalent implication resembling the contrapositive? I suggest: "If the midnight visitor was in the stable and was someone the dog did not know, then the dog would have barked." Thus, the relevant contrapositive of the implication $p \wedge q \Rightarrow r$ is $p \wedge \neg r \Rightarrow \neg q$. I call this the *partial contrapositive* of $p \wedge q \Rightarrow r$.

To prove that $p \wedge q \Rightarrow r$ and $p \wedge \neg r \Rightarrow \neg q$ are logically equivalent, consider the negations of both implications:

$$\begin{aligned} p \wedge q \Rightarrow r &\qquad \text{has negation} \qquad p \wedge q \wedge \neg r, \\ p \wedge \neg r \Rightarrow \neg q &\qquad \text{has negation} \qquad p \wedge \neg r \wedge q. \end{aligned}$$

To form the second negation, we used that $\neg(\neg q)$ is equivalent to q. Since the two negations are logically equivalent, the original implications are equivalent.

Another way to establish the equivalence is to present a truth table:

p	q	r	$\neg q$	$\neg r$	$p \wedge q$	$p \wedge \neg r$	$p \wedge q \Rightarrow r$	$p \wedge \neg r \Rightarrow \neg q$
F	F	F	T	T	F	F	T	T
F	F	T	T	F	F	F	T	T
F	T	F	F	T	F	F	T	T
F	T	T	F	F	F	F	T	T
T	F	F	T	T	F	T	T	T
T	F	T	T	F	F	F	T	T
T	T	F	F	T	T	T	F	F
T	T	T	F	F	T	F	T	T

Implications of the form $p \wedge q \Rightarrow r$ occur often in mathematics. As before, it is good practice, whenever we encounter them, to write down and internalize the meaning of their partial contrapositives. Note that there are two equivalent forms: $p \wedge \neg r \Rightarrow \neg q$ and $\neg r \wedge q \Rightarrow \neg p$.

2.2.5. Implication Combined with Disjunction. Occasionally, we encounter implications of the form $p \Rightarrow q \vee r$. A logically equivalent implication is $p \wedge \neg q \Rightarrow r$. To prove that these two implications are logically equivalent, we analyze their negations.

By De Morgan's Law, the negation of $q \vee r$ is logically equivalent to $\neg q \wedge \neg r$. Hence, the negations of the two implications are:

$$\begin{array}{rcl} p \Rightarrow q \vee r & \text{has negation} & p \wedge \neg q \wedge \neg r, \\ p \wedge \neg q \Rightarrow r & \text{has negation} & p \wedge \neg q \wedge \neg r. \end{array}$$

Since the negations are logically equivalent, the original implications must be logically equivalent as well.

Notice that the contrapositive of the proposition $p \Rightarrow q \vee r$ is $\neg(q \vee r) \Rightarrow \neg p$, which, by De Morgan's Law, is equivalent to $\neg q \wedge \neg r \Rightarrow \neg p$. This last implication is of the form studied in Subsection 2.2.4: **Implication Combined with Conjunction**. A partial contrapositive of this implication is $p \wedge \neg q \Rightarrow r$, which is exactly the logically equivalent proposition from the beginning of this subsection.

One could argue that Subsections 2.2.4 and 2.2.5 present the same content from two different points of view. The moral of this section of the chapter **A Brief Review of Mathematical Logic** is to understand that there are multiple, logically equivalent ways to express an implication. When we encounter an implication to be proved or applied, we can explore and benefit from these different, logically equivalent formulations.

2.2.6. Before Moving On. Since this section referenced the story *Silver Blaze* as a source of implicit logical reasoning, it is worth pointing out an explicit appearance of a formal rule of logic, called the *disjunctive syllogism*, in the same story. This rule has not yet been introduced in these notes.

Sherlock Holmes says: "I have already said that he must have gone to Kings Pyland or to Mapleton. He is not at Kings Pyland. Therefore he is at Mapleton."

Observe the following implication present in Holmes' reasoning:

$$\big(p \vee q\big) \wedge \neg p \;\; \Rightarrow \;\; q.$$

Regardless of the truth values of the propositions p and q, this implication is always true. A compound proposition that is true for all possible truth values of its components is called a *tautology*. To verify that Holmes' reasoning is a tautology, we construct its truth table:

p	q	$p \vee q$	$\neg p$	$\big(p \vee q\big) \wedge \neg p$	$\big(p \vee q\big) \wedge \neg p \Rightarrow q$
F	F	F	T	F	T
F	T	T	T	T	T
T	F	T	F	F	T
T	T	T	F	F	T

2.3. Propositional functions

2.3.1. Propositional functions. The next level in the hierarchy of logical constructs is that of propositional functions. A *propositional function* or *predicate* is an expression which involves one or more variables and which becomes a proposition when variables are replaced by specific values from a particular set of values called the *universe of discourse*.

For example: $2x^2 - x > 0$ is not a proposition; it is a propositional function whose universe of discourse can be any set of numbers. Denote the the propositional function $2x^2 - x > 0$ by $Q(x)$. Let $S = \{-1, 0, 1\}$ be the universe of discourse for $Q(x)$. Then we can form three specific propositions $Q(-1), Q(0), Q(1)$:

x	$Q(x)$	meaning	T or F
-1	$Q(-1)$	$3 > 0$	T
0	$Q(0)$	$0 > 0$	F
1	$Q(1)$	$1 > 0$	T

Now we can form the conjunction of these three propositions $Q(-1) \wedge Q(0) \wedge Q(1)$ which is

$$(3 > 0) \wedge (0 > 0) \wedge (1 > 0)$$

which is false since not all of the three propositions are true. The negation of this compound proposition

$$(3 \leq 0) \vee (0 \leq 0) \vee (1 \leq 0)$$

is true since the proposition $0 \leq 0$ is true.

Similarly we can form the disjunction of the propositions in the table: $Q(-1) \vee Q(0) \vee Q(1)$ which is

$$(3 > 0) \vee (0 > 0) \vee (1 > 0)$$

which is true since not all of the three propositions are false. The negation of the preceding compound proposition

$$(3 \leq 0) \wedge (0 \leq 0) \wedge (1 \leq 0)$$

is false since not all three propositions in the conjunction are true; the proposition $1 \leq 0$ is false.

2.3.2. Quantifiers. If we consider the universe of discourse of the propositional function $Q(x)$ considered above to be the set $\mathbb{R}$ of all real numbers, then forming the conjunction and the disjunction of all the propositions $Q(x)$ with $x \in \mathbb{R}$ would be impossible in a way that we did above. Therefore we introduce the concept of the *universal quantifier* and the *existential quantifier*.

The symbol for the *universal quantifier* is $\forall$. We read it "for all." If $P(x)$ is any propositional function with the universe of discourse U then

$$\forall x \in U \quad P(x) \quad \text{means:} \quad P(x) \text{ is true for all } x \in U.$$

Consider a specific example. Use the propositional function $2x^2 - x > 0$ and the universal quantifier. The statement

$$\forall x \in \mathbb{R} \quad 2x^2 - x > 0$$

is a proposition. The truth value of the displayed proposition is F (false) since $2x^2 - x > 0$ is not true for all $x \in \mathbb{R}$. For example, with $x = 0$ we have $2(0)^2 - 0 = 0 > 0$, which is false.

The symbol for the *existential quantifier* is $\exists$. We read it "there exists." If $P(x)$ is any propositional function with the universe of discourse U then

$$\exists x \in U \ P(x) \quad \text{means that} \quad P(x) \text{ is true for at least one } x \in U.$$

or, in other words,

$$\exists x \in U \ P(x) \quad \text{means that} \quad \text{there exists } x \in U \ \text{ such that } \ P(x) \text{ is true.}$$

Consider a specific example. Use, again, the propositional function $2x^2 - x > 0$ and, this time, the existential quantifier. The statement

$$\exists x \in \mathbb{R} \quad 2x^2 - x > 0$$

is a proposition. The truth value is T (true) since for $x = 1 \in \mathbb{R}$ we have $2(1)^2 - 1 = 1 > 0$, which is true.

2.3.3. General De Morgan's Laws. The general form of De Morgan's laws describes how to negate quantified statements involving the universal quantifier ($\forall$) and the existential quantifier ($\exists$). The meanings of these quantified statements and their negations are summarized in the following table:

Statement	**When true?**	**When false?**	**Negation**
$\forall x \in U \ P(x)$	$P(x)$ is true for all $x \in U$	There exists an $x \in U$ such that $P(x)$ is false	$\exists x \in U \ \neg P(x)$
$\exists x \in U \ P(x)$	There exists an $x \in U$ such that $P(x)$ is true	$P(x)$ is false for all $x \in U$	$\forall x \in U \ \neg P(x)$

Below, the first general form of De Morgan's Law is expressed in logical notation:

$$\neg\big(\forall x \in U \quad P(x)\big) \quad \Leftrightarrow \quad \exists x \in U \quad \neg P(x).$$

We read this as: The negation of the proposition for all x in U, $P(x)$ is true is logically equivalent to the proposition there exists an x in U such that $P(x)$ is not true.

To illustrate this law, let us consider an example where the universe of discourse U is the set of all students in this class, and let $P(x)$ be the following statement:

$$P(x) \text{ is } \quad \boxed{x \text{ likes chocolate}}.$$

Then the proposition $\forall x \in U \quad P(x)$, in ordinary language, becomes:

$$\boxed{\text{Every student in this class likes chocolate}}.$$

This statement is either true or false. Regardless of its truth value, we can formulate its negation. The negation of the boxed proposition above is:

There exists a student in this class who does not like chocolate.

This example demonstrates how the negation of a proposition involving a universal quantifier produces a statement asserting the existence of a counterexample, in accordance with De Morgan's Law. In other words, to disprove the proposition Every student in this class likes chocolate, it is enough to find a single counterexample: that is, a student who does not like chocolate.

Below, the second general form of De Morgan's Law is expressed in logical notation:

$$\neg\big(\exists x \in U \quad P(x)\big) \quad \Leftrightarrow \quad \forall x \in U \quad \neg P(x).$$

We read this as: The negation of the proposition there exists an x in U such that $P(x)$ is true is logically equivalent to the proposition for all x in U, $P(x)$ is not true.

To illustrate this law, we continue with the universe of discourse U as the set of all students in this class, and we take $P(x)$ to be the same statement as above:

$P(x)$ is x likes chocolate.

Then the proposition $\exists x \in U \quad P(x)$, in everyday language, becomes:

There exists a student in this class who likes chocolate.

This statement is either true or false. Regardless of its truth value, we can formulate its negation. The negation of the boxed proposition above is:

Every student in this class does not like chocolate.

2.3.4. Statements with multiple quantifiers. To make things even more interesting mathematical statements often come with two or more quantifiers. Consider the statement

$$\forall a \in \mathbb{R} \ \ \exists x \in \mathbb{R} \quad ax^2 - x > 0. \tag{2.3.1}$$

Is this statement true or false? Its negation is

$$\exists a \in \mathbb{R} \ \ \forall x \in \mathbb{R} \quad ax^2 - x \leq 0.$$

The statement in (2.3.1) is true. We can prove it by considering two cases. If $a \geq 0$, then we can take $x = -1$. We get $a(-1)^2 - (-1) = a + 1 \geq 1 > 0$, which is true. If $a < 0$, then take $x = 1/(2a)$. Then, since $a < 0$, we have

$$a\left(\frac{1}{2a}\right)^2 - \frac{1}{2a} = \frac{1}{4a} - \frac{1}{2a} = -\frac{1}{4a} > 0.$$

So, in both cases the statement in (2.3.1) is true.

It is very important to realize that the order of quantifiers matters. The reversal of the quantifiers in (2.3.1) leads to the following statement

$$\exists x \in \mathbb{R} \ \ \forall a \in \mathbb{R} \quad ax^2 - x > 0.$$

This statement is false since its negation

$$\forall x \in \mathbb{R} \ \ \exists a \in \mathbb{R} \quad ax^2 - x \leq 0$$

is true. To prove the negation we consider two cases. The first case $x \geq 0$. Setting $a = 0$ we have $0(x)^2 - x = -x \leq 0$ which is true. The second case is $x < 0$. Assume $x < 0$. Setting $a = 2/x$ we have

$$\frac{2}{x}x^2 - x = x \leq 0,$$

which is a true statement since $x < 0$ in this case.

The simplest unsolved problem in mathematics is Goldbach's Conjecture which states a natural relationship between the even positive integers greater than 2 and the primes. The common notation for the set of all positive integers is $\mathbb{N}$. By $\mathbb{P}$ we denote the set of all prime numbers. Using the notation of mathematical logic, one of the most famous unsolved problems in mathematics, Goldbach's Conjecture, can be stated as follows:

$$\forall k \in \mathbb{N}\backslash\{1\} \;\; \exists p \in \mathbb{P} \;\; \exists q \in \mathbb{P} \quad \text{such that} \quad 2k = p + q.$$

Using sophisticated programming this proposition has been verified for all $k \in \mathbb{N}\backslash\{1\}$ such that than $k \leq 2 \times 10^{18}$. However, despite significant effort of generations of mathematicians there is no proof that Goldbach's Conjecture holds for all integers greater than 1.

2.4. Examples

In the proof of Proposition 2.4.2 below, I will demonstrate how I use colors to enhance the clarity of proofs. Most mathematical statements are implications of the form $p \Rightarrow q$. To prove such a statement, we assume p and use some previous mathematical knowledge, which I refer to as the *Background Knowledge*, along with logic to deduce q. To emphasize what is **assumed** and which **Background Knowledge** is used, these statements are colored **green**. The statement that needs to be proved, q, is colored **red**. As the proof progresses, more statements become **green**. Finally, at the end of the proof, the conclusion q becomes q, signifying that it has been **greenified** and thus proven.

Another way that I use colors is to color terms of different kind in algebraic expressions. For example, in your previous mathematical experience you have learned how to solve a quadratic equation: Let $a, b, c \in \mathbb{R}$ and assume that $a \neq 0$. Find $x \in \mathbb{R}$ such that

$$ax^2 + bx + c = 0.$$

Here $a, b, c \in \mathbb{R}$ are assumed to be known numbers (called coefficients) and x is unknown. To emphasise this **known**–**unknown** dichotomy, the coefficients $a, b, c \in \mathbb{R}$ are colored **green** and the unknown $x \in \mathbb{R}$ is colored **red**. So, the proper, colorful way to write a quadratic equation is as follows:

$$ax^2 + bx + c = 0.$$

With this coloring convention, solving the equation becomes just separating the colors. That is, we aim to express the **red** unknown in terms of **green** "knowns"

$$x = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \qquad \text{or} \qquad x = \frac{-b + \sqrt{b^2 - 4ac}}{2a}.$$

Using the notation of mathematical logic this short review of quadratic equation can be summarized as a proposition:

PROPOSITION 2.4.1. *Let $a, b, c, x \in \mathbb{R}$. Assume that $a \neq 0$ and $b^2 - 4ac \geq 0$. Then*

$$ax^2 + bx + c = 0 \quad \Leftrightarrow \quad x = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \;\bigvee\; x = \frac{-b + \sqrt{b^2 - 4ac}}{2a}.$$

PROOF. Think about how a proof of this proposition might proceed. □

The following proposition states something that you might have learned in your previous mathematical experience. Here I state the proposition using the notation of mathematical logic and use colors to make proof easier to follow.

PROPOSITION 2.4.2. *For all $a, b, c \in \mathbb{R}$ the following implication holds*

$$\forall x \in \mathbb{R} \quad ax^2 + bx + c \geq 0 \quad \Rightarrow \quad a \geq 0.$$

PROOF. Let $a, b, c \in \mathbb{R}$ be arbitrary. The colored version of the implication in the proposition is

$$\boxed{\forall x \in \mathbb{R} \quad ax^2 + bx + c \geq 0} \quad \Rightarrow \quad \boxed{a \geq 0}. \tag{2.4.1}$$

The implication in (2.4.1) is equivalent to its contrapositive

$$\boxed{a < 0} \quad \Rightarrow \quad \boxed{\exists x \in \mathbb{R} \quad ax^2 + bx + c < 0}. \tag{2.4.2}$$

Proof of (2.4.2) must start with the following phrase: Let $a \in \mathbb{R}$ be arbitrary and assume

$$\boxed{a < 0}. \tag{2.4.3}$$

To prove the implication in (2.4.2) we need to understand the coloring of the variables involved. The real numbers $a, b, c \in \mathbb{R}$ are given. They are arbitrary but fixed real numbers. To emphasise this fact I color them **green**: $a, b, c \in \mathbb{R}$. We only know that $\boxed{a < 0}$. The variable $x \in \mathbb{R}$ which is under the existential quantifier is **red** since we **have to find** a real number x with the property specified in the conclusion in (2.4.2):

$$\boxed{\exists x \in \mathbb{R} \quad \text{such that} \quad ax^2 + bx + c < 0}.$$

In fact the redness of $x \in \mathbb{R}$ is the core of the redness of the conclusion in the implication in (2.4.2).

The colors indicate how to prove (2.4.2). We need to find a formula for a red number x in terms of the green numbers. To be more specific: For the given $a, b, c \in \mathbb{R}$ of which you know only that $\boxed{a < 0}$, our task is to find a formula for $x \in \mathbb{R}$ in terms of $a, b, c \in \mathbb{R}$ for which we will be able to prove the following inequality

$$ax^2 + bx + c < 0.$$

This is the shortest proof that I found. It is somewhat cryptic. I want you to understand that what I present is a correct proof. I will try to explain how I found it in the comment box after the proof.

Since the numbers $b, c \in \mathbb{R}$ are green, the following number is also green

$$d = \max\{b, c, 0\}.$$

In words, d is the largest of the three numbers b, c, and 0. In particular we have the following inequalities

$$\boxed{b \leq d}, \quad \boxed{c \leq d}, \quad \text{and} \quad \boxed{0 \leq d}. \tag{2.4.4}$$

Set

$$x = x_0 = 1 - \frac{d}{a}.$$

Now calculate

$$\begin{aligned} a(x_0)^2 + bx_0 + c &= a\left(1 - \frac{d}{a}\right)^2 + b\left(1 - \frac{d}{a}\right) + c \\ &= a\left(1 - 2\frac{d}{a} + \frac{d^2}{a^2}\right) + b - \frac{bd}{a} + c \\ &= a - 2d + \frac{d^2}{a} + b - \frac{bd}{a} + c \\ &= a + (b-d) + (c-d) + \frac{d(d-b)}{a}. \end{aligned}$$

Next, we consider the four summands in the preceding sum.

We have that $\boxed{a < 0}$ by assumption. We have that $\boxed{b - d \le 0}$, since $\boxed{b \le d}$ by (2.4.4). We have that $\boxed{c - d \le 0}$, since $\boxed{c \le d}$ by (2.4.4). We have that $\boxed{d(d-b) \ge 0}$, since $\boxed{d \ge 0}$ and $\boxed{d \ge b}$ by (2.4.4). Since $\boxed{a < 0}$ we have $\boxed{d(d-b)/a \le 0}$. To summarize, we have

$$\boxed{a < 0}, \quad \boxed{b - d \le 0}, \quad \boxed{c - d \le 0}, \quad \boxed{\frac{d(d-b)}{a} \le 0}, \quad \boxed{x_0 = 1 - \frac{d}{a}}.$$

Therefore

$$\boxed{a(x_0)^2 + bx_0 + c = a + (b-d) + (c-d) + \frac{d(d-b)}{a} < 0}.$$

Since

$$x_0 = 1 - \frac{\max\{b, c, 0\}}{a}$$

is given in terms of a, b, and c, and

$$\boxed{a(x_0)^2 + bx_0 + c < 0},$$

the proposition is proved. □

Thinking that led to the above proof.

I first decided to study a simpler version of the quadratic expression with $a = -1$. Can I find a simple value of x for which

$$-x^2 + bx + c < 0?$$

That turned out to be hard. So, I decided to simplify further by making b and c equal and I called it d. Can I find a simple value of x for which

$$-x^2 + dx + d < 0?$$

This turned out to be simpler. I discovered that for $x = 1 + d$ we have

$$-(1+d)^2 + d(1+d) + d = -1 - 2d - d^2 + d + d^2 + d = -1.$$

This was great. How to connect it to $-x^2 + bx + c$? If I take $b \le d$ and $c \le d$, then I will have

$$-x^2 + bx + c \le -x^2 + dx + d \quad \text{provided that } x \text{ is positive.}$$

Since I need a positive x and we used $x = 1 + d$, if we choose a positive d this could all work out. This inspired me to take

$$d = \max\{b, c, 0\}.$$

Then

$$\begin{aligned} -(1+d)^2 + b(1+d) + c &= -1 - 2d - d^2 + b + bd + c \\ &= -1 + (b-d) + (c-d) + d(b-d) \\ &\leq -1. \end{aligned}$$

How to incorporate a into this? Above, I considered $a = -1$. Where is -1 hiding in the above formulas? It turned out that $x = 1 - d/(-1)$ is the right answer. So, that is where $x = 1 - d/a$ comes from.

Proof of Proposition 2.4.2 that I wrote above is long. In that proof, I incorporated some thinking that could have been skipped. Below I present a shorter, so called bare-bones, version of the proof with just the essentials.

Bare-bones proof.

PROOF. Let $a, b, c \in \mathbb{R}$ be arbitrary. We will prove the contrapositive of the implication in the proposition. Assume $a < 0$. Set

$$d = \max\{b, c, 0\}.$$

As a consequence of this definition, the following inequalities hold

$$b \leq d, \quad c \leq d, \quad \text{and} \quad 0 \leq d. \tag{2.4.5}$$

Next we define a real number x_0. Set

$$x_0 = 1 - \frac{d}{a}.$$

At x_0 the quadratic expression takes a negative value. To prove this claim, calculate:

$$a(x_0)^2 + bx_0 + c = a\left(1 - 2\frac{d}{a} + \frac{d^2}{a^2}\right) + b - \frac{bd}{a} + c = a + (b-d) + (c-d) + \frac{d(d-b)}{a}.$$

By assumption we have $a < 0$ and from (2.4.5) we have $b - d \leq 0$, $c - d \leq 0$, and $d(d-b) \geq 0$. Therefore

$$a(x_0)^2 + bx_0 + c = a + (b-d) + (c-d) + \frac{d(d-b)}{a} < 0.$$

The proposition is proved. □

The preceding proposition is a part of the proof of the first equivalence in the following theorem which states three equivalences related to quadratic polynomials.

THEOREM 2.4.3. *Let $a, b, c \in \mathbb{R}$. Then the following equivalences hold*

$$\forall x \in \mathbb{R} \;\; ax^2 + bx + c \geq 0 \quad \Leftrightarrow \quad a \geq 0 \;\wedge\; c \geq 0 \;\wedge\; b^2 - 4ac \leq 0, \tag{2.4.6}$$

$$\forall x \in \mathbb{R} \;\; ax^2 + bx + c \leq 0 \quad \Leftrightarrow \quad a \leq 0 \;\wedge\; c \leq 0 \;\wedge\; b^2 - 4ac \leq 0, \tag{2.4.7}$$

$$\Big(\exists s \in \mathbb{R} \;\; as^2 + bs + c < 0\Big) \wedge \Big(\exists t \in \mathbb{R} \;\; at^2 + bt + c > 0\Big) \quad \Leftrightarrow \quad b^2 - 4ac > 0. \tag{2.4.8}$$

A complete proof of the equivalence (2.4.6) is on this webpage.

2.5. Exercises

In some exercises below and later in the notes, we use the following interval notation: For a real number a we set

$$\mathbb{R}_{<a} = \{x \in \mathbb{R} : x < a\},$$
$$\mathbb{R}_{>a} = \{x \in \mathbb{R} : a < x\}.$$

EXERCISE 2.5.1. In this exercise, the universe of discourse is the set of all positive real numbers, that is the set $\mathbb{R}_{>0}$. For each proposition formulate its negation and decide which is true, the stated proposition or its negation.

(a) For every $x \in \mathbb{R}_{>0}$, there exists $y \in \mathbb{R}_{>0}$ such that $xy < 1$.
(b) For every $x \in \mathbb{R}_{>0}$, there exists $y \in \mathbb{R}_{>0}$ such that $xy > 1$.
(c) There exists $x \in \mathbb{R}_{>0}$ such that for all $y \in \mathbb{R}_{>0}$, we have $xy \leq 1$.

As a second part of this exercise, explore the analogous statements when the universe of discourse is replaced by the set $\mathbb{R}$ of all real numbers. $\triangleleft$

EXERCISE 2.5.2. In this exercise, the universe of discourse is the set of all real numbers. Each of the quantified statements below represents a famous mathematical fact. State that fact in English without using formal logic.

(a) $\exists x \in \mathbb{R} \;\; \forall y \in \mathbb{R} \;\; x + y = y$
(b) $\exists x \in \mathbb{R} \;\; \forall y \in \mathbb{R} \;\; xy = y$
(c) $\forall x \in \mathbb{R} \;\; \forall y \in \mathbb{R} \quad x = 0 \wedge y = 0 \Leftrightarrow xy = 0$
(d) $\forall x \in \mathbb{R} \;\; \forall y \in \mathbb{R} \quad x < 0 \wedge y < 0 \quad \Rightarrow \quad xy > 0$
(e) $\forall x \in \mathbb{R} \;\; \forall y \in \mathbb{R} \;\; \exists z \in \mathbb{R} \quad x + y = z$

$\triangleleft$

EXERCISE 2.5.3. The universe of discourse in this exercise is the set of all real numbers. This will not be repeated in the statements. For each statement, formulate its negation and determine which one is true, the original statement or its negation. Provide a proof of your claim.

(i) $\forall x \; \exists y \;\; x^2 = y$
(ii) $\forall x \; \exists y \;\; x = y^2$
(iii) $\exists x \; \forall y \;\; xy = 0$
(iv) $\exists x \; \forall y \;\; y \neq 0 \;\Rightarrow\; xy = 1$
(v) $\forall x \; \exists y \;\; x + y = 1$ $\triangleleft$

EXERCISE 2.5.4. The universe of discourse in this exercise is the set of all positive real numbers. This will not be repeated in the statements. For each statement, formulate its negation and determine which one is true, the original statement or its negation. Provide a proof of your claim.

(i) $\forall a \; \forall b \; \exists c \quad cb > a$
(ii) $\forall x \; \exists y \; \forall z \quad z > y \quad \Rightarrow \quad zx > 1$
(iii) $\forall x \; \exists y \; \forall z \quad z > y \quad \Rightarrow \quad z^2 > x$ $\triangleleft$

EXERCISE 2.5.5. The universe of discourse in this exercise is the set of all positive real numbers. This will not be repeated in the statements. For each statement, formulate its negation and determine which one is truethe original statement or its negation. Provide a proof of your claim.

(i) $\forall a \; \forall b \; \exists c \quad cb > a$

(ii) $\forall x\ \exists y\ \forall z \quad z > y \ \Rightarrow\ zx > 1$

(iii) $\forall x\ \exists y\ \forall z \quad z > y \ \Rightarrow\ z^2 > x^2$ ◁

EXERCISE 2.5.6. With the interval notation introduced at the beginning of this section, prove the following proposition:

$$\forall y \in \mathbb{R}_{<1}\ \ \exists x \in \mathbb{R}_{>0} \quad \text{such that} \quad y < \frac{x}{x+1}.$$ ◁

EXERCISE 2.5.7. Let $a, b \in \mathbb{R}$. With the interval notation introduced at the beginning of this section, consider the following logical equivalence:

$$\forall x \in \mathbb{R}_{<a}\ \ x < b \quad \Leftrightarrow \quad \forall x \in \mathbb{R}_{<a}\ \ x \leq b.$$

Is this logical equivalence valid? Justify your answer by proving the equivalence of the two quantified statements, or by providing a counterexample if they are not logically equivalent. ◁

CHAPTER 3

Sets and functions

3.1. Sets

3.1.1. Sets, element of a set, empty set. By a *set* S we mean a well-defined collection of objects such that it can be determined whether or not any particular object is an element of S or not. If x is an object in the set S we say that x is an *element* of S and write $x \in S$. The negation of $x \in S$ is $x \notin S$.

The *empty set* is the unique set which contains no elements. The empty set is denoted by the symbol $\emptyset$.

Generally, capital letters will be used to denote sets of objects and lower case letters to denote objects themselves. However, watch for deviations of this rule. We will be concerned mainly with sets of real numbers. The specially designed letters $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ denote the following important sets of real numbers:

$\mathbb{Z}$ denotes the set of all *integers*,
$\mathbb{N}$ denotes the set of all *positive integers* (also caleed *natural numbers*),
$\mathbb{Q}$ denotes the set of all *rational numbers*,
$\mathbb{R}$ denotes the set of all *real numbers*.

A set can be described by:

- a statement such as "Let A be the set of real solutions of the equation $x^2 - x = 0$."
- a listing of all the elements; for example $A = \{0, 1\}$.
- the set builder notation: $A = \bigl\{x \in \mathbb{R} : x^2 = x\bigr\}$, which builds a set from a description of its elements .

Notice the usage of the braces (or curly brackets) $\{$ and $\}$ in the above examples to delimit sets. Notice the difference between a set and its elements. The number 0 is an important real number. However, $\{0\}$ denotes the set whose only element is 0.

The expression $\bigl\{x \in \mathbb{R} : x^2 = x\bigr\}$ is read as "the set of all real numbers x such that $x^2 = x$". In this context, the colon ":" is used as an abbreviation for the phrase "such that". However, many books use the vertical bar symbol $|$ as an alternative to colon.

3.1.2. Subset, power set, equal sets. It is common to illustrate relationships among sets using pictures. In this section, we represent sets with elliptical shapes. To distinguish between different sets, we use parallel hatchings with positive and negative slopes. Such illustrations are known as Venn diagrams.

Definition 3.1.1. A set B is a *subset* of a set A if every element of B is also an element of A. In this case we write $B \subseteq A$ or $A \supseteq B$. Using the logic notation this definition is expressed as follows:

$$B \subseteq A \quad \Leftrightarrow \quad \forall x \in U \; \big(x \in B \quad \Rightarrow \quad x \in A\big). \tag{3.1.1}$$

The set U in equivalence (3.1.1) stands for the universe of discourse. That is, U is a set that contains all the objects that are considered in a given situation. For example, in the figures in which we illustrate relations between sets, the universe of discourse is bounded by the black rectangle.

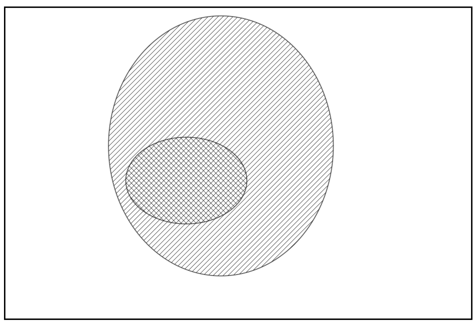


FIG. 3.1.1. The inclusion

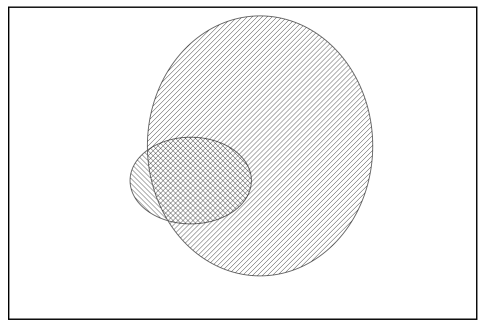


FIG. 3.1.2. Not a subset

Writing mathematical statements using formal logic notation, as in (3.1.1), enhances clarity. Also, it makes stating the negation easier. What does it mean that a set B is not a subset of A? The negation of $B \subseteq A$, denoted by $B \nsubseteq A$, holds if and only if there exists $b \in B$ such that $b \notin A$. Or, stated as the negation of each side of the statement in (3.1.1),

$$B \nsubseteq A \quad \Leftrightarrow \quad \exists x \in U \; \big(x \in B \; \wedge \; x \notin A\big).$$

Recall that the implication $\mathsf{F} \Rightarrow q$ is always true, regardless of the truth value of q. Therefore, the statement

$$\forall x \in U \quad x \in \emptyset \; \Rightarrow \; x \in A$$

is true. Consequently, the empty set is a subset of each set. In particular, $\emptyset \subseteq \emptyset$. However, notice that $\emptyset \in \emptyset$ is not true.

DEFINITION 3.1.2. Two sets A and B are *equal*, denoted $A = B$, if they contain precisely the same elements, that is, if and only if $A \subseteq B$ and $B \subseteq A$. Or, using formal logic notation,

$$A = B \quad \Leftrightarrow \quad \big(A \subseteq B \; \wedge \; B \subseteq A\big). \tag{3.1.2}$$

The negation of each side of the equivalence in (3.1.2) yields to a characterization of two sets not being equal:

$$A \neq B \quad \Leftrightarrow \quad \big(A \nsubseteq B \; \vee \; B \nsubseteq A\big).$$

The preceding definition of the equality of sets implies that the sets $\{0, 1, 0\}$ and $\{0, 1\}$ are equal. These sets are equal since they have the same elements. The fact that 0 is written twice in the first set is irrelevant. Also, the order in which elements are listed is irrelevant: $\{3, 2, 1\} = \{1, 2, 3\}$.

Equality is allowed in the definition of a subset. That is, a set is a subset of itself. If we wish to exclude this possibility we say B is a *proper subset* of A and we write $B \subsetneq A$ or $B \subset A$. Formally, $B \subsetneq A$ if and only if $x \in B$ implies $x \in A$ and there exists $a \in A$ such that $a \notin B$.

DEFINITION 3.1.3. Let S be a set. The *power set* of S, denoted by $\mathcal{P}(S)$, is the set of all subsets of S.

Below is the power set of the set $\{-1, 0, 1\}$:

$$\Big\{\emptyset, \{-1\}, \{0\}, \{1\}, \{-1, 0\}, \{-1, 1\}, \{0, 1\}, \{-1, 0, 1\}\Big\}.$$

We have

$$\begin{aligned} \mathcal{P}(\emptyset) &= \{\emptyset\} \\ \mathcal{P}\big(\{\emptyset\}\big) &= \Big\{\emptyset, \{\emptyset\}\Big\} \\ \mathcal{P}\Big(\Big\{\emptyset, \{\emptyset\}\Big\}\Big) &= \Big\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\Big\}. \end{aligned}$$

3.1.3. Set operations. In this subsection, we introduce the following set operations: the set complement, the union, the intersection, the set difference, and the set symmetric difference. In the Venn diagrams, the resulting set of a set operation is represented in yellow with a teal border. All sets in this section are subsets of a universal set U. In the Venn diagrams, the rectangle bounded by the black line represents the universal set.

DEFINITION 3.1.4. The *complement* of A is the set of all $x \in U$ such that $x \notin A$. The complement of A is denoted by A^c. Thus

$$A^c = \big\{x \in U : \ x \notin A\big\}.$$

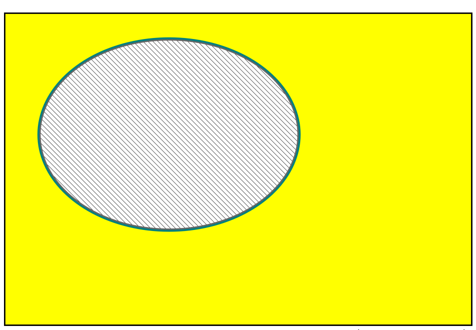


FIG. 3.1.3. The complement (in yellow)

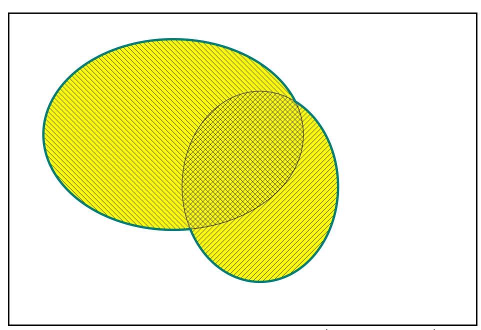


FIG. 3.1.4. The union (in yellow)

DEFINITION 3.1.5. The *union* of A and B is the set of all $x \in U$ such that $x \in A$ or $x \in B$. The union of A and B is denoted $A \cup B$. Thus

$$A \cup B \stackrel{\text{def}}{=} \big\{x \in U : \ x \in A \ \vee \ x \in B\big\}.$$

REMARK 3.1.6. The conjunction "or" in mathematics is always in an inclusive sense, that is, it is allowed in the definition that x belong to both A and B. For example, $\{0,1,2,3\} \cup \{2,3,4,5\} = \{0,1,2,3,4,5\}$. $\triangleleft$

DEFINITION 3.1.7. The *intersection* of A and B is the set of all x such that x is an element of A and x is an element of B. It is denoted $A \cap B$. Thus

$$A \cap B \stackrel{\text{def}}{=} \{x \in U : \; x \in A \;\wedge\; x \in B\}.$$

Two sets A and B are said to be *disjoint* if their intersection is the empty set, that is if $A \cap B = \emptyset$.

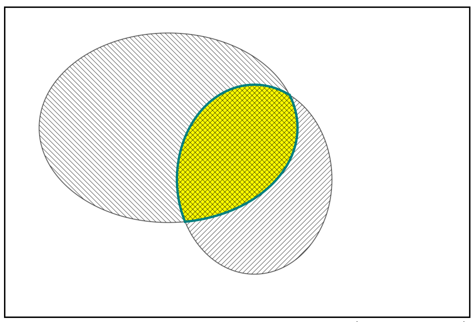


FIG. 3.1.5. The intersection (in yellow)

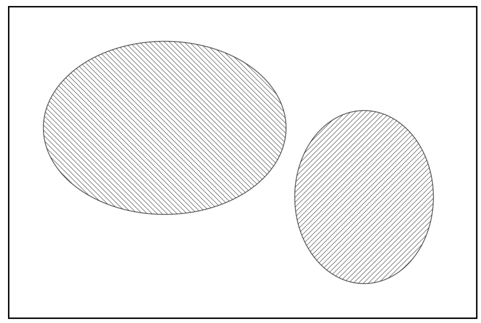


FIG. 3.1.6. Disjoint sets

DEFINITION 3.1.8. The *set difference* between the sets A and B, denoted by $A \setminus B$, is the set of all x such that $x \in A$ and $x \notin B$. That is,

$$A \setminus B \stackrel{\text{def}}{=} \{x \in U : \; x \in A \;\wedge\; x \notin B\}.$$

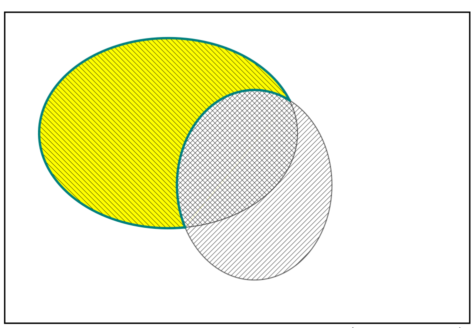


FIG. 3.1.7. A set difference (in yellow)

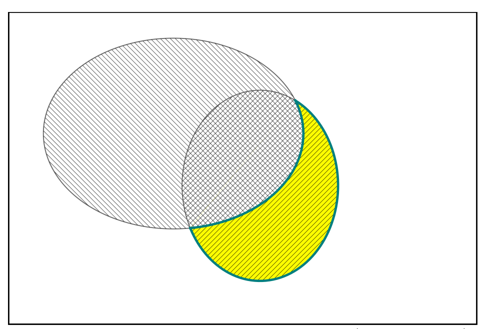


FIG. 3.1.8. A set difference (in yellow)

EXERCISE 3.1.9. Let $A, B \subseteq U$. Prove that the following three statements are equivalent

(i) $A \subseteq B$.
(ii) $B^c \subseteq A^c$.
(iii) $A \cap B^c = \emptyset$. $\triangleleft$

PROOF. Notice that for any set A we have $x \in A^c$ if and only if $x \notin A$. Or, what is equivalent, $x \notin A^c$ if and only if $x \in A$.

We will prove that (i) is equivalent to (iii) by proving the equivalent statement: $A \cap B^c \neq \emptyset$ if and only if $A \not\subseteq B$. The following sequence of equivalences proves the claim:

$$\begin{aligned} A \cap B^c \neq \emptyset \quad &\Leftrightarrow \quad \exists x \in U \quad x \in A \cap B^c \\ &\Leftrightarrow \quad \exists x \in U \quad x \in A \wedge x \in B^c \\ &\Leftrightarrow \quad \exists x \in U \quad x \in A \wedge x \notin B \\ &\Leftrightarrow \quad A \not\subseteq B. \end{aligned}$$

The equivalence of (ii) and (iii) is proved similarly. □

EXERCISE 3.1.10. Let $A, B \subseteq U$. Prove De Morgan's laws:

$$(A \cup B)^c = A^c \cap B^c \qquad \text{and} \qquad (A \cap B)^c = A^c \cup B^c. \qquad \triangleleft$$

EXERCISE 3.1.11. Let $A, B, C \subseteq U$. Prove the distributive properties of intersection and union.

Show that intersection distributes over union:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

and that union distributes over intersection:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C). \qquad \triangleleft$$

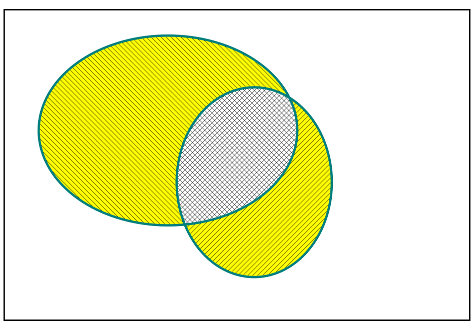


FIG. 3.1.9. A symmetric set difference (in yellow)

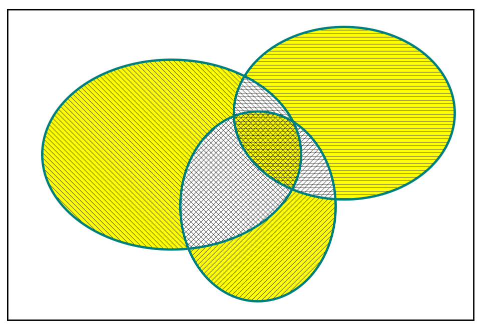


FIG. 3.1.10. A three set symmetric difference (in yellow)

EXERCISE 3.1.12. Let $A, B \subseteq U$. The *symmetric difference of two sets* is defined as

$$A \Delta B \stackrel{\text{def}}{=} (A \setminus B) \cup (B \setminus A).$$

Prove that the symmetric difference is associative

$$(A \Delta B) \Delta C = A \Delta (B \Delta C). \qquad \triangleleft$$

PROOF. One way to prove the given equality is to construct a *membership table*, in which we indicate the membership status of an arbitrary element x in all relevant sets.

$x \in$	A	B	C	$A\Delta B$	$B\Delta C$	$(A\Delta B)\Delta C$	$A\Delta(B\Delta C)$
	F	F	F	F	F	F	F
	F	F	T	F	T	T	T
	F	T	F	T	T	T	T
	F	T	T	T	F	F	F
	T	F	F	T	F	T	T
	T	F	T	T	T	F	F
	T	T	F	F	T	F	F
	T	T	T	F	F	T	T

□

3.1.4. Generalized unions and intersections. In this class we primarily consider sets of real numbers. Occasionally, we will also consider sets whose elements are sets of real numbers. When discussing sets of sets, it is customary to use the term "*family*" or "*collection*" instead of "set." For examples, see Section 4.3.

For any nonempty family of sets we can define the concepts of union and intersection. Let $\mathcal{F}$ be a nonempty family of sets. We define the *union of the family* $\mathcal{F}$ as

$$\bigcup_{A\in\mathcal{F}} A = \bigcup\{A : A \in \mathcal{F}\} \stackrel{\text{def}}{=} \{x \in U : \exists A \in \mathcal{F} \ \ x \in A\}.$$

We define the *intersection of the family* $\mathcal{F}$ as

$$\bigcap_{A\in\mathcal{F}} A = \bigcap\{A : A \in \mathcal{F}\} \stackrel{\text{def}}{=} \{x \in U : \forall A \in \mathcal{F} \ \ x \in A\}.$$

The concept of disjoint sets extends to families of sets. A family of sets is *pairwise disjoint* or *mutually disjoint* if, given any two distinct sets in the family, those two sets are disjoint.

Exercise 3.1.13. Let $\mathcal{F}$ be a nonempty family of subsets of U. Prove De Morgan's laws:

$$\Big(\bigcup\{A : A \in \mathcal{F}\}\Big)^c = \bigcap\{A^c : A \in \mathcal{F}\},$$
$$\Big(\bigcap\{A : A \in \mathcal{F}\}\Big)^c = \bigcup\{A^c : A \in \mathcal{F}\}.$$

◁

3.1.5. Ordered pairs and the Cartesian product of sets. In this subsection, we introduce the concept of an ordered pair and the Cartesian product of two sets.

Definition 3.1.14. An ordered pair is a collection of two elements, which are not necessarily distinct, one of which is distinguished as the first coordinate (or the first entry) and the other as the second coordinate (or the second entry). The common notation for an ordered pair with the first coordinate a and the second coordinate b is (a, b). Ordered pairs are also called two-tuples.

The definition of an ordered pair given above is somewhat informal. The formal definition is: Let a and b be given objects. The ordered pair (a, b) of objects a and b is the set $\{\{a\}, \{a, b\}\}$. However, it seems to me that this definition obscures more than it clarifies.

The ordered pairs $(0,1)$ and $(1,0)$ are different since their first entries are different. The ordered pairs $(0,0)$ and $(0,1)$ are different since their second entries are different. In general, $(a,b) = (x,y)$ if and only if $a = x$ and $b = y$.

Notice the usage of the round brackets (and) in the definition of an ordered pair. Please distinguish between $\{0,1\}$ and $(0,1)$: $\{0,1\}$ is a set with two elements, $(0,1)$ is an ordered pair, an object defined by Definition 3.1.14.

DEFINITION 3.1.15. The *Cartesian product* (or *direct product*) of two sets A and B, denoted $A \times B$, is the set of all possible ordered pairs whose first entry is a member of A and whose second entry is a member of B:

$$A \times B = \Big\{ (a,b) \ : \ a \in A \ \wedge \ b \in B \Big\}.$$

The main example of a Cartesian product is $\mathbb{R} \times \mathbb{R}$, which provides a coordinate system for the plane.

EXAMPLE 3.1.16. Let $A = \{1,2,3,4\}$ and let $C = \{\texttt{R}, \texttt{G}, \texttt{B}\}$ be the set of primary colors where $\texttt{R}$ stands for red, $\texttt{G}$ for green, and $\texttt{B}$ for blue. Then

$$A \times C = \Big\{ (1,\texttt{R}), (1,\texttt{G}), (1,\texttt{B}), (2,\texttt{R}), (2,\texttt{G}), (2,\texttt{B}), \\ (3,\texttt{R}), (3,\texttt{G}), (3,\texttt{B}), (4,\texttt{R}), (4,\texttt{G}), (4,\texttt{B}) \Big\}. \qquad \triangleleft$$

In a small example like this, the table below illustrates how the ordered pairs arise as entries in a grid, analogous to how points are formed in the coordinate plane $\mathbb{R} \times \mathbb{R}$:

	1	2	3	4
R	(1, R)	(2, R)	(3, R)	(4, R)
G	(1, G)	(2, G)	(3, G)	(4, G)
B	(1, B)	(2, B)	(3, B)	(4, B)

However, when one component of an ordered pair is a color, we can stretch our imagination a bit and think of a colored number as an ordered pair consisting of a number and a color:

$$\{1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4\}.$$

REMARK 3.1.17. (ON NOTATION) Ideally, mathematical terminology and notation should be completely free of ambiguities. We strive for absolute certainty. However, we will soon introduce the concept of an open interval, for which we will use the same notation as for an ordered pair. The intended meaning should be clear from the context. Whenever you are uncertain, look for the resolution of the uncertainty. $\triangleleft$

3.2. Functions

We will start with a definition of a function as commonly presented in undergraduate mathematics classes. Let A and B be two nonempty sets. A function from A to B is a rule f that assigns exactly one element of B to each element of A.

A limitation of this definition is that it relies on the undefined concept of a "rule." It is not clear what constitutes a valid rule defining a function. To address this limitation, in the next subsection, we identify a function with its graph and describe the properties that a graph must have.

3.2.1. A formal definition of a function.

Here we present the formal definition of a function.

DEFINITION 3.2.1. Let A and B be nonempty sets. A *function* f with *domain* A and *codomain* B is a subset of the Cartesian product $A \times B$ such that the following two conditions are satisfied:

TOTAL. For every $x \in A$ there exists $y \in B$ such that the pair (x, y) belongs to f.

UNIVAL. For all pairs (x_1, y_1), (x_2, y_2) in f the following implication holds: $x_1 = x_2$ implies $y_1 = y_2$.

A function f with domain A and codomain B is denoted by $f : A \to B$. If $f : A \to B$ is a function then instead of $(x, y) \in f$ we often write $y = f(x)$.

REMARK 3.2.2. (ON TERMINOLOGY) Stated in words, the condition TOTAL means that the function $f : A \to B$ is defined on the entire domain A. While the term TOTAL is not commonly used in the context of functions, it is used with exactly this meaning in the context of relations, that is, subsets of a Cartesian product. See the Wikipedia page Total relations.

To state the condition UNIVAL in words, let us first express its negation: there exist pairs (x_1, y_1) and (x_2, y_2) in f such that $x_1 = x_2$ and $y_1 \neq y_2$. In other words, UNIVAL rules out the possibility that the function f assigns two different values to the same input. This condition is commonly referred to in Precalculus as the *vertical line test*. In the Precalculus setting, A and B are nonempty subsets of $\mathbb{R}$, and the condition UNIVAL states that no two distinct points on a vertical line can lie on a graph of a function. I coined the term UNIVAL to emphasize that this condition means that $f : A \to B$ assigns unique values. $\triangleleft$

In the notation of formal logic, conditions TOTAL and UNIVAL read:

TOTAL. $\forall x \in A \;\; \exists y \in B$ such that $(x, y) \in f$.

UNIVAL. $\forall (x_1, y_1) \in f \;\; \forall (x_2, y_2) \in f$ we have $x_1 = x_2 \Rightarrow y_1 = y_2$.

Since the contrapositive is equivalent to the original implication, condition UNIVAL can be expressed in an equivalent way:

UNIVAL. $\forall (x_1, y_1) \in f \;\; \forall (x_2, y_2) \in f$ we have $y_1 \neq y_2 \Rightarrow x_1 \neq x_2$,

which is more convenient in some settings.

DEFINITION 3.2.3. Let A and B be nonempty sets and let $f : A \to B$ be a function. The set

$$\operatorname{ran}(f) \stackrel{\text{def}}{=} \Big\{ y \in B : \exists x \in A \ \ (x, y) \in f \Big\} = \Big\{ y \in B : \exists x \in A \ \ y = f(x) \Big\}$$

is called the *range* of f.

DEFINITION 3.2.4. Let A and B be nonempty sets. The set of all functions with domain A and codomain B is denoted by B^A.

The reason for the notation in the preceding definition will become clearer when you read Example 3.2.7 below.

In precalculus and calculus classes the sets A and B are typically subsets of $\mathbb{R}$, and functions are often defined by formulas. However, the sets A and B are often not explicitly specified. If this is the case, we consider that the domain of such a function is the largest subset of $\mathbb{R}$ on which the formula is defined. This domain is called the natural domain of a function. For example, the natural domain of the reciprocal function $r(x) = 1/x$ is the set $\mathbb{R} \backslash \{0\}$.

DEFINITION 3.2.5. Let A be a nonempty set. The *identity function* with domain A and codomain A, denoted by id_A, is the function defined as

$$\mathrm{id}_A \stackrel{\text{def}}{=} \big\{ (x, x) \in A \times A : x \in A \big\},$$

or, in the traditional notation,

$$\forall x \in A \quad \mathrm{id}_A(x) = x.$$

The identity function maps each element of A to itself.

EXAMPLE 3.2.6. Consider the sets $A = \{1, 2, 3, 4\}$ and $C = \{\mathtt{R}, \mathtt{G}, \mathtt{B}\}$ as in Example 3.1.16. The subset

$$\big\{ (1, \mathtt{G}), (2, \mathtt{R}), (3, \mathtt{G}), (4, \mathtt{B}) \big\}.$$

of $A \times C$ is a function in the sense of Definition 3.2.1. In the traditional notation this function is given by $f(1) = \mathtt{G}, f(2) = \mathtt{R}, f(3) = \mathtt{G}, f(4) = \mathtt{B}$. In contrast, the subset

$$\big\{ (1, \mathtt{B}), (2, \mathtt{G}), (2, \mathtt{R}), (3, \mathtt{R}), (4, \mathtt{G}) \big\}$$

of $A \times C$ is not a function since $(2, \mathtt{G}), (2, \mathtt{R})$ are in the subset and $\mathtt{G} \neq \mathtt{R}$. Therefore, condition UNIVAL in Definition 3.2.1 does not hold for this subset. ◁

EXAMPLE 3.2.7. Let $A = \{0, 1\}$ and let $C = \{\mathtt{R}, \mathtt{G}, \mathtt{B}\}$. The following is the list of all functions with domain A and codomain C:

$$\begin{array}{lll} \big\{ (0, \mathtt{R}), (1, \mathtt{R}) \big\}, & \big\{ (0, \mathtt{R}), (1, \mathtt{G}) \big\}, & \big\{ (0, \mathtt{R}), (1, \mathtt{B}) \big\}, \\ \big\{ (0, \mathtt{G}), (1, \mathtt{R}) \big\}, & \big\{ (0, \mathtt{G}), (1, \mathtt{G}) \big\}, & \big\{ (0, \mathtt{G}), (1, \mathtt{B}) \big\}, \\ \big\{ (0, \mathtt{B}), (1, \mathtt{R}) \big\}, & \big\{ (0, \mathtt{B}), (1, \mathtt{G}) \big\}, & \big\{ (0, \mathtt{B}), (1, \mathtt{B}) \big\}. \end{array}$$

Thus, there are 3^2 functions with domain $\{0, 1\}$ and codomain $\{\mathtt{R}, \mathtt{G}, \mathtt{B}\}$. This can be a coloring exercise: In how many different ways one can color the integers 0 and

1 using three primary colors **red**, **green**, and **blue**? And the answer is:

$$\{0,1\},\ \{0,1\},\ \{0,1\},\ \{0,1\},\ \{0,1\},\ \{0,1\},\ \{0,1\},\ \{0,1\},\ \{0,1\}. \qquad \triangleleft$$

REMARK 3.2.8. Let A and B be any nonempty sets. Exercise 3.2.7 invites us to stretch our imagination a bit and think of a function $f : A \to B$ as a way of coloring each element of A using the "colors" available in B. In this view, an element $x \in A$ is colored by the "color" $f(x) \in B$. The fundamental properties of a function can now be rephrased as: every element of A is colored, and no element is assigned multiple colors. $\triangleleft$

3.2.2. New functions from old.

DEFINITION 3.2.9. Let A, B, C, and D be nonempty sets such that $B \subseteq C$, and let $f\colon A \to B$ and $g\colon C \to D$ be functions. The *composition* of f and g is the function $h\colon A \to D$ defined by

$$\forall x \in A \qquad h(x) = g\big(f(x)\big).$$

We denote the composition $h\colon A \to D$ by $h = g \circ f$.

DEFINITION 3.2.10. Let A, B, C, and D be nonempty sets such that $A \cap C = \emptyset$, and let $f\colon A \to B$ and $g\colon C \to D$ be functions. Then the function defined by

$$\forall x \in A \cup C \qquad h(x) = \begin{cases} f(x) & \text{if} \quad x \in A, \\ g(x) & \text{if} \quad x \in C, \end{cases}$$

is called a *piecewise defined* function with domain $A \cup C$ and codomain $B \cup D$.

The preceding two definitions relay on the standard functional notation $y = f(x)$ as an equivalent expression to $(x, y) \in f$. To be completely rigorous we would need to prove that such defined functions are truly functions in the sense of Definition 3.2.1. The formal statements corresponding to the above definitions and corresponding proofs are given in the propositions below.

PROPOSITION 3.2.11. *Let A and B be nonempty sets and let $f : A \to B$ be a function with domain A and codomain B. Let C and D be nonempty sets and let $g : C \to D$ be a function with domain C and codomain D. Assume that $B \subseteq C$. Then the set*

$$h = \Big\{(x, z) \in A \times D : \exists y \in B \ \ (x, y) \in f \ \wedge \ (y, z) \in g\Big\} \tag{3.2.1}$$

is a function with domain A and codomain D.

PROOF. Assume the following facts:

- ➢ A, B, C and D are nonempty sets,
- ➢ $f : A \to B$ is a function,
- ➢ $g : C \to D$ is a function,
- ➢ $B \subseteq C$.

It is very helpful to write down the specific **green statements** in the notation of formal logic:

TOTALf $\forall x \in A \ \ \exists y \in B$ such that $(x,y) \in f$.

UNIVALf $\forall (x_1, y_1) \in f \ \ \forall (x_2, y_2) \in f \quad x_1 = x_2 \Rightarrow y_1 = y_2$.

TOTALg $\forall y \in C \ \ \exists z \in D$ such that $(x,y) \in g$.

UNIVALg $\forall (y_1, z_1) \in g \ \ \forall (y_2, z_2) \in g \quad y_1 = y_2 \Rightarrow z_1 = z_2$.

G1 $B \subseteq C$.

G2 The definition of $h \subseteq A \times D$ in (3.2.1) is green.

We need to prove that $h : A \to D$ is a function. It is helpful to be specific what we need to prove:

TOTALh $\forall x \in A \ \ \exists z \in D$ such that $(x,z) \in h$.

UNIVALh $\forall (x_1, z_1) \in h \ \ \forall (x_2, z_2) \in h \quad x_1 = x_2 \Rightarrow z_1 = z_2$.

We first prove TOTALh. Since the statement TOTALh involves the universal quantifier, we prove it by universal generalization. Let $x \in A$ be arbitrary. By TOTALf, there exists $y \in B$ (label **G3**) such that $(x,y) \in f$ (label **G4**). By **G1** and **G3**, we deduce $y \in C$ (label **G5**). By **G5** and TOTALg, there exists $z \in D$ (label **G6**) such that $(y,z) \in g$ (label **G7**). By **G3**, **G4**, **G7** and **G2**, we deduce that $(x,z) \in h$. This completes the proof of TOTALh.

Now we prove UNIVALh. Since statement UNIVALh involves the universal quantifiers, we prove it by universal generalization. Let $(x_1, z_1) \in h$ (label **G8**) and $(x_2, z_2) \in h$ (label **G9**) be arbitrary.

To prove UNIVALh, we must prove

$$x_1 = x_2 \quad \Rightarrow \quad z_1 = z_2.$$

Assume $x_1 = x_2$ (label **G10**).

By **G8** and **G2**, there exists $y_1 \in B$ such that $(x_1, y_1) \in f$ (label **G11**) and $(y_1, z_1) \in g$ (label **G12**).

By **G9** and **G2**, there exists $y_2 \in B$ such that $(x_2, y_2) \in f$ (label **G13**) and $(y_2, z_2) \in g$ (label **G14**).

By **G10**, **G11**, **G13**, and UNIVALf, we deduce $y_1 = y_2$ (label **G15**).

By **G15**, **G12**, **G14**, and UNIVALg, we deduce $z_1 = z_2$. Thus, we have greenified the conclusion $z_1 = z_2$ in UNIVALh. This completes the proof of UNIVALh.

This completes the proof that h is a function. □

PROPOSITION 3.2.12. *Let A and B be nonempty sets and let $f : A \to B$ be a function with domain A and codomain B. Let C and D be nonempty sets and let $g : C \to D$ be a function with domain C and codomain D. Assume that B and C are disjoint. Then the set*

$$h = f \cup g \subseteq (A \cup C) \times (B \cup D) \tag{3.2.2}$$

is a function with domain $A \cup C$ and codomain $B \cup D$.

Only the beginning of the proof below uses color. The remainder is written to remain clear without it.

PROOF. Let A, B, C and D be nonempty sets and assume:

TOTALf $\forall x \in A \ \ \exists y \in B$ such that $(x,y) \in f$.

UNIVALf $\forall (x_1,y_1) \in f \ \ \forall (x_2,y_2) \in f \quad x_1 = x_2 \Rightarrow y_1 = y_2$.

TOTALg $\forall y \in C \ \ \exists z \in D$ such that $(x,y) \in g$.

UNIVALg $\forall (y_1,z_1) \in g \ \ \forall (y_2,z_2) \in g \quad y_1 = y_2 \Rightarrow z_1 = z_2$.

G1 $A \cap C = \emptyset$.

G2 The definition of $h \subseteq (A \cup C) \times (B \cup D)$ in (3.2.2) is green.

We need to prove:

TOTALh $\forall x \in A \cup C \ \ \exists y \in B \cup D$ such that $(x,y) \in h$.

UNIVALh $\forall (x_1,y_1) \in h \ \ \forall (x_2,y_2) \in h \quad x_1 = x_2 \Rightarrow y_1 = y_2$.

We first prove TOTALh. Let $x \in A \cup C$ be arbitrary. By definition of $A \cup C$ we have $x \in A$ or $x \in C$. We consider two cases.

Case 1. $x \in A$. By TOTALf, there exists $y \in B$ such that $(x,y) \in f$. Clearly $y \in B \cup D$ and $(x,y) \in f \cup g = h$. Thus, TOTALh is proved in this case.

Case 2. $x \in C$. By TOTALg, there exists $y \in D$ such that $(x,y) \in g$. Clearly $y \in B \cup D$ and $(x,y) \in f \cup g = h$. Thus, TOTALh is proved in this case as well.

Before moving to the next part, recall some facts. By Exercise 3.1.9 we have that **G1** implies that $A \subseteq C^c$ and $C \subseteq A^c$. Also, since $f \subseteq A \times B$, we have

$$(x,y) \in f \quad \Rightarrow \quad (x,y) \in A \times B \quad \Rightarrow \quad x \in A.$$

Passing to contrapositives we obtain

$$x \notin A \quad \Rightarrow \quad (x,y) \notin A \times B \quad \Rightarrow \quad (x,y) \notin f.$$

Now we prove UNIVALh. Let $(x_1,y_1) \in h$ and $(x_2,y_2) \in h$ be arbitrary. Assume $x_1 = x_2$. Since $x_1 = x_2 \in A \cup C$, we consider two cases.

Case 1. $x_1 = x_2 \in A$. Since $A \cap C = \emptyset$, we have that $A \subseteq C^c$. Therefore, $x_1 = x_2 \notin C$, and consequently, $(x_1,y_1) \notin g$ and $(x_2,y_2) \notin g$. Notice that

$$(x_1,y_1) \in h = f \cup g \ \wedge \ (x_1,y_1) \notin g \quad \Rightarrow \quad (x_1,y_1) \in f,$$

and

$$(x_2,y_2) \in h = f \cup g \ \wedge \ (x_2,y_2) \notin g \quad \Rightarrow \quad (x_2,y_2) \in f.$$

Hence, we have $(x_1,y_1) \in f$ and $(x_2,y_2) \in f$ and $x_1 = x_2$. By UNIVALf, we deduce $y_1 = y_2$. Thus, UNIVALh is proved in this case.

Case 2 is very similar to Case 1.

Case 2. $x_1 = x_2 \in C$. Since $A \cap C = \emptyset$, we have that $C \subseteq A^c$. Therefore, $x_1 = x_2 \notin A$, and consequently, $(x_1,y_1) \notin f$ and $(x_2,y_2) \notin f$. Notice that

$$(x_1,y_1) \in h = f \cup g \ \wedge \ (x_1,y_1) \notin f \quad \Rightarrow \quad (x_1,y_1) \in g,$$

and

$$(x_2,y_2) \in h = f \cup g \ \wedge \ (x_2,y_2) \notin f \quad \Rightarrow \quad (x_2,y_2) \in g.$$

Hence, we have $(x_1,y_1) \in g$ and $(x_2,y_2) \in g$ and $x_1 = x_2$. By UNIVALg, we deduce $y_1 = y_2$. Thus, UNIVALh is proved in this case as well.

Proof that h is a function is complete. $\square$

EXERCISE 3.2.13. Let A, B, and C be nonempty sets. Let $f : A \to B$ and $g : A \to C$ be functions. The function $H : A \to B \times C$ with domain A and codomain $B \times C$ defined as

$$\forall x \in A \quad H(x) = \big(f(x), g(x)\big)$$

is called the *product extension* of the pair of functions f and g.

Provide a formal definition of the product extension and prove that it is a function. ◁

3.2.3. Surjection, injection, bijection (via flip). Let A and B be two nonempty sets. The flip operation on the Cartesian product $A \times B$ is an interesting one to consider. This operation consists of swapping the elements in an ordered pair (x, y) to obtain (y, x). Thus, the flip of the ordered pair $(x, y) \in A \times B$ is the ordered pair $(y, x) \in B \times A$.

Let us now apply the flip to a function $f : A \to B$, that is apply the flip to the subset $f \subseteq A \times B$. Denote by $p \subseteq B \times A$ the result of the flip. That is

$$p = \big\{(y, x) \in B \times A : (x, y) \in f\big\}.$$

Or, in the notation of formal logic,

$$\forall x \in A \;\; \forall y \in B \qquad (y, x) \in p \quad \Leftrightarrow \quad (x, y) \in f. \tag{3.2.3}$$

Is the flip $p \subseteq B \times A$ a function?

We answer this question by invoking Definition 3.2.1:

The flip $p \subseteq B \times A$ of a function $f \subseteq A \times B$ is a function with domain B and codomain A if and only if the following two conditions are satisfied:

FLIP-TOT. $\forall y \in B \;\; \exists x \in A$ such that $(y, x) \in p$.

FLIP-UNI. $\forall (y_1, x_1) \in p \;\; \forall (y_2, x_2) \in p$ we have $y_1 = y_2 \Rightarrow x_1 = x_2$.

Since we began by considering the function $f \subseteq A \times B$, it is of interest to restate conditions FLIP-TOT and FLIP-UNI in the preceding framed box in terms of the function $f \subseteq A \times B$. To do this, we use the equivalence in (3.2.3).

In the framed box below, the restatement of FLIP-TOT is labeled SURJECT, and the restatement of FLIP-UNI is labeled INJECT. In the restatement INJECT of FLIP-UNI, the implication is replaced by its contrapositive. The reasoning behind the labels SURJECT and INJECT will become clear in Definition 3.2.14.

The conditions for the flip of a function to be a function.

The flip $p \subseteq B \times A$ of a function $f \subseteq A \times B$ is a function with domain B and codomain A if and only if the following two conditions are satisfied:

SURJECT. $\forall y \in B \;\; \exists x \in A$ such that $(x, y) \in f$.

INJECT. $\forall (x_1, y_1) \in f \;\; \forall (x_2, y_2) \in f$ we have $x_1 \neq x_2 \Rightarrow y_1 \neq y_2$.

Conditions SURJECT and INJECT are properties of the function $f \subseteq A \times B$. Recall the definition of the range of $f \subseteq A \times B$:

$$\operatorname{ran}(f) = \big\{y \in B : \exists x \in A \;\; (x, y) \in f\big\},$$

and consider SURJECT. Clearly SURJECT tells us that $\operatorname{ran}(f) = B$. That is, the range of $f : A \to B$ is the entire codomain.

Condition INJECT tells us that the function $f : A \to B$ assigns distinct values in the codomain to distinct values of the domain. Both of these properties are important features of $f : A \to B$. Therefore the following definitions.

DEFINITION 3.2.14. Let A and B be two nonempty sets. A function $f : A \to B$ is said to be a *surjection* if for every $y \in B$ there exists $x \in A$ such that $(x, y) \in f$. A function $f : A \to B$ is said to be an *injection* if for all $(x_1, y_1) \in f$, for all $(x_2, y_2) \in f$ we have $x_1 \neq x_2$ implies $y_1 \neq y_2$. A function $f : A \to B$ is said to be a *bijection* if it is both a surjection and an injection.

REMARK 3.2.15. (ON TERMINOLOGY) We could have stated briefly: A function $f : A \to B$ is called a *surjection* if it satisfies the condition SURJECT; alternatively we say f is *surjective*. A function $f : A \to B$ is called an *injection* if it satisfies condition INJECT; alternatively we say f is *injective*. Some mathematicians use the synonyms *onto* function for a surjection and *one-to-one* function for an injection. I recommend that you ignore these synonyms. $\lhd$

In the preceding framed box, two conditions for $f \subseteq A \times B$ to be a function are "hidden." I want to have all four conditions for both $f \subseteq A \times B$ and its flip $p \subseteq B \times A$ to be functions.

Four conditions for a subset of $A \times B$ and its flip to be functions.

Both $f \subseteq A \times B$ and its flip $p \subseteq B \times A$ are functions if and only if the following **four** conditions are satisfied:

TOTAL. $\forall x \in A \ \ \exists y \in B$ such that $(x, y) \in f$.
UNIVAL. $\forall (x_1, y_1) \in f \ \ \forall (x_2, y_2) \in f \quad x_1 = x_2 \Rightarrow y_1 = y_2$.
SURJECT. $\forall y \in B \ \ \exists x \in A$ such that $(x, y) \in f$.
INJECT. $\forall (x_1, y_1) \in f \ \ \forall (x_2, y_2) \in f \quad x_1 \neq x_2 \Rightarrow y_1 \neq y_2$.

Exploring the conditions in the preceding box.

Assume that all four conditions TOTAL, UNIVAL, SURJECT, and INJECT in the preceding framed box are satisfied. I want to verbalize what each of these conditions is saying. The condition SURJECT says that the range of $f : A \to B$ is the entire codomain B. That is, f is surjective. Condition INJECT says that $f : A \to B$ maps distinct inputs into distinct outputs. That is, f is injective. The condition TOTAL means that $f : A \to B$ is defined on the entire domain A. The condition UNIVAL means that no input has two distinct outputs. I can state each of these conditions in terms of the flip $p : B \to A$, which is now a function. The conditions read

- $\forall x \in A \ \ \exists y \in B$ such that $(y, x) \in p$. ____ p is surjective
- $\forall (y_1, x_1) \in p \ \forall (y_2, x_2) \in p \quad y_1 \neq y_2 \Rightarrow x_1 \neq x_2$. ____ p is injective
- $\forall y \in B \ \ \exists x \in A$ such that $(y, x) \in p$. ____ p is total
- $\forall (y_1, x_1) \in p \ \forall (y_2, x_2) \in p \quad y_1 = y_2 \Rightarrow x_1 = x_2$. ____ p is unival

Now the equivalent conditions in terms of $f : A \to B$, in the same order and with the names for f:

➢ $\forall x \in A \;\; \exists y \in B$ such that $(x,y) \in f$. ——— f is total
➢ $\forall (x_1,y_1) \in f \;\; \forall (x_2,y_2) \in f \quad x_1 = x_2 \Rightarrow y_1 = y_2$. —— f is unival
➢ $\forall y \in B \;\; \exists x \in A$ such that $(x,y) \in f$. ——— f is surjective
➢ $\forall (x_1,y_1) \in f \;\; \forall (x_2,y_2) \in f \quad x_1 \neq x_2 \Rightarrow y_1 \neq y_2$. —— f is injective

The information from the preceding framed box is collected in Table 3.2.1.

verbalize	$f : A \to B$		$p : B \to A$	verbalize
$f : A \to B$ is total	$\forall x \in A \;\; \exists y \in B$ such that $(x,y) \in f$	$\Leftrightarrow$	$\forall x \in A \;\; \exists y \in B$ such that $(y,x) \in g$	$p : B \to A$ is surjective
$f : A \to B$ is unival	$\forall (x_1,y_1)\in f \; \forall (x_2,y_2)\in f$ $x_1 = x_2 \Rightarrow y_1 = y_2$.	$\Leftrightarrow$	$\forall (y_1,x_1)\in p \; \forall (y_2,x_2)\in p$ $y_1 \neq y_2 \Rightarrow x_1 \neq x_2$.	$p : B \to A$ is injective
$f : A \to B$ is surjective	$\forall y \in B \;\; \exists x \in A$ such that $(x,y) \in f$	$\Leftrightarrow$	$\forall y \in B \;\; \exists x \in A$ such that $(y,x) \in p$	$p : B \to A$ is total
$f : A \to B$ is injective	$\forall (x_1,y_1)\in f \; \forall (x_2,y_2)\in f$ $x_1 \neq x_2 \Rightarrow y_1 \neq y_2$.	$\Leftrightarrow$	$\forall (y_1,x_1)\in p \; \forall (y_2,x_2)\in p$ $y_1 = y_2 \Rightarrow x_1 = x_2$.	$p : B \to A$ is unival

TABLE 3.2.1. Bijection f and its flip p

3.2.4. Invertible function, inverse.

DEFINITION 3.2.16. Let A and B be two nonempty sets. A function $f : A \to B$ is said to be *invertible* if there exists a function $g : B \to A$ such that

$$\forall x \in A \quad g\big(f(x)\big) = x \qquad \text{and} \qquad \forall y \in B \quad f\big(g(y)\big) = y. \tag{3.2.4}$$

The function $g : B \to A$ that satisfies the conditions in (3.2.4) is called the *inverse* of $f : A \to B$. The inverse of $f : A \to B$ is denoted by $f^{-1} : B \to A$.

Recall that we defined the identity functions on nonempty sets:

$$\mathrm{id}_A = \big\{(x,x) \in A \times A : x \in A\big\}, \qquad \mathrm{id}_B = \big\{(y,y) \in B \times B : y \in B\big\}.$$

Notice that (3.2.4) can be restated as

$$g \circ f = \mathrm{id}_A \qquad \text{and} \qquad f \circ g = \mathrm{id}_B\,. \tag{3.2.5}$$

THEOREM 3.2.17. *Let A and B be nonempty sets. A function $f : A \to B$ is invertible if and only if $f : A \to B$ is a bijection.*

PROOF. First we prove

$$\boxed{f : A \to B \;\text{ is a bijection}} \quad \Rightarrow \quad \boxed{f : A \to B \;\text{ is invertible}}$$

Assume that $\boxed{f : A \to B \text{ is a bijection}}$. Then, by Definition 3.2.14, f satisfies conditions SURJECT and INJECT introduced on page 36. Recall that conditions SURJECT and INJECT were introduced on page 36 to make sure that the flip of f, defined as

$$p = \big\{(y,x) \in B \times A : (x,y) \in f\big\},$$

is a function with domain B and codomain A.

Recall the definition in Proposition 3.2.11,

$$p \circ f = \big\{(x, z) \in A \times A : \exists y \in B \ \ (x, y) \in f \wedge (y, z) \in p\big\}.$$

Next we prove two set equalities $p \circ f = \mathrm{id}_A$ and $f \circ p$. First, $p \circ f = \mathrm{id}_A$. Since this is a set equality, we need to prove two inclusions $\mathrm{id}_A \subseteq p \circ f$ and $p \circ f \subseteq \mathrm{id}_A$. Let $x \in A$ be arbitrary. By the condition TOTAL for f there exists $y \in B$ such that $(x, y) \in f$. By definition of p we have $(y, x) \in p$. Since we have $y \in B$, $(x, y) \in f$, and $(y, x) \in p$, by definition of $p \circ f$ we have $(x, x) \in p \circ f$. This proves $\mathrm{id}_A \subseteq p \circ f$.

Next we prove $p \circ f \subseteq \mathrm{id}_A$. Let $(x, z) \in p \circ f$ be arbitrary. By definition of $p \circ f$, this means that there exists $y \in B$ such that $(x, y) \in f$ and $(y, z) \in p$. Since $(x, y) \in f$ is equivalent to $(y, x) \in p$, we have that $(y, x) \in p$ and $(y, z) \in p$. Since p is a function, that is by the condition FLIP-UNI, which is equivalent to INJECT, we deduce that $z = x$. Hence $(x, z) \in p \circ f$ implies $(x, z) = (x, x)$, proving $p \circ f \subseteq \mathrm{id}_A$.

Proof of the equality $f \circ p = \mathrm{id}_B$ is similar and is left as an exercise. Thus, we proved that the function p satisfies $p \circ f = \mathrm{id}_A$ and $f \circ p = \mathrm{id}_B$. By Definition 3.2.16, f is invertible, and its flip is its inverse.

Now we prove the converse

$$\boxed{f : A \to B \ \text{ is invertible}} \quad \Rightarrow \quad \boxed{f : A \to B \ \text{ is a bijection}}$$

Assume $\boxed{f : A \to B \text{ is invertible}}$. That is, assume that there exists a function $g : B \to A$ such that (3.2.4) holds.

We will prove that $f : A \to B$ is a surjection. Let $y \in B$ be arbitrary. Since $g : B \to A$ is a function, set $x = g(y)$. By the second identity in (3.2.4), we have $f(x) = f\big(g(y)\big) = y$. Thus, for every $y \in B$ there exists $x \in A$ such that $f(x) = y$, proving that f is a surjection.

We will prove that $f : A \to B$ is a injection. We need to prove that for arbitrary $x_1, x_2 \in A$ the following implication holds: $f(x_1) = f(x_2)$ implies $x_1 = x_2$. Assume, $f(x_1) = f(x_2)$ and apply g to both sides of the last equality. Then we get $g\big(f(x_1)\big) = g\big(f(x_2)\big)$. By the first identity in (3.2.4), we have $x_1 = x_2$, proving that f is an injection.

Since f is both, a surjection and an injection, we proved that f is a bijection. □

In the last proof, I mixed the graph notation for a function $f \subseteq A \times B$, where we write $(x, y) \in f$, with the traditional notation for a function $f : A \to B$, in which we write $y = f(x)$ as equivalent to $(x, y) \in f$. In the rest of these notes, we will predominantly use the traditional notation $f : A \to B$ to denote a function with domain A and codomain B and write $y = f(x)$ instead of $(x, y) \in f$.

In the next proposition, we prove that the composition of two bijections is a bijection.

THEOREM 3.2.18. *Let A, B and C be nonempty sets. Let $f : A \to B$ and $g : B \to C$ be bijections. Then $g \circ f : A \to C$ is a bijection and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

PROOF. Assume that $f : A \to B$ and $g : B \to C$ be bijections. By Theorem 3.2.17, $f : A \to B$ is invertible, with inverse $f^{-1} : B \to A$, and $g : B \to C$ is invertible, with inverse $g^{-1} : C \to B$. Denote by $h : C \to A$ the function $f^{-1} \circ g^{-1}$.

Let $x \in A$ be arbitrary, set $y = f(x)$ and calculate, starting with the definition of the composition $h \circ (g \circ f)$,

$$\begin{aligned}
\big(h \circ (g \circ f)\big)(x) &= h\big((g \circ f)(x)\big) \\
\boxed{\text{by the def. of } g \circ f} &= h\big(g(f(x))\big) \\
\boxed{g, h \text{ are funs and } y = f(x)} &= h\big(g(y)\big) \\
\boxed{h = f^{-1} \circ g^{-1}} &= (f^{-1} \circ g^{-1})\big(g(y)\big) \\
\boxed{\text{by the def. of } f^{-1} \circ g^{-1}} &= f^{-1}\Big(g^{-1}\big(g(y)\big)\Big) \\
\boxed{\text{Definition 3.2.16 applied to } g} &= f^{-1}(y) \\
\boxed{f^{-1} \text{ is a fun and } y = f(x)} &= f^{-1}\big(f(x)\big) \\
\boxed{\text{Definition 3.2.16 applied to } f} &= x.
\end{aligned}$$

Hence,

$$\forall x \in A \qquad \big(h \circ (g \circ f)\big)(x) = x, \tag{3.2.6}$$

that is $h \circ (g \circ f) = \mathrm{id}_A$.

Similarly, we can show that $(g \circ f) \circ h = \mathrm{id}_B$. Let $z \in C$ be arbitrary, set $y = g^{-1}(z)$, notice that $h(z) = f^{-1}\big(g^{-1}(z)\big) = f^{-1}(y)$, and calculate, starting with the definition of the composition $(g \circ f) \circ h$,

$$\begin{aligned}
\big((g \circ f) \circ h\big)(z) &= (g \circ f)\big(h(z)\big) \\
\boxed{g \circ f \text{ is a fun and } h(z) = f^{-1}(y)} &= (g \circ f)\big(f^{-1}(y)\big) \\
\boxed{\text{by the def. of } g \circ f} &= g\Big(f\big(f^{-1}(y)\big)\Big) \\
\boxed{\text{Definition 3.2.16 applied to } f} &= g(y) \\
\boxed{g \text{ is a fun and } y = g^{-1}(z)} &= g\big(g^{-1}(z)\big) \\
\boxed{\text{Definition 3.2.16 applied to } g} &= z.
\end{aligned}$$

Thus,

$$\forall z \in C \qquad \big((g \circ f) \circ h\big)(z) = z, \tag{3.2.7}$$

that is $(g \circ f) \circ h = \mathrm{id}_C$.

Since we have established (3.2.6) and (3.2.7), it follows from Definition 3.2.16 that $g \circ f$ is invertible, with $h = f^{-1} \circ g^{-1}$ as its inverse. Consequently, Theorem 3.2.17 yields that $g \circ f : A \to C$ is also a bijection. □

3.2.5. More on Surjections, Injections, and Bijections. Next we restate the definitions of a surjection, an injection, and a bijection using the traditional notation of a function.

A function $f : A \to B$ with domain A and codomain B is a *surjection* if for every $y \in B$ there exist $x \in A$ such that $y = f(x)$. Another way of saying that $f : A \to B$ is a surjection is to say that the range of f is equal to B.

What does it mean that $f : A \to B$ is not a surjection? A function $f : A \to B$ is not a surjection if there exists $y \in B$ such that for all $x \in A$ we have $f(x) \neq y$.

A function $f : A \to B$ is an *injection* if, for all $x_1, x_2 \in A$, we have that $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$. Note that the contrapositive of this implication is: for all $x_1, x_2 \in A$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$.

A function $f : A \to B$ is not an injection if there exist $x_1, x_2 \in A$ such that $x_1 \neq x_2$ and $f(x_1) = f(x_2)$.

A function $f : A \to B$ which is both an injection and a surjection is called *bijection*.

Recall that Theorem 3.2.17 states that a function $f : A \to B$ is a bijection if and only there exists a function $g : B \to A$ such that (3.2.4) holds. Thus, if we need to prove that a given function $f : A \to B$ is a bijection, an efficient way to accomplish this is to discover a function $g : B \to A$ for which we can verify (3.2.4). This, of course, amounts finding the inverse of f.

EXAMPLE 3.2.19. There are only three functions listed in Example 3.2.7 which are not injections. Find them! There are no bijections. There are no surjections. ◁

In mathematics, functions are often defined by formulas without specific names. Therefore, we also use the notation $f : x \mapsto f(x)$, where $x \in A$. For example, the notation $x \mapsto x^2$, where $x \in \mathbb{R}$, denotes the square function defined on $\mathbb{R}$ without specifying its codomain.

EXAMPLE 3.2.20. In this mixture of an example and exercise, we consider the squaring operation $x \mapsto x^2$ with $x \in \mathbb{R}$. We explore how properties of square being a function, surjective, injective and bijective depend on the domain and codomain. We will consider the following sets as domains and codomains: the set $\mathbb{R}$ of real numbers, the set $\mathbb{Q}$ of rational numbers, the set $\mathbb{R}_{\geq 0}$ of all nonnegative real numbers, the set $\mathbb{Q}_{\geq 0}$ of all nonnegative rational numbers.

In Table 3.2.2, I list various domains and codomains and ask you to fill in the blank boxes. Under each property, there are two boxes. The first box is for answering *Yes* or *No*, indicating whether the property listed in the header holds. The second box is for indicating the level of mathematical sophistication needed to justify your answer in the first box. For this, you may use a digit from the set $\{1, 2, 3, 4, 5\}$; however, to be thorough, you should also provide a brief explanation in English.

Some rigorous justifications for the answers in this table will be developed later in these notes. Ideally, at this stage, you should identify those answers you can justify now. I will refer back to Table 3.2.2 as we progress. ◁

3.2.6. Cantor's theorem. Let S be a nonempty set. In this subsection we explore the power set of S, the family of all subsets of S, denoted by $\mathcal{P}(S)$. We first prove that there exists a bijection with domain $\mathcal{P}(S)$ and range $\{0, 1\}^S$, the set of all functions with domain S and codomain $\{0, 1\}$.

DEFINITION 3.2.21. Let S be a nonempty set. For a subset $A \subseteq S$ we define the *indicator function of* A, denoted by $\chi_A : S \to \{0, 1\}$, to be the function

$$\chi_A(x) = \begin{cases} 1 & \text{if} \quad x \in A, \\ 0 & \text{if} \quad x \in S \setminus A. \end{cases}$$

A synonym for the term 'indicator function' is 'characteristic function'.

dom	codom	function		surjection		injection		bijection	
$\mathbb{Z}$	$\mathbb{Z}$	Yes	1	Y N	level	Y N	level	Y N	level
$\mathbb{Z}$	$\mathbb{Z}_{\geq 0}$	Yes	2	Y N	level	Y N	level	Y N	level
$\mathbb{Z}_{\geq 0}$	$\mathbb{Z}_{\geq 0}$	Yes	2	Y N	level	Y N	level	Y N	level
$\mathbb{Q}$	$\mathbb{Q}$	Yes	1	Y N	level	Y N	level	Y N	level
$\mathbb{Q}$	$\mathbb{Q}_{\geq 0}$	Yes	2	Y N	level	Y N	level	Y N	level
$\mathbb{Q}_{\geq 0}$	$\mathbb{Q}_{\geq 0}$	Yes	2	Y N	level	Y N	level	Y N	level
$\mathbb{R}$	$\mathbb{R}$	Yes	1	Y N	level	Y N	level	Y N	level
$\mathbb{R}$	$\mathbb{R}_{\geq 0}$	Yes	2	Y N	level	Y N	level	Y N	level
$\mathbb{R}_{\geq 0}$	$\mathbb{R}_{\geq 0}$	Yes	2	Y N	level	Y N	level	Y N	level

TABLE 3.2.2. Exploring square functions

Recall that for nonempty sets X and Y by Y^X, we denote the set of all functions with domain X and codomain Y, see Definition 3.2.4. Thus, $\{0,1\}^S$ is the set of all functions with domain S and codomain $\{0,1\}$.

PROPOSITION 3.2.22. *Let S be a nonempty set. The function $\Phi : \mathcal{P}(S) \to \{0,1\}^S$ defined by*

$$\forall A \in \mathcal{P}(S) \qquad \Phi(A) = \chi_A$$

is a bijection.

PROOF. Define the function $\Psi : \{0,1\}^S \to \mathcal{P}(S)$ by

$$\forall f \in \{0,1\}^S \qquad \Psi(f) = \{x \in S : f(x) = 1\}.$$

Now we prove

$$\forall A \in \mathcal{P}(S) \qquad \Psi\big(\Phi(A)\big) = A.$$

Let $A \subseteq S$ be arbitrary. Then

$$\Psi\big(\Phi(A)\big) = \Psi(\chi_A) = \big\{x \in S : \chi_A(x) = 1\big\} = A,$$

where the first equality sign holds by the definition of Φ, the second by the definition of Φ, and the third by the definition of χ_A, which states that for $x \in S$ $\chi_A(x) = 1$ if and only if $x \in A$.

Now we prove

$$\forall f \in \{0,1\}^S \qquad \Phi\big(\Psi(f)\big) = f.$$

Let $f \in \{0,1\}^S$ be arbitrary. Set

$$A = \Psi(f) = \{x \in S : f(x) = 1\}.$$

Then,

$$\Phi\big(\Psi(f)\big) = \Phi(A) = \chi_A = f.$$

where the first equality follows from the definition of A, the second by the definition of Φ, and the third again from the definition of A, and the fact that $f : S \to \{0,1\}$

has only two values 0 and 1, yielding

$$\big(f(x) = 1 \;\Leftrightarrow\; x \in A\big) \wedge \big(f(x) = 0 \;\Leftrightarrow\; x \in S\backslash A\big).$$

Since we have established that

$$\forall A \in \mathcal{P}(S) \quad \Psi\big(\Phi(A)\big) = A \qquad \text{and} \qquad \forall f \in \{0,1\}^S \qquad \Phi\big(\Psi(f)\big) = f,$$

it follows from Definition 3.2.16 that Φ is invertible, with Ψ as its inverse. Consequently, by Theorem 3.2.17, Φ is a bijection. □

Exercise 3.2.23. Let A, B and C be nonempty sets. Let $f : A \to B$ and $g : B \to C$ be surjections. Prove that $g \circ f : A \to C$ is a surjection. ◁

Solution. Assume that $f : A \to B$ and $g : B \to C$ are surjections. To prove that $g \circ f : A \to C$ is a surjection we have to prove that for every $z \in C$ there exists $x \in A$ such that $g(f(x)) = z$. Let $z \in C$ be arbitrary. Since $g : B \to C$ is a surjection, there exists $y \in B$ such that $g(y) = z$. Since $y \in B$ and as $f : A \to B$ is a surjection, there exists $x \in A$ such that $f(x) = y$. Now we have $g(f(x)) = g(y) = z$, where the first equality holds since $f(x) = y$ and g is a function, and the second since $g(y) = z$. □

Exercise 3.2.24. Let A and B be nonempty sets. Consider the following statement: There exists a surjective function $f : A \to B$.

(a) Translate the above statement from mathematical English into the notation of formal logic, using sets and quantifiers.
(b) State the negation of the given statement, both in mathematical English and in the notation of formal logic, using sets and quantifiers. ◁

Solution. It is convenient to use the notation B^A for the set of all functions with domain A and codomain B. Then the given statement in the notation of formal logic reads:

$$\exists f \in B^A \quad \forall y \in B \quad \exists x \in A \quad y = f(x).$$

The negation in English is: For an arbitrary function $g : A \to B$, g is not a surjection. But the statement "g is not a surjection" is itself a negation which means: There exists $b \in B$ such that for all $x \in A$ we have $g(x) \neq b$. Hence the negation of the given statement in the notation of formal logic is:

$$\forall g \in B^A \quad \exists b \in B \quad \forall x \in A \quad f(x) \neq b.$$

It is important to note that $b \in B$ in the last statement depends on g. In a proof of the last statement one would start from an arbitrary g and then try to construct $b \in B$ with the desired property. □

Theorem 3.2.25 (Cantor's Theorem). *Let S be a nonempty set. Then there is no surjection with domain S and codomain $\mathcal{P}(S)$.*

Proof. We have to prove

$$\forall \Theta : S \to \mathcal{P}(S) \quad \exists A \subseteq S \quad \forall x \in S \quad \Theta(x) \neq A.$$

Let $\Theta : S \to \mathcal{P}(S)$ be an arbitrary function. We will prove that Θ is not a surjection. We need to discover a subset A of S such that

$$\forall x \in S \quad \Theta(x) \neq A.$$

The definition of such subset A is truly ingenious, due to Georg Cantor,

$$A = \{s \in S : s \notin \Theta(s)\}.$$

Now we need to prove that for all $x \in S$ we have $\Theta(x) \neq A$. Let $x \in S$ be arbitrary. We consider two cases: **Case 1**: $x \in A$ and **Case 2**: $x \notin A$. First consider **Case 1** in which we assume $x \in A$. In this case $x \notin \Theta(x)$. Hence, we have $x \in A$ and $x \notin \Theta(x)$. Consequently $A \nsubseteq \Theta(x)$. Therefore $\Theta(x) \neq A$, in this case.

Now consider **Case 2**. In this case $x \notin A$. By the definition of A we have $x \in \Theta(x)$. Thus, $x \notin A$ and $x \in \Theta(x)$. Consequently, $\Theta(x) \nsubseteq A$. Therefore, $\Theta(x) \neq A$, in this case as well.

In both cases, $\Theta(x) \neq A$. Since $x \in S$ was arbitrary, we have proved that for all $x \in S$ we have $\Theta(x) \neq A$. Therefore, the subset A of S is not in the range of Θ. Since $\Theta : S \to \mathcal{P}(S)$ was an arbitrary function with domain S and codomain $\mathcal{P}(S)$, we have proved that there does not exist a surjection from S to $\mathcal{P}(S)$. □

EXERCISE 3.2.26. Prove that the function $f : \mathbb{R}\setminus\{1\} \to \mathbb{R}\setminus\{1\}$, defined by $f(x) = \dfrac{x}{x-1}$ for all $x \in \mathbb{R}\setminus\{1\}$, is a bijection. ◁

EXERCISE 3.2.27. Let A, B, and C be nonempty sets and let $f : A \to B$ and $g : A \to C$ be functions. We define the *product extension* of the pair (f, g) to be the function $h : A \to B \times C$ defined as follows

$$\forall x \in A \quad \text{we define} \quad h(x) = \big(f(x), g(x)\big).$$

(i) Show that if f or g is injective, then H is also injective.
(ii) Is it true that if f and g are surjective, then H is surjective? ◁

EXERCISE 3.2.28. Let A, B and C be nonempty sets. Let $f : A \to B$ and $g : B \to C$ be injections. Prove that $g \circ f : A \to C$ is an injection. ◁

EXERCISE 3.2.29. Let A, B and C be nonempty sets. Let $f : A \to B$, $g : B \to C$. Prove that if $g \circ f$ is an injection, then f is an injection. ◁

EXERCISE 3.2.30. Let A, B and C be nonempty sets. Let $f : A \to B$ and $g : B \to C$ be surjections. Prove that $g \circ f : A \to C$ is a surjection. ◁

EXERCISE 3.2.31. Let A, B and C be nonempty sets and let $f : A \to B$, $g : B \to C$. Prove that if $g \circ f$ is a surjection, then g is a surjection. ◁

EXERCISE 3.2.32. Let A, B and C be nonempty sets and let $f : A \to B$, $g : B \to C$ and $h : C \to A$ be three functions. Prove that if any two of the functions $h \circ g \circ f$, $g \circ f \circ h$, $f \circ h \circ g$ are injections and the third is a surjection, or if any two of them are surjections and the third is an injection, then f, g, and h are bijections. ◁

3.3. Cardinality of sets

Counting is the most fundamental mathematical operation used in everyday life. It allows us to compare the numbers of elements in different sets and identify sets with the same number of elements. In everyday life, this concept is typically applied to finite sets. However, in mathematics, we seek to extend the concept of having the 'same number of elements' to infinite sets. The concept of bijection plays a crucial role in achieving this extension. Instead of stating that two sets have the same number of elements, we express this equivalence using the term 'same cardinality'.

DEFINITION 3.3.1. Two nonempty sets A and B have the same *cardinality* if there exists a bijection $f\colon A \to B$. The fact that two sets have the same cardinality is denoted by $\operatorname{card}(A) = \operatorname{card}(B)$ or $A \sim B$.

PROPOSITION 3.3.2. *The equality of cardinalities is reflexive, symmetric and transitive. That is, for arbitrary nonempty sets A, B, and C we have*

(i) $\operatorname{card}(A) = \operatorname{card}(A)$.
(ii) $\operatorname{card}(A) = \operatorname{card}(B)$ *if and only if* $\operatorname{card}(B) = \operatorname{card}(A)$.
(iii) *If* $\operatorname{card}(A) = \operatorname{card}(B)$ *and* $\operatorname{card}(B) = \operatorname{card}(C)$, *then* $\operatorname{card}(A) = \operatorname{card}(C)$.

PROOF. (i) Since $\mathrm{id}_A : A \to A$ is a bijection, we have $\operatorname{card}(A) = \operatorname{card}(A)$.

(ii) Assume $\operatorname{card}(A) = \operatorname{card}(B)$. Then there exists a bijection $f : A \to B$. By Theorem 3.2.17, f is invertible, and its inverse $f^{-1} : B \to A$ is a bijection. Therefore, $\operatorname{card}(B) = \operatorname{card}(A)$. The converse is proved similarly.

(iii) Assume $\operatorname{card}(A) = \operatorname{card}(B)$ and $\operatorname{card}(B) = \operatorname{card}(C)$. Then there exist bijections $f : A \to B$ and $g : B \to C$. By Theorem 3.2.18, the composition $g \circ f : A \to C$ is a bijection. Therefore, $\operatorname{card}(A) = \operatorname{card}(C)$. □

PROPOSITION 3.3.3. *Let A, B, C, D be nonempty sets such that*

$$A \cap C = \emptyset, \qquad B \cap D = \emptyset$$

and

$$\operatorname{card}(A) = \operatorname{card}(B), \qquad \operatorname{card}(C) = \operatorname{card}(D).$$

Then $\operatorname{card}(A \cup C) = \operatorname{card}(B \cup D)$.

PROOF. Assume that A, B, C, D be nonempty sets such that $A \cap C = \emptyset$, $B \cap D = \emptyset$, $\operatorname{card}(A) = \operatorname{card}(B)$, and $\operatorname{card}(C) = \operatorname{card}(D)$. Then there exist bijections $f : A \to B$ and $g : C \to D$.

By Theorem 3.2.17, $f : A \to B$ and $g : C \to D$ are invertible. Let $f^{-1} : B \to A$ and $g^{-1} : D \to C$ be their inverses. Let

$$h : A \cup C \to B \cup D, \qquad \forall x \in A \cup C \qquad h(x) = \begin{cases} f(x) & \text{if} \quad x \in A, \\ g(x) & \text{if} \quad x \in C \end{cases}$$

be a piecewise defined function as in Definition 3.2.10, see also Proposition 3.2.12. Since $B \cap D = \emptyset$, we can define the piecewise function

$$k : B \cup D \to A \cup C, \qquad \forall y \in B \cup D \qquad k(y) = \begin{cases} f^{-1}(y) & \text{if} \quad y \in B, \\ g^{-1}(y) & \text{if} \quad y \in D. \end{cases}$$

The following two functional equalities hold: $k \circ h = \mathrm{id}_{A\cup C}$ and $h \circ k = \mathrm{id}_{B\cup D}$. We leave proofs of these equalities to the reader. With these equalise proven, Theorem 3.2.17 yields that $h : A \cup C \to B \cup D$ is a bijection. Therefore,

$$\mathrm{card}(A \cup C) = \mathrm{card}(B \cup D). \qquad \square$$

DEFINITION 3.3.4. Let A and B be sets. If there exists a subset $C \subseteq B$ such that $\mathrm{card}(A) = \mathrm{card}(C)$, then we write $\mathrm{card}(A) \leq \mathrm{card}(B)$ and say that the *cardinality of A is less than or equal to the cardinality of B*. If $\mathrm{card}(A) \neq \mathrm{card}(B)$ and $\mathrm{card}(A) \leq \mathrm{card}(B)$, then we write $\mathrm{card}(A) < \mathrm{card}(B)$.

A fundamental result about cardinalities is the following theorem of Cantor.

THEOREM 3.3.5 (Cantor's Theorem). *Let S be a nonempty set. Then*

$$\mathrm{card}(S) < \mathrm{card}\big(\mathcal{P}(S)\big).$$

That is, there exists an injection $F : S \to \mathcal{P}(S)$ and there is no bijection from S to $\mathcal{P}(S)$.

PROOF. Consider the function $\Psi : S \to \mathcal{P}(S)$ defined by

$$\forall x \in S \qquad \Psi(x) = \{x\}.$$

The function $\Psi : S \to \mathcal{P}(S)$ is an injection. To prove this claim, let $x_1, x_2 \in S$ be such that $x_1 \neq x_2$. Then, $\Psi(x_1) \neq \Psi(x_1)$, since $x_1 \in \Psi(x_1)$ and $x_1 \notin \Psi(x_2)$.

Denote by $\mathcal{R}$ the range of Ψ. That is,

$$\mathcal{R} = \Big\{ \{x\} \in \mathcal{P}(S) : x \in S \Big\}.$$

Then the function $\Psi : S \to \mathcal{R}$ is a bijection. Therefore, $\mathrm{card}(S) = \mathrm{card}(\mathcal{R})$. Since $\mathcal{R} \subseteq \mathcal{P}(S)$, this proves that $\mathrm{card}(S) \leq \mathrm{card}\big(\mathcal{P}(S)\big)$.

In Theorem 3.2.25 we proved that there is no surjection with domain S and codomain $\mathcal{P}(S)$. Consequently, there is no bijection with domain S and codomain $\mathcal{P}(S)$. Hence, $\mathrm{card}(S) \neq \mathrm{card}\big(\mathcal{P}(S)\big)$. Therefore, $\mathrm{card}(S) < \mathrm{card}\big(\mathcal{P}(S)\big)$. $\square$

EXERCISE 3.3.6. Let A and B be nonempty sets. Prove that $\mathrm{card}(A) \leq \mathrm{card}(B)$ if and only if there exists an injection $f : A \to B$. $\triangleleft$

EXERCISE 3.3.7. Let A and B be nonempty sets. Prove that $\mathrm{card}(A) \leq \mathrm{card}(B)$ if and only if there exists a surjection $f : B \to A$. $\triangleleft$

EXERCISE 3.3.8. Let A, B, C be nonempty sets. Prove that $\mathrm{card}(A) \leq \mathrm{card}(B)$ and $\mathrm{card}(B) \leq \mathrm{card}(C)$ implies that $\mathrm{card}(A) \leq \mathrm{card}(C)$. $\triangleleft$

REMARK 3.3.9. The following natural implication holds:

$$\mathrm{card}(A) \leq \mathrm{card}(B) \quad \text{and} \quad \mathrm{card}(B) \leq \mathrm{card}(A) \quad \Rightarrow \quad \mathrm{card}(A) = \mathrm{card}(B).$$

A proof of this implication is not easy. This implication is known as the Cantor-Bernstein-Schröder theorem, which we state below in an equivalent form. $\triangleleft$

THEOREM 3.3.10 (Cantor-Bernstein-Schröder). *Let A and B be nonempty sets. If there exist injections $f : A \to B$ and $g : B \to A$, then there exists a bijection $h : A \to B$.*

A proof of this theorem requires more background knowledge than is currently available to us, so we will postpone it.

CHAPTER 4

The set $\mathbb{R}$ of real numbers

4.1. Axioms for the set $\mathbb{R}$ of real numbers

All concepts that we will study in this course have their roots in the set of real numbers. We assume that you are familiar with some basic properties of the real numbers $\mathbb{R}$ and of its subsets $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$. However, in order to clarify exactly what we need to know about $\mathbb{R}$, we set down its basic properties (called axioms) and some of their consequences. Below are sixteen axioms of $\mathbb{R}$.

The set of real numbers is a nonempty set $\mathbb{R}$ that satisfies the following axioms:

Axiom 1 (**AE**: Addition Exists). There exists a function $+ : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, called *addition*. For $a, b \in \mathbb{R}$, the value of $+$ at the pair (a, b), called the *sum* of a and b, is denoted by $a + b$.

Axiom 2 (**AA**: Addition is Associative). For all $a, b, c \in \mathbb{R}$ we have $a + (b + c) = (a + b) + c$.

Axiom 3 (**AC**: Addition is Commutative). For all $a, b \in \mathbb{R}$ we have $a + b = b + a$.

Axiom 4 (**AZ**: Addition has 0 (Zero)). There exists $0 \in \mathbb{R}$ such that for all $a \in \mathbb{R}$ we have $a + 0 = a$.

Axiom 5 (**AO**: Addition has Opposites). For every $a \in \mathbb{R}$ there exists $x \in \mathbb{R}$ such that $a + x = 0$.

Axiom 6 (**ME**: Multiplication Exists). There exists a function $\cdot : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, called *multiplication*. For $a, b \in \mathbb{R}$, the value of $\cdot$ at the pair (a, b), called the *product* of a and b, is denoted by $a \cdot b$ or simply ab.

Axiom 7 (**MA**: Multiplication is Associative). For all $a, b, c \in \mathbb{R}$ we have $a(bc) = (ab)c$.

Axiom 8 (**MC**: Multiplication is Commutative). For all $a, b \in \mathbb{R}$ we have $ab = ba$.

Axiom 9 (**MO**: Multiplication has 1 (One)). There exists $1 \in \mathbb{R}$ such that $1 \neq 0$ and for all $a \in \mathbb{R}$ we have $1 \cdot a = a$

Axiom 10 (**MR**: Multiplication has Reciprocals). For every $a \in \mathbb{R}\setminus \{0\}$ there exists $x \in \mathbb{R}$ such that $ax = 1$.

Axiom 11 (**DL**: Distributive Law). For all $a, b, c \in \mathbb{R}$ we have $a(b + c) = ab + ac$.

AXIOM 12 (**OE**: Order Exists). For all $a, b \in \mathbb{R}$ such that $a \neq b$ we have $a < b$ exclusive or $b < a$. (The symbol $a \leq b$ stands for $a < b$ exclusive or $a = b$.)

AXIOM 13 (**OT**: Order is Transitive). For all $a, b, c \in \mathbb{R}$ we have that $a < b$ and $b < c$ implies $a < c$.

AXIOM 14 (**OA**: Order respects Addition). For all $a, b, c \in \mathbb{R}$ we have that $a < b$ implies $a + c < b + c$.

AXIOM 15 (**OM**: Order respects Multiplication). For all $a, b, c \in \mathbb{R}$ we have that $a < b$ and $0 < c$ implies $ac < bc$.

AXIOM 16 (**CA**: Completeness Axiom). If A and B are nonempty subsets of $\mathbb{R}$ such that for all $a \in A$ and for all $b \in B$ we have $a \leq b$, then there exists $c \in \mathbb{R}$ such that for all $a \in A$ and for all $b \in B$ we have $a \leq c$ and $c \leq b$.

The end of axioms

All statements about real numbers that are studied in beginning mathematical courses can be deduced from these sixteen axioms.

The formulation of the **Completeness Axiom** given as **Axiom 16** is not standard. This formulation is from the book *Mathematical analysis* by Vladimir Zorich, published by Springer in 2004. Zorich's formulation is easier to state and it is equivalent to the standard formulation of the Completeness Axiom. The Completeness Axiom in the notation of formal logic reads as follows:

AXIOM 16 (**CA**). $A \subset \mathbb{R}, \ A \neq \emptyset, \quad B \subset \mathbb{R}, \ B \neq \emptyset,$

$$\forall a \in A \ \forall b \in B \qquad a \leq b \quad \Rightarrow \quad \exists c \in \mathbb{R} \ \forall a \in A \ \forall b \in B \quad a \leq c \wedge c \leq b$$

4.1.1. Axioms of a field. The first eleven axioms relate to the addition and multiplication in $\mathbb{R}$. These axioms are called the axioms of a field. A triple of a nonempty set, an addition operation written as $a+b$, and a multiplication operation written as $a \cdot b$, which satisfy Axioms 1 through 11 is called a *field*.

Notice that the only specific real numbers mentioned in the axioms are 0 and 1.

EXAMPLE 4.1.1. Consider the set $\{0, 1\}$. Define functions $\oplus$ (a special addition instead of $+$) and $\odot$ (a special multiplication instead of $\cdot$) as follows

$$0 \oplus 0 = 1 \oplus 1 = 0, \ 0 \oplus 1 = 1 \oplus 0 = 1 \quad \text{and} \quad 0 \odot 0 = 0 \odot 1 = 1 \odot 0 = 0, \ 1 \odot 1 = 1.$$

The set $\{0, 1\}$ with the operations $\oplus$ and $\odot$ satisfies all Axioms 1 through 11. In the language of algebra, $\{0, 1\}$ with operations $\oplus$ and $\odot$ is a field. $\triangleleft$

Axioms AA and MA are called *associative laws* and Axioms AC and MC are *commutative laws*. Axiom DL is the *distributive law*; this law justifies 'factorization' and 'multiplying out' which you learned in college algebra. The basic algebraic properties of $\mathbb{R}$ can be proved solely on the basis of the field axioms. We illustrate this claim by the following exercise.

THEOREM 4.1.2 (Uniqueness Theorem). *The following uniqueness statements hold:*

(i) *Recall Axiom* AZ*: There exists* $0 \in \mathbb{R}$ *such that for all* $a \in \mathbb{R}$ *we have* $a + 0 = a$. *Such* $0 \in \mathbb{R}$ *is unique. That is, if for* $z \in \mathbb{R}$ *we have* $a + z = a$ *that for all* $a \in \mathbb{R}$, *then* $z = 0$.

(ii) *For all* $a, x, y \in \mathbb{R}$ *the following implication holds: If* $a+x = 0$ *and* $a + y = 0$, *then* $x = y$.

(iii) *Recall Axiom* MO*: There exists* $1 \in \mathbb{R}$ *with* $1 \neq 0$ *such that for all* $a \in \mathbb{R}$ *we have* $a \cdot 1 = a$. *Such* $1 \in \mathbb{R}$ *is unique. That is, if for* $z \in \mathbb{R}$ *we have* $a \cdot z = a$ *for all* $a \in \mathbb{R}$, *then* $z = 1$.

(iv) *For all* $a, x, y \in \mathbb{R}$ *with* $a \neq 0$, *the following implication holds: If* $a \cdot x = 1$ *and* $a \cdot y = 1$, *then* $x = y$.

PROOF. (i) Suppose that 0 and z are elements of $\mathbb{R}$ such that for all $a \in \mathbb{R}$, $a + 0 = a$ and $a + z = a$. Taking $a = 0$, we obtain $0 + z = 0$. By Axiom AC (commutativity of addition), $0+z = z+0$. By Axiom AZ (existence of 0), $z+0 = z$. Thus, $z = z + 0 = 0 + z = 0$. Therefore, the additive identity is unique.

(ii) Let $a, x, y \in \mathbb{R}$ be arbitrary and assume that $a + x = 0$ and $a + y = 0$. Then the following equalities hold, the first one being a cosequence of Axiom AZ,

$$
\begin{aligned}
x &= x + 0 \\
\boxed{\text{by assumption } a + y = 0} &= x + (a + y) \\
\boxed{\text{by Axiom AA}} &= (x + a) + y \\
\boxed{\text{by Axiom AC}} &= (a + x) + y \\
\boxed{\text{by assumption } a + x = 0} &= 0 + y \\
\boxed{\text{by Axiom AZ}} &= y.
\end{aligned}
$$

(iii) Suppose that 1 and z are elements of $\mathbb{R}$ such that for all $a \in \mathbb{R}$, $a \cdot 1 = a$ and $a \cdot z = a$. Taking $a = 1$, we obtain $1 \cdot z = 1$. By Axiom MC (commutativity of multiplication), $1 \cdot z = z \cdot 1$. By Axiom MO (existence of 1), $z \cdot 1 = z$. Thus, $z = 1$. Therefore, the multiplicative identity is unique.

(iv) Let $a, x, y \in \mathbb{R}$ with $a \neq 0$ and suppose that $a \cdot x = 1$ and $a \cdot y = 1$. Then the following equalities hold:

$$
\begin{aligned}
x &= 1 \cdot x \\
\boxed{\text{Why?}} &= x \cdot 1 \\
\boxed{\text{Why?}} &= x(ay) \\
\boxed{\text{Why?}} &= (xa)y \\
\boxed{\text{Why?}} &= (ax)y \\
\boxed{\text{Why?}} &= 1 \cdot y \\
\boxed{\text{Why?}} &= y.
\end{aligned}
$$

□

In Theorem 4.1.2, we proved four uniqueness statements: in (i), the uniqueness of the additive identity; in (ii), the uniqueness of the additive inverse; in (iii), the uniqueness of the multiplicative identity; and in (iv), the uniqueness of the multiplicative inverse. This theorem justifies the following definition.

DEFINITION 4.1.3. The element $0 \in \mathbb{R}$ which appears in Axiom AZ, and whose uniqueness is proved in Theorem 4.1.2(i), is called the *zero* in $\mathbb{R}$. Given $a \in \mathbb{R}$, the unique solution of the equation in Axiom AO is denoted by $-a$ and is called the *opposite* of a. The element $1 \in \mathbb{R}$ which appears in Axiom MO, and whose uniqueness is proved in Theorem 4.1.2(iii), is called the *number one*. Given $a \in \mathbb{R} \setminus \{0\}$, the unique solution of the equation in Axiom MR is denoted by $\dfrac{1}{a}$ and is called the *reciprocal* of a.

Let $a, b \in \mathbb{R}$. Instead of $a + (-b)$ we write $a - b$. With $b \neq 0$, instead of $a\dfrac{1}{b}$ we write $\dfrac{a}{b}$ or a/b or ab^{-1}.

The theorem below includes fourteen items. All the items are proved using the axioms of a field, that is Axioms 1 through 11. The first four items rely solely on the five addition axioms. Items (v) and (vi) use only the multiplication axioms. The remaining items depend on all eleven axioms of a field. Item (vii) presents the "if" direction of item (viii) and is stated separately for emphasis.

THEOREM 4.1.4. *The following statements hold:*

(i) $-0 = 0$.

(ii) *For all* $a, b, c \in \mathbb{R}$ *we have:* $a + c = b + c$ *implies* $a = b$.

(iii) *For all* $a \in \mathbb{R}$ *we have* $-(-a) = a$.

(iv) *For all* $a, b \in \mathbb{R}$ *we have* $-(a + b) = (-a) + (-b)$.

(v) $\dfrac{1}{1} = 1$.

(vi) *For all* $a, b, c \in \mathbb{R}$ *we have:* $ac = bc$ *and* $c \neq 0$ *implies* $a = b$.

(vii) *For all* $a \in \mathbb{R}$ *we have* $a \cdot 0 = 0$.

(viii) *For all* $a, b \in \mathbb{R}$ *we have:* $ab = 0$ *if and only if* $a = 0$ *or* $b = 0$.

(ix) *For all* $a \in \mathbb{R}\setminus\{0\}$ *we have:* $\dfrac{1}{a} \neq 0$ *and* $\dfrac{1}{\frac{1}{a}} = a$.

(x) *For all* $a, b \in \mathbb{R} \setminus \{0\}$ *we have* $\dfrac{1}{ab} = \dfrac{1}{a}\dfrac{1}{b}$.

(xi) *For all* $a, c \in \mathbb{R}$ *and all* $b, d \in \mathbb{R} \setminus \{0\}$ *we have* $\dfrac{ac}{bd} = \dfrac{a}{b}\dfrac{c}{d}$.

(xii) *For all* $a, c \in \mathbb{R}$ *and all* $b, d \in \mathbb{R} \setminus \{0\}$ *the following equivalence holds*

$$\frac{a}{b} = \frac{c}{d} \quad \Leftrightarrow \quad ad = bc.$$

(xiii) *For all* $a \in \mathbb{R}$ *and all* $b, c \in \mathbb{R} \setminus \{0\}$ *we have* $\dfrac{ac}{bc} = \dfrac{a}{b}$.

(xiv) *For all* $a \in \mathbb{R}$ *and all* $b, c, d \in \mathbb{R} \setminus \{0\}$ *we have* $\dfrac{\frac{a}{b}}{\frac{c}{d}} = \dfrac{ad}{bc}$.

(xv) $(-1)(-1) = 1$.

(xvi) *For all* $a \in \mathbb{R}$ *we have* $-a = (-1)a$.

(xvii) *For all* $a, b \in \mathbb{R}$ *we have* $(-a)b = -(ab)$ *and* $a(-b) = -(ab)$.

(xviii) *For all* $a, b \in \mathbb{R}$ *we have* $(-a)(-b) = ab$.

(xix) *For all* $a \in \mathbb{R}\setminus\{0\}$ *we have* $-\dfrac{1}{a} = \dfrac{1}{-a}$.

PROOF. (i) By Axiom AZ we have $0 + 0 = 0$. Definition 4.1.3 yields $-0 = 0$.

(ii) Let $a, b, c \in \mathbb{R}$ be arbitrary. Assume that $a + c = b + c$. By Axiom ME adding any $x \in \mathbb{R}$ to both sides of the equality leads to $(a + c) + x = (b + c) + x$. It follows from Axiom AA that $a + (c + x) = b + (c + x)$. By Axiom AO, there exists $-c \in \mathbb{R}$ such that $c + (-c) = 0$. Choose $x = -c$. Then $a = a + 0 = a + (c + (-c)) = b + (c + (-c)) = b + 0 = b$.

(iii) Let $a \in \mathbb{R}$ be arbitrary. By Definition 4.1.3, we have $a + (-a) = 0$ and $(-a) + \big(-(-a)\big) = 0$. By Axiom AC, $(-a) + a = (-a) + \big(-(-a)\big)$. Now (ii), yields $a = -(-a)$.

(iv) Let $a, b \in \mathbb{R}$ be arbitrary, and start the calculation by applying Axiom AA,

$$
\begin{aligned}
(a + b) + \big((-a) + (-b)\big) &= a + \Big(b + \big((-a) + (-b)\big)\Big) \\
\boxed{\text{by Axiom AC}} \quad &= a + \Big(\big((-a) + (-b)\big) + b\Big) \\
\boxed{\text{by Axiom AA, twice}} \quad &= \big(a + (-a)\big) + \big((-b) + b\big) \\
\boxed{\text{by Definition 4.1.3 and Axiom AC}} \quad &= 0 + \big(b + (-b)\big) \\
\boxed{\text{by Definition 4.1.3}} \quad &= 0 + 0 \\
\boxed{\text{by Axiom AZ}} \quad &= 0.
\end{aligned}
$$

The preceding calculation and Definition 4.1.3 yield $-(a + b) = (-a) + (-b)$.

(v) By Axiom MO we have $1 \cdot 1 = 1$. Definition 4.1.3 states thaT the real number $\dfrac{1}{1}$ satisfies $1 \cdot \dfrac{1}{1} = 1$. By Theorem 4.1.2 the equation $1 \cdot x = 1$ has a unique solution. Consequently, $\dfrac{1}{1} = 1$.

(vi) Let $a, b, c \in \mathbb{R}$ be arbitrary. Assume that $c \neq 0$ and $ac = bc$. By Axiom MR and Definition 4.1.3 there exists $1/c \in \mathbb{R}$ such that $c(1/c) = 1$. Then, by Axiom ME we have $(ac)(1/c) = (bc)(1/c)$. Axiom MA yields, $a\big(c(1/c)\big) = b\big(c(1/c)\big)$, that is $a \cdot 1 = b \cdot 1$, and (vi) follows from Axiom MO.

(vii) Let $a \in \mathbb{R}$ be arbitrary. By Axiom ME, the product $a \cdot 0$ is a real number. Set $a \cdot 0 = w \in \mathbb{R}$. By Axiom AO, there exists $-w \in \mathbb{R}$ such that $w + (-w) = 0$. By Axiom AZ, we have $0 + 0 = 0$. By Axiom ME, $a \cdot (0 + 0) = a \cdot 0$. By Axiom DL, we have $a \cdot 0 + a \cdot 0 = a \cdot 0$. Since we set $a \cdot 0 = w$, we proved $w + w = w$. By Axiom AE, $(w + w) + (-w) = w + (-w)$. By Axiom AA, $w + \big(w + (-w)\big) = w + (-w)$. Since by Axiom AO, $w + (-w) = 0$, we have $w + 0 = 0$. By Axiom AZ, $w = 0$. Hence, $a \cdot 0 = 0$.

(viii) The 'if' part of (viii) follows from (vii). Now we prove the 'only if' part of (viii). That is, we let $a, b \in \mathbb{R}$ be arbitrary and prove the implication:

$$ab = 0 \quad \Rightarrow \quad a = 0 \ \vee \ b = 0.$$

The implication $p \Rightarrow (q \vee r)$ is equivalent to $p \wedge (\neg r) \Rightarrow q$. Therefore we will prove the equivalent implication

$$ab = 0 \ \wedge \ b \neq 0 \quad \Rightarrow \quad a = 0.$$

From (vii) and Axiom AC it follows that $0 = 0 \cdot b$. Therefore we need to prove

$$ab = 0 \cdot b \ \wedge \ b \neq 0 \quad \Rightarrow \quad a = 0.$$

This implication is proved in (vi).

(ix) Let $a \in \mathbb{R} \setminus \{0\}$. Then by Axiom MR and Definition 4.1.3 we have $a(1/a) = 1$. Since $1 \neq 0$, part (viii) of this theorem yields $(1/a) \neq 0$. Since by Axiom MC we have $(1/a)a = 1$, we see that a solves the equation $(1/a)x = 1$. Hence Definition 4.1.3 implies that $\dfrac{1}{1/a} = a$.

(x) Let $a, b \in \mathbb{R} \setminus \{0\}$. By (viii) we have $ab \neq 0$. Next we will show that $\frac{1}{a}\frac{1}{b}$ is the solution of the equation $(ab)x = 1$. To that end, calculate, starting by applying Axiom MA,

$$\begin{aligned}
(ab)\Big(\frac{1}{a}\frac{1}{b}\Big) &= \Big((ab)\frac{1}{a}\Big)\frac{1}{b} \\
\boxed{\text{by Axiom MC}} &= \Big(\frac{1}{a}(ab)\Big)\frac{1}{b} \\
\boxed{\text{by Axiom MA}} &= \Big(\Big(\frac{1}{a}a\Big)b\Big)\frac{1}{b} \\
\boxed{\text{by Definition 4.1.3}} &= (1 \cdot b)\frac{1}{b} \\
\boxed{\text{by Axiom MO}} &= b\frac{1}{b} \\
\boxed{\text{by Definition 4.1.3}} &= 1.
\end{aligned}$$

Now (x) follows from Definition 4.1.3.

(xi) Let $a, c \in \mathbb{R}$ and $b, d \in \mathbb{R} \setminus \{0\}$ be arbitrary. Then, starting with an application of Definition 4.1.3,

$$\begin{aligned}
\frac{ac}{bd} &= (ac)\frac{1}{bd} \\
\boxed{\text{By (x)}} &= (ac)\left(\frac{1}{b}\frac{1}{d}\right) \\
\boxed{\text{Apply Axioms MA and MC}} &= \left(a\frac{1}{b}\right)\left(c\frac{1}{d}\right) \\
\boxed{\text{Definition 4.1.3}} &= \frac{a}{b}\frac{c}{d}.
\end{aligned}$$

(xii) Let $a, c \in \mathbb{R}$ and $b, d \in \mathbb{R} \setminus \{0\}$ be arbitrary. By (viii) we have $bd \neq 0$. The following equivalences hold, the first one being the consequence of (vi) and

Axiom ME,

$$\begin{aligned}
\frac{a}{b} = \frac{c}{d} \quad &\Leftrightarrow \quad \frac{a}{b}(bd) = \frac{c}{d}(bd) \\
\boxed{\text{Why?}} \quad &\Leftrightarrow \quad \left(\frac{a}{b}b\right) d = \frac{c}{d}(db) \\
\boxed{\text{Why?}} \quad &\Leftrightarrow \quad \left(\left(a\frac{1}{b}\right) b\right) d = \left(\frac{c}{d}d\right) b \\
\boxed{\text{Why?}} \quad &\Leftrightarrow \quad \left(a\left(\frac{1}{b}b\right)\right) d = \left(c\left(\frac{1}{d}d\right)\right) b \\
\boxed{\text{Why?}} \quad &\Leftrightarrow \quad ad = bc
\end{aligned}$$

To prove (xiii) and (xiv) apply (xii).

(xv) By Definition 4.1.3 we have $1 + (-1) = 0$. By Axiom ME we have

$$\big(1 + (-1)\big)\big(1 + (-1)\big) = 0 \cdot 0. \tag{4.1.1}$$

Calculate the left-hand side of (4.1.1), starting by applying Axioms DL,

$$\begin{aligned}
\big(1 + (-1)\big)\big(1 + (-1)\big) &= \big(1 + (-1)\big) \cdot 1 + \big(1 + (-1)\big)(-1) \\
\boxed{\text{Why?}} &= \big(1 + (-1)\big) + \big((-1) + (-1)(-1)\big) \\
\boxed{\text{Why?}} &= 0 + \big((-1) + (-1)(-1)\big) \\
\boxed{\text{Why?}} &= (-1) + (-1)(-1)
\end{aligned}$$

Using just proved equality, (4.1.1), and $0 \cdot 0 = 0$ proved in (vii), we obtain

$$(-1) + (-1)(-1) = 0.$$

Now, by Definition 4.1.3 and Axiom AC we have $(-1) + 1 = 0$. The last two equalities and part (ii) of this proof yield $(-1)(-1) = 1$.

(xvi) Let $a \in \mathbb{R}$ be arbitrary and calculate

$$\begin{aligned}
a + (-1)a &= 1a + (-1)a \\
\boxed{\text{Why?}} &= a \cdot 1 + a(-1) \\
\boxed{\text{Why?}} &= a\big(1 + (-1)\big) \\
\boxed{\text{Why?}} &= a \cdot 0 \\
\boxed{\text{Why?}} &= 0.
\end{aligned}$$

Thus, $a + (-1)a = 0$, and Definition 4.1.3 implies $-a = (-1)a$.

(xvii) Let $a, b \in \mathbb{R}$ be arbitrary and calculate

$$\begin{aligned}
(-a)b &= \big((-1)a\big)b \\
\boxed{\text{Why?}} &= (-1)(ab) \\
\boxed{\text{Why?}} &= -(ab).
\end{aligned}$$

(xviii) Let $a, b \in \mathbb{R}$ be arbitrary and calculate

$$\begin{aligned}
(-a)(-b) &= \big((-1)a\big)\big((-1)b\big) \\
\boxed{\text{Why?}} &= \Big(\big((-1)a\big)(-1)\Big)b
\end{aligned}$$

$$
\begin{aligned}
\boxed{\text{Why?}} &= \Big((-1)\big((-1)a\big)\Big)b \\
\boxed{\text{Why?}} &= \Big(\big((-1)(-1)\big)a\Big)b \\
\boxed{\text{Why?}} &= (1 \cdot a)b \\
\boxed{\text{Why?}} &= ab.
\end{aligned}
$$

(xix) Let $a \in \mathbb{R} \setminus \{0\}$ be arbitrary. Then $-a \neq 0$ and we will show that $-(1/a)$ solves the equation $(-a)x = 1$. To that end, calculate

$$
\begin{aligned}
(-a)\left(-\frac{1}{a}\right) &= \big((-1)a\big)\left((-1)\frac{1}{a}\right) \\
\boxed{\text{Why?}} &= \Big(\big((-1)a\big)(-1)\Big)\,\frac{1}{a} \\
\boxed{\text{Why?}} &= \Big((-1)\big((-1)a\big)\Big)\,\frac{1}{a} \\
\boxed{\text{Why?}} &= \Big(\big((-1)(-1)\big)a\Big)\,\frac{1}{a} \\
\boxed{\text{Why?}} &= (1 \cdot a)\,\frac{1}{a} \\
\boxed{\text{Why?}} &= a\,\frac{1}{a} \\
\boxed{\text{Why?}} &= 1.
\end{aligned}
$$

Now (xix) follows from Definition 4.1.3. □

REMARK 4.1.5. It is interesting to point out that Axioms 1 through 11 are not sufficient to prove the following implication. For all $a \in \mathbb{R}$ we have: If $-a = a$, then $a = 0$. To prove this claim consider Example 4.1.1. This example satisfies Axioms 1 through 11 but, in this example, $-1 = 1$, while $1 \neq 0$.

We will prove that $-a = a \Rightarrow a = 0$ in Theorem 4.1.8(vi). ◁

4.1.2. Axioms of order in a field. Axioms OE, OT, OA, and OM establish the total order on the set $\mathbb{R}$. Axiom OE, referred to as the *trichotomy law*, distinguishes exactly one of three possible relations between any pair of numbers. The property of order in Axiom OT is known as *transitive law*. Axiom OA highlights the harmony between addition and order, ensuring that adding numbers preserves the order. Axiom OM defines a special relationship between multiplication and order. A field with an order satisfying Axioms 12 through 15 is called an *ordered field*.

REMARK 4.1.6. (ON NOTATION) For $a, b \in \mathbb{R}$, the notation $a \leq b$ stands for the statement: $a < b$ exclusive or $a = b$. The notation $a > b$ means $b < a$, while $a \geq b$ means $b \leq a$. For $a, b, c \in \mathbb{R}$, the notation $a < b < c$ stands for $a < b$ and $b < c$. Similarly, $a \leq b < c$ stands for $a \leq b$ and $b < c$, $a < b \leq c$ stands for $a < b$ and $b \leq c$, $a \leq b \leq c$ stands for $a \leq b$ and $b \leq c$. ◁

DEFINITION 4.1.7. A number $x \in \mathbb{R}$ is *positive* if $x > 0$. A number $x \in \mathbb{R}$ is *negative* if $x < 0$. A number $x \in \mathbb{R}$ is *nonnegative* if $x \geq 0$. A number $x \in \mathbb{R}$

is *nonpositive* if $x \leq 0$. We introduce the following notation:

$$\mathbb{R}_{>0} \stackrel{\text{def}}{=} \bigl\{x \in \mathbb{R} : x > 0\bigr\}, \qquad \mathbb{R}_{<0} \stackrel{\text{def}}{=} \bigl\{x \in \mathbb{R} : x < 0\bigr\},$$
$$\mathbb{R}_{\geq 0} \stackrel{\text{def}}{=} \bigl\{x \in \mathbb{R} : x \geq 0\bigr\}, \qquad \mathbb{R}_{\leq 0} \stackrel{\text{def}}{=} \bigl\{x \in \mathbb{R} : x \leq 0\bigr\}.$$

THEOREM 4.1.8. *The following statements hold:*

(i) *For all $a, b \in \mathbb{R}$ we have: $a < b$ implies $-b < -a$.*

(ii) *For all $a, b, c \in \mathbb{R}$ we have: $a < b$ and $c < 0$ implies $bc < ac$.*

(iii) *For all $a \in \mathbb{R}_{>0}$, for all $b \in \mathbb{R}\backslash\{0\}$ we have: $b > 0$ if and only if $ab > 0$.*

(iv) *For all $a \in \mathbb{R}\backslash\{0\}$ we have $aa > 0$.*

(v) *$0 < 1$ and $0 < 1 + 1$.*

(vi) *For all $a \in \mathbb{R}$ we have: $-a = a$ if and only if $a = 0$.*

(vii) *For all $a \in \mathbb{R}\backslash\{0\}$ we have: $a > 0$ if and only if $\dfrac{1}{a} > 0$.*

(viii) *For all $a, b \in \mathbb{R}\backslash\{0\}$ we have: $0 < a < b$ if and only if $0 < \dfrac{1}{b} < \dfrac{1}{a}$.*

PROOF. Let $a, b, c \in \mathbb{R}$ be arbitrary. (i) Assume $a < b$. By Axiom OA we have $a + (-b) < b + (-b)$. Hence, $(-b) + a < 0$. Using Axiom OA again, we conclude that $\bigl((-b) + a\bigr) + (-a) < 0 + (-a)$, and consequently $-b < -a$.

(ii) Let $a, b, c \in \mathbb{R}$ be arbitrary. Assume $a < b$ and $c < 0$. By preceding item (i) and Theorem 4.1.4(i) we have $0 < -c$. Now Axiom OM applied to $a < b$ and $0 < -c$ yields $a(-c) < b(-c)$. By Theorem 4.1.4(xvii) and Axiom MC the last inequality becomes $-ac < -bc$. Finally, by preceding item (i) and Theorem 4.1.4(iii) we get $bc < ac$.

Now we prove (iii). Assume $a > 0$ and $b \neq 0$. This assumption is used throughout this part of the proof. Since $a \neq 0$ and $b \neq 0$, by Theorem 4.1.4 (viii), it follows that $ab \neq 0$. The implication: 'If $b > 0$, then $ab > 0$' is a special case of Axiom OM. Next we deal with the implication 'If $ab > 0$, then $b > 0$.' It turns out that the contrapositive is easier to prove. The negation of $b > 0$ is $b \leq 0$. But, it is assumed that $b \neq 0$. Thus, with this assumption, the negation of $b > 0$ is $b < 0$. Similarly, the negation of $ab > 0$ is $ab < 0$. Hence, the contrapositive of 'If $ab > 0$, then $b > 0$.' is 'If $b < 0$, then $ab < 0$.' The last implication follows directly from part (ii). This completes the proof of (iii).

(iv) Consider two cases: $a > 0$ and $a < 0$. If $a > 0$, then (iii) implies that $aa > 0$. If $a < 0$, then, by (i), $-0 < -a$, and since $-0 = 0$ we have $-a > 0$. By the first part of this proof, we conclude that $(-a)(-a) > 0$. By part (xviii) of Theorem 4.1.4 we have $(-a)(-a) = a\,a$. Therefore $aa > 0$ for all $a \neq 0$.

(v) By Axiom MO we have $1 \neq 0$. Therefore, by previous item (iv), $1 \cdot 1 > 0$. Since by Axiom MO, $1 \cdot 1 = 1$ the first claim in (v) is proved. By Axiom OA, $0 < 1$ implies $0 + 1 < 1 + 1$. That is $1 < 1 + 1$. Now, $0 < 1$ and $1 < 1 + 1$, and Axiom OT imply $0 < 1 + 1$.

(vi) The "if" part of (vi) is proved in Theorem 4.1.4(i). To prove the "only if" part, let $a \in \mathbb{R}$ be arbitrary and assume $-a = a$. By Axiom AE, $a + (-a) = a + a$. By Definition 4.1.3, $0 = a + a$. By Axioms MO and DL, $(1 + 1)a = 0$. By Theorem 4.1.4(viii), $(1 + 1)a = 0$ implies that $1 + 1 = 0$ or $a = 0$. Since, by item (v), $1 + 1 \neq 0$, disjunctive syllogism yields $a = 0$.

To prove (vii) we assume $a > 0$. By Axiom MR, $a\dfrac{1}{a} = 1$. By Axiom MO, $1 \neq 0$. Hence, $a\dfrac{1}{a} \neq 0$. By Theorem 4.1.4 (viii) $\dfrac{1}{a} \neq 0$ and by (v) $1 > 0$. Now we can apply the 'if' part of (iii). (Take $b = 1/a$ in (iii).) We conclude that $a\dfrac{1}{a} = 1 > 0$ implies $\dfrac{1}{a} > 0$. This proves (vii).

Do (viii) as an exercise. □

Remark 4.1.9. (A Mnemonic Tool: Pizza-Party) The content of Theorem 4.1.8(viii) I call the Pizza-Party Principle. Why? In the inequality

$$\frac{1}{b} < \frac{1}{a},$$

the number 1 represents a pizza, and a and b represent the number of people at two different parties. A smaller party results in more pizza for each attendant; there is no doubt about that. ◁

Definition 4.1.10. We define the following eight real numbers

$$2 \stackrel{\text{def}}{=} 1+1, \qquad 3 \stackrel{\text{def}}{=} 2+1, \qquad 4 \stackrel{\text{def}}{=} 3+1, \qquad 5 \stackrel{\text{def}}{=} 4+1,$$

$$6 \stackrel{\text{def}}{=} 5+1, \qquad 7 \stackrel{\text{def}}{=} 6+1, \qquad 8 \stackrel{\text{def}}{=} 7+1, \qquad 9 \stackrel{\text{def}}{=} 8+1.$$

The real numbers $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ are called *digits*.

In the preceding definition we implied that the digits are distinct numbers. The next exercise justifies this claim.

Exercise 4.1.11. Prove the inequalities:

$$0 < 1 < 2 < 3 < 4 < 5 < 6 < 7 < 8 < 9. \qquad ◁$$

Exercise 4.1.12. Let $a, b \in \mathbb{R}$. If $a < b$, then there exists $c \in \mathbb{R}$ such that $a < c$ and $c < b$. ◁

Solution. Let $a, b \in \mathbb{R}$ be arbitrary and assume that $a < b$. By Axiom OA, adding a and then b to both sides of the inequality yields

$$2a < a + b \quad \wedge \quad a + b < 2b.$$

Since $\dfrac{1}{2} > 0$, by Axiom OM we deduce

$$a < \frac{1}{2}(a+b) \quad \wedge \quad \frac{1}{2}(a+b) < b.$$

Hence,

$$a < \frac{1}{2}(a+b) < b.$$

Thus, we can take $c = \dfrac{1}{2}(a+b)$. □

DEFINITION 4.1.13. For $a \in \mathbb{R}$, the product aa is called the *square* of a and is denoted by a^2.

The following four exercises deal with squares of real numbers.

EXERCISE 4.1.14. Let $a \in \mathbb{R}$. Prove that the equation $x^2 = a$ has at most two solutions in $\mathbb{R}$. ◁

SOLUTION. Consider the set

$$S = \{x \in \mathbb{R} : x^2 = a\}.$$

If $S = \emptyset$, then the statement is true. Now assume that $S \neq \emptyset$ and let $b \in S$. From $b \in S$, we deduce that $b \in \mathbb{R}$ and $b^2 = a$. Since $b \in \mathbb{R}$, by Axiom AO and Definition 4.1.3, $-b \in \mathbb{R}$. Next we will prove

$$S = \{b, -b\}. \tag{4.1.2}$$

Let $c \in S$. Then $c^2 = a$, and therefore $c^2 = b^2$. Consequently, $c^2 - b^2 = 0$. Using Axioms 2 through 11 and properties in Theorem 4.1.4 we can prove that $(c-b)(c+b) = c^2 - b^2$. Therefore $(c-b)(c+b) = c^2 - b^2 = 0$. Theorem 4.1.4 (viii) implies that $c - b = 0$ or $c + b = 0$. Thus $c = b$ or $c = -b$. This proves

$$S \subseteq \{b, -b\}. \tag{4.1.3}$$

Next we prove $\{b, -b\} \subseteq S$. By assumption, $b \in S$. Since $(-b)^2 = b^2$, we have $(-b)^2 = a$. Hence $-b \in S$. Therefore

$$\{b, -b\} \subseteq S. \tag{4.1.4}$$

The inclusions in (4.1.3) and (4.1.4) imply equality (4.1.2). Since the set $\{b, -b\}$ has at most two elements the statement is proved. □

EXERCISE 4.1.15. For all $a, b \in \mathbb{R}$ we have $(a+b)^2 = a^2 + 2ab + b^2$. ◁

EXERCISE 4.1.16. For all $x, y \in \mathbb{R}_{\geq 0}$ the following equivalences hold:

(i) $x = y$ if and only if $x^2 = y^2$.

(ii) $x < y$ if and only if $x^2 < y^2$. ◁

SOLUTION. (i) Let $x, y \in \mathbb{R}_{\geq 0}$ be arbitrary. The "only if" part, that is, the implication $x = y \ \Rightarrow \ x^2 = y^2$ follows from Axiom ME.

To prove the "if" part assume $x^2 = y^2$. Consider two cases $y = 0$ and $y > 0$. If $y = 0$, then, by Theorem 4.1.4(vii) $y^2 = 0$. Hence, $x^2 = 0$. By Theorem 4.1.4(viii), $x^2 = 0$ implies $x = 0$. Thus, $x = y$.

Now assume $y > 0$. Then by Theorem 4.1.8(i) and Theorem 4.1.4(i), we have $-y < 0$. Since $x \geq 0$ we have $x \neq -y$. The assumption $x^2 = y^2$ yields

$$0 = x^2 - y^2 = (x-y)(x+y).$$

By Theorem 4.1.4(viii), the equality $(x-y)(x+y) = 0$ implies $x = y$ or $x = -y$. Since we proved $x \neq -y$, disjunctive syllogism implies $x = y$. Thus, $x^2 = y^2$ implies $x = y$.

(ii) □

EXERCISE 4.1.17. For all $a, x \in \mathbb{R}$ we have: $a > 1$ and $a > x^2$ imply $a > x$. ◁

EXERCISE 4.1.18. For all $s, t \in \mathbb{R}$ we have: $s \neq t$ implies $(s+t)^2 > 4st$. ◁

EXERCISE 4.1.19. Let $a, b, c, d \in \mathbb{R}$.

(i) Prove or disprove the statement: If $a < b$ and $c < d$, then $a - c < b - d$.
(ii) If you disproved the statement in (i), change the assumptions about c and d to make a correct statement. Prove your new statement. $\triangleleft$

EXERCISE 4.1.20. Let $\alpha \in \mathbb{R}$. Prove the following statement

$$\Big(\forall x > 0 \quad \text{we have} \quad \alpha \leq x\Big) \quad \Rightarrow \quad \alpha \leq 0. \qquad \triangleleft$$

PROOF. Let $\alpha \in \mathbb{R}$ be arbitrary and prove the contrapositive:

$$\alpha > 0 \quad \Rightarrow \quad \Big(\exists x > 0 \quad \text{such that} \quad x < \alpha\Big).$$

Assume $\alpha > 0$. In Exercise 4.1.11 we proved $0 < 1 < 2$. By Theorem 4.1.8(viii) and Theorem 4.1.4(v), $0 < \frac{1}{2} < 1$. Now $\alpha > 0$, Axiom OM, Theorem 4.1.4(vii), and Axiom MO yield $0 < \frac{1}{2}\alpha < \alpha$. Thus, the claim of the contrapositive is true for $x = \frac{1}{2}\alpha$. $\square$

REMARK 4.1.21. The properties of real numbers proved in this section are essential. Many of them are truly elementary (although sometimes hard to prove) and you can (and I will) use such properties in proofs without any justification. But, when you are using more subtle properties (like ones in Exercises 4.1.16, 4.1.17, or 4.1.19) you should state explicitly which property you are using and either prove it, or post it as a question to the class. $\triangleleft$

4.1.3. The number line. An exceptionally useful tool for internalizing the real numbers and their properties is their visualization on a number line. We think of points on a straight line as representing real numbers, as shown in Figure 4.1.1, where the digits $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and their opposites are marked by black tick marks, and the real number line itself is depicted as a slightly thicker gray line. I marked 0 and 1 in boldface, as they are the only numbers explicitly mentioned in the Axioms. This emphasis appears only in this first visualization of the real number line.

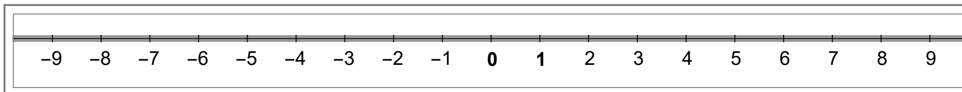


FIG. 4.1.1. A portion of the real number line with the digits and their opposites

4.2. Five functions to start with

In the preceding section, we explored the algebra of real numbers in some depth. The next natural topic is to investigate functions with $\mathbb{R}$ as both domain and codomain. With nonempty sets A and B, we defined in Section 3.2 a function with domain A and codomain B as a special subset of the Cartesian product $A \times B$. To visualize functions whose domain and codomain are both $\mathbb{R}$, we use the Cartesian product $\mathbb{R} \times \mathbb{R}$. A visualization of $\mathbb{R} \times \mathbb{R}$, known as the Cartesian plane, is given in Figure 4.2.1.

In Definition 3.2.5, we introduced the identity function on a nonempty set A. The first of the five functions presented in this section is the identity function on $\mathbb{R}$:

$$\mathrm{id}_{\mathbb{R}} \stackrel{\text{def}}{=} \big\{(x,x) \in \mathbb{R}\times\mathbb{R} : x \in \mathbb{R}\big\},$$

or, using traditional notation:

$$\forall x \in \mathbb{R} \qquad \mathrm{id}_{\mathbb{R}}(x) \stackrel{\text{def}}{=} x.$$

A graph of the identity function on $\mathbb{R}$ is shown in Figure 4.2.2. Note that the notation $\mathrm{id}_{\mathbb{R}}$, introduced above, is not standard. In fact, no universally accepted notation exists; most mathematics textbooks use this function without assigning it a specific symbol. I have chosen the abbreviation id , as most named mathematical functions use two- or three-letter abbreviations, and the subscript $\mathbb{R}$ to indicate its domain and codomain.

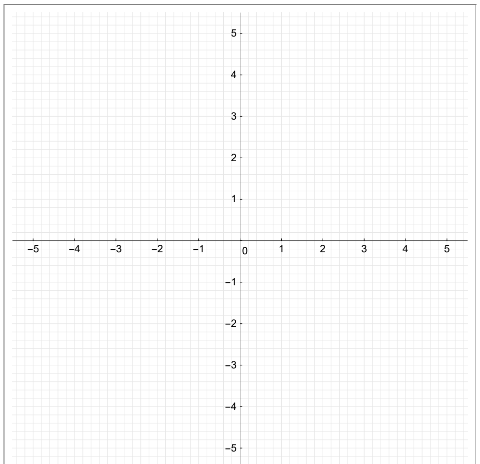


FIG. 4.2.1. The Cartesian plane $\mathbb{R}\times\mathbb{R}$

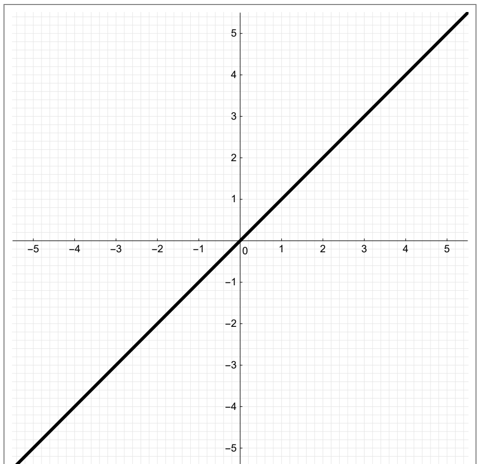


FIG. 4.2.2. The identity function

There are only two specific real numbers mentioned in Axioms 1 through 16. These real numbers are 0 (zero) and 1 (one). The number -1 is implicitly mentioned in Axiom AO. Therefore the following two functions are of interest.

DEFINITION 4.2.1. The *unit step* function, denoted by $\mathrm{us} : \mathbb{R} \to \mathbb{R}$ and shown in Figure 4.2.3, is defined as follows:

$$\forall x \in \mathbb{R} \qquad \mathrm{us}(x) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if} \quad x < 0, \\ 1 & \text{if} \quad x \geq 0. \end{cases}$$

DEFINITION 4.2.2. The *sign* function, denoted by $\mathrm{sgn} : \mathbb{R} \to \mathbb{R}$ and shown in Figure 4.2.4, is defined as follows

$$\forall x \in \mathbb{R} \qquad \mathrm{sgn}(x) \stackrel{\text{def}}{=} \begin{cases} -1 & \text{if} \quad x < 0, \\ 0 & \text{if} \quad x = 0, \\ 1 & \text{if} \quad x > 0. \end{cases}$$

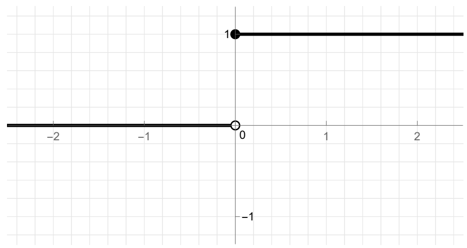


FIG. 4.2.3. The unit step function

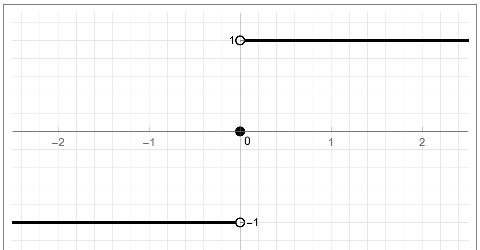


FIG. 4.2.4. The sign function

REMARK 4.2.3. The graph of the unit step function shown in Figure 4.2.3 consists of two half-lines:

$$\big\{(x,0) \in \mathbb{R}\times\mathbb{R} : x < 0\big\} \quad \text{and} \quad \big\{(x,1) \in \mathbb{R}\times\mathbb{R} : x \geq 0\big\}.$$

To visually emphasize that the point $(0,1)$ belongs to the second half-line, we place a *solid disc* at $(0,1)$ and an *open circle* at $(0,0)$. The open circle indicates that the point $(0,0)$ is *not* part of the graph.

Similarly, the graph of the sign function shown in Figure 4.2.4 consists of two half-lines and a single point:

$$\big\{(x,-1) \in \mathbb{R}\times\mathbb{R} : x < 0\big\}, \quad \big\{(x,1) \in \mathbb{R}\times\mathbb{R} : x > 0\big\}, \quad \text{and} \quad (0,0).$$

To clarify which points are included, we place open circles at $(0,-1)$ and $(0,1)$, indicating those points are not part of the graph, and a solid disc at $(0,0)$, indicating that it is.

This use of *solid discs* and *open circles* helps readers clearly see which points are included in a graph and which are not. It is a common visual convention used to avoid ambiguity, especially when values change abruptly. ◁

DEFINITION 4.2.4. The *absolute value* function, denoted by $\text{abs} : \mathbb{R} \to \mathbb{R}$ and shown in Figure 4.2.5, is defined as

$$\forall x \in \mathbb{R} \qquad \text{abs}(x) \stackrel{\text{def}}{=} x\,\text{sgn}(x).$$

We will also use the standard notation $\text{abs}(x) = |x|$. The number $|x|$ is called the *absolute value* of the number x.

DEFINITION 4.2.5. The *Rectified Linear Unit* function, abbreviated to ReLU, denoted here by $\text{rlu} : \mathbb{R} \to \mathbb{R}$ and shown in Figure 4.2.6, is defined as

$$\forall x \in \mathbb{R} \qquad \text{rlu}(x) \stackrel{\text{def}}{=} x\,\text{us}(x).$$

The pronunciation of ReLU, and our notation rlu, is /'rEl.u:/.

The most important of these four functions is the absolute value function. Why is the absolute value so important? Its importance lies in its geometric significance. The common geometric representation of real numbers is as points on a straight line. This representation is achieved by choosing a point on the line to represent 0 and another point to represent 1. Consequently, every real number corresponds to a point on this line, commonly called the real number line, and every point on the number line corresponds to a real number. With this geometric representation, the

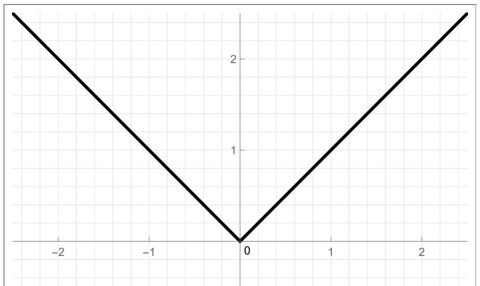


FIG. 4.2.5. The absolute value

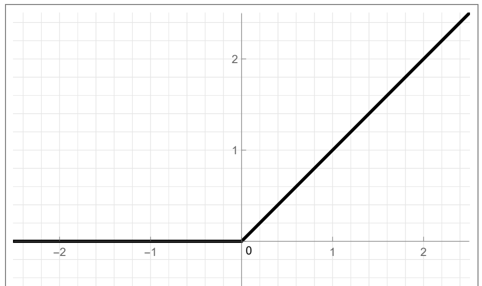


FIG. 4.2.6. The Rectified Linear Unit

absolute value of x, denoted $|x|$, represents the distance of the number x from the number 0. Generally, we introduce the following definition:

DEFINITION 4.2.6. Let $a, b \in \mathbb{R}$. We define

$$\operatorname{dist}(a, b) \overset{\text{def}}{=} |b - a|.$$

The number $\operatorname{dist}(a, b)$ is called the *distance between* the real numbers a and b.

The other three functions are there for the absolute value not to feel lonely. Additionally, it is beneficial to have a library of simple functions that can be used in examples. The sign function is in such nice harmony with the absolute value that I defined the sign function first, then used it to define the absolute value. Furthermore, one could argue that the sign function is simpler, as it takes only three values; its range is the set $\{-1, 0, 1\}$.

The unit step function is a bridge to indicator functions introduced in Definition 3.2.21. It is the indicator function of the set $\mathbb{R}_{\geq 0}$ of all nonnegative real numbers. Although, I have not seen the unit step function discussed in undergraduate mathematics classes, it does appear in physical and engineering applications, sometimes under the name Heaviside function, and sometimes assuming a different value at 0.

Notice the similarity in the definitions of the absolute value function and the *Rectified Linear Unit* (ReLU) function (see Figures 4.2.5 and 4.2.6): the absolute value function is defined as the product of the identity function of $\mathbb{R}$ and the sign function, while the ReLU function is the product of the identity function of $\mathbb{R}$ and the unit step function. The ReLU function has become quite famous due to its extensive use in deep learning applications.

The concept of a minimum and a maximum is in general useful, and it is closely related to the absolute value, as we can see in the theorem and exercises below.

DEFINITION 4.2.7. Let $a, b \in \mathbb{R}$. We define

$$\min\{a, b\} \overset{\text{def}}{=} \begin{cases} a & \text{if } a < b, \\ b & \text{if } b \leq a, \end{cases} \qquad \max\{a, b\} \overset{\text{def}}{=} \begin{cases} b & \text{if } a < b, \\ a & \text{if } b \leq a. \end{cases}$$

The number $\min\{a, b\}$ is called the *minimum of* a and b, and the number $\max\{a, b\}$ is called the *maximum of* a and b.

THEOREM 4.2.8. *The following statements hold:*

(i) *For all* $x \in \mathbb{R}$ *we have* $|x| = \max\{x, -x\}$.

(ii) *For all* $x \in \mathbb{R}$ *we have* $|x| \geq 0$.

(iii) *For all* $x \in \mathbb{R}$ *we have:* $|x| = 0$ *if and only if* $x = 0$.

(iv) *For all* $x \in \mathbb{R}$ *we have* $\big||x|\big| = |x|$.

(v) *For all* $x \in \mathbb{R}$ *we have* $|-x| = |x|$.

(vi) *For all* $x, y \in \mathbb{R}$ *we have* $|xy| = |x|\,|y|$.

(vii) *For all* $x \in \mathbb{R}$, *for all* $y \in \mathbb{R}\backslash\{0\}$ *we have* $\left|\dfrac{x}{y}\right| = \dfrac{|x|}{|y|}$.

(viii) *For all* $x \in \mathbb{R}$ *we have* $x^2 = |x|^2$.

(ix) *For all* $x \in \mathbb{R} \setminus \{0\}$ *we have* $|x| < 1$ *if and only if* $x^2 < |x|$.

REMARK 4.2.9. Many of the proofs involving the absolute value function are done by cases. In simple settings these are efficient proofs. However, in more complicated situations listing all possible cases might be challenging. Therefore, pay attention to the proofs which are not done by cases. ◁

PROOF OF THEOREM 4.2.8. (i) Let $x \in \mathbb{R}$ be arbitrary. We consider three cases, $x = 0$, $x > 0$, and $x < 0$. Assume $x = 0$. By the definition of the absolute value we have $|0| = 0$, and by the definition of the maximum we have $\max\{0, -0\} = 0$. Hence, the claim holds in this case. Assume $x > 0$. Then $|x| = x$, and $-x < 0$, so $\max\{x, -x\} = x$. Thus, $|x| = \max\{x, -x\}$, in this case as well. Assume $x < 0$, then $|x| = -x$, and $-x > 0$, so $\max\{x, -x\} = -x$. Hence, $|x| = \max\{x, -x\}$ in this case as well, completing the proof of (i).

(vii) Let $y \in \mathbb{R}\backslash\{0\}$ be arbitrary and first prove that $\big|1/y\big| = 1/|y|$. If $y > 0$, by Theorem 4.1.8(vii) we have $1/y > 0$, hence $\big|1/y\big| = 1/y = 1/|y|$. If $y < 0$, then $-y > 0$, so by the first part of the proof, $\big|1/(-y)\big| = 1/|-y|$. By Theorem 4.1.4(xix) $1/(-y) = -(1/y)$. With this equality, part (v) yields, $\big|1/y\big| = 1/|y|$ in this case as well. Now the general case of (vii) follows from (vi). Let $x \in \mathbb{R}$ be arbitrary. Then

$$\left|\frac{x}{y}\right| = \left|x\,\frac{1}{y}\right| = |x|\left|\frac{1}{y}\right| = |x|\frac{1}{|y|} = \frac{|x|}{|y|},$$

where the first equality sign follows from Definition 4.1.3, the second equality sign follows from (vi), the third from the first part of this proof for (vii), and the last again from Definition 4.1.3. □

THEOREM 4.2.10. *The following statements hold:*

(i) *For all* $x \in \mathbb{R}$ *we have* $x \leq |x|$ *and* $-x \leq |x|$.

(ii) *For all* $x \in \mathbb{R}$ *we have* $-|x| \leq x \leq |x|$.

(iii) *For all* $x, a \in \mathbb{R}$ *we have:* $|x| \leq a$ *if and only if* $-a \leq x \leq a$.

(iv) *For all* $x, a \in \mathbb{R}$ *we have:* $|x| \geq a$ *if and only if* $x \leq -a$ *or* $x \geq a$.

PROOF. □

THEOREM 4.2.11 (Triangle Inequality). *The following statements hold:*

(I) *For all* $a, b \in \mathbb{R}$ *we have* $|a + b| \leq |a| + |b|$.

(II) *For all* $a, b \in \mathbb{R}$ *we have:* $|a + b| = |a| + |b|$ *if and only if* $ab \geq 0$.

(III) *For all* $a, b \in \mathbb{R}$ *we have:* $|a + b| < |a| + |b|$ *if and only if* $ab < 0$.

(IV) *For all* $x, y \in \mathbb{R}$ *we have:* $\big||x| - |y|\big| \leq |x - y|$.

(V) *For all* $x, y \in \mathbb{R}$ *we have:* $\big||x| - |y|\big| = |x - y|$ *if and only if* $xy \geq 0$.

(VI) *For all* $x, y, z \in \mathbb{R}$ *we have* $|x - z| \leq |x - y| + |y - z|$. *In terms of the distance introduced in Definition* 4.2.6 *this inequality can be expressed as:*

$$\operatorname{dist}(x, z) \leq \operatorname{dist}(x, y) + \operatorname{dist}(y, z).$$

PROOF. (I) Let $a, b \in \mathbb{R}$ be arbitrary. By Theorem 4.2.10 (i) we have $a \leq |a|$ and $b \leq |b|$. By an straightforward consequence of Axiom OA we have $a + b \leq |a| + |b|$. Similarly, $-a \leq |a|$, $-b \leq |b|$, and $(-a) + (-b) \leq |a| + |b|$. Since by Theorem 4.1.4(iv, $(-a) + (-b) = -(a + b)$, we have $-(a + b) \leq |a| + |b|$. Hence, we proved two inequalities $a + b \leq |a| + |b|$ and $-(a + b) \leq |a| + |b|$. Consequently,

$$\max\{a + b, -(a + b)\} \leq |a| + |b|.$$

Now, (I) follows from Theorem 4.2.8(i).

(II) By Exercise 4.1.16(i) we have $|a + b| = |a| + |b|$ if and only if $|a + b|^2 = \big(|a| + |b|\big)^2$. By Theorem 4.2.8(viii) we have $|a+b|^2 = (a+b)^2$, and by Exercise 4.1.15, $|a + b|^2 = \big(|a| + |b|\big)^2$ becomes

$$a^2 + 2ab + b^2 = |a|^2 + 2|a||b| + |b|^2.$$

Since, $a^2 = |a|^2$ and $b^2 = |b|^2$, the last equality is equivalent to $ab = |ab|$. Thus, $|a+b| = |a|+|b|$ is equivalent to $ab = |ab|$. Finally, by the definition of the absolute value function, $ab = |ab|$ is equivalent to $ab \geq 0$.

(IV) Let $x, y \in \mathbb{R}$ be arbitrary. Set $a = x - y$ and $b = y$ in the triangle inequality proved in (I). Since by Axioms AA, AZ and AO we have

$$a + b = (x - y) + y = x + \big((-y) + y\big) = x.$$

the triangle inequality in (I) yields

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

By Axioms AA, AZ and AO we deduce

$$|x| - |y| \leq |x - y|. \tag{4.2.1}$$

Using similar ideas we prove

$$-\big(|x| - |y|\big) \leq |x - y|. \tag{4.2.2}$$

By properties of the maximum, inequalities (4.2.1) (4.2.2) imply

$$\max\big\{|x| - |y|, -\big(|x| - |y|\big)\big\} \leq |x - y|.$$

Now, (IV) follows from Theorem 4.2.8(i). □

The inequalities in Theorem 4.2.11 are different forms the *Triangle Inequality*. The inequalities in (I) and (VI) are called the *Triangle Inequality*, while (IV) is called the *reverse triangle inequality*. Item (II) provides the necessary and sufficient condition for the equality to hold in the triangle inequality, while item (III) provides the necessary and sufficient condition for the strict inequality. Item (V) provides the necessary and sufficient condition for the equality to hold in the reverse triangle inequality. The reader can formulate the equivalent condition for the strict inequality.

EXERCISE 4.2.12. For all $c, x \in \mathbb{R}$ and all $d \in \mathbb{R}_{>0}$ the following equivalence holds: $|x - c| \leq d$ if and only if $c - d \leq x \leq c + d$. ◁

EXERCISE 4.2.13. For all $a, x \in \mathbb{R}$ we have: If $|x - a| \leq 1$, then $|x| \leq 1 + |a|$. ◁

EXERCISE 4.2.14. For all $a, x \in \mathbb{R}$ we have: If $|x - a| \leq 1$, then $|x + a| \leq 1 + 2|a|$. ◁

EXERCISE 4.2.15. For all $a, x \in \mathbb{R}$ the following implication holds:

$$|x - a| \leq 1 \quad \Rightarrow \quad \bigl|x^2 - a^2\bigr| \leq \bigl(1 + 2|a|\bigr)|x - a|. \qquad ◁$$

EXERCISE 4.2.16. For all $x \in \mathbb{R}$ and all $a \in \mathbb{R}\setminus\{0\}$ the following implication holds:

$$|x - a| < \frac{|a|}{2} \quad \Rightarrow \quad |x| > \frac{|a|}{2}. \qquad ◁$$

EXERCISE 4.2.17. Prove the following identities:

(i) For all $a, b \in \mathbb{R}$ we have $|a - b| = \max\{a, b\} - \min\{a, b\}$.

(ii) For all $a, b \in \mathbb{R}$ we have $a + b = \max\{a, b\} + \min\{a, b\}$.

(iii) For all $a, b \in \mathbb{R}$ we have $\max\{a, b\} = \dfrac{1}{2}\Bigl(a + b + |a - b|\Bigr)$.

(iv) For all $a, b \in \mathbb{R}$ we have $\min\{a, b\} = \dfrac{1}{2}\Bigl(a + b - |a - b|\Bigr)$. ◁

EXERCISE 4.2.18. Prove the following identities:

(i) For all $x, y \in \mathbb{R}$ we have $\max\{x, y\} = x + (y - x)\ \mathrm{us}(y - x)$

(ii) For all $x, y \in \mathbb{R}$ we have $\min\{x, y\} = y + (x - y)\ \mathrm{us}(y - x)$. ◁

EXERCISE 4.2.19. For all $x \in \mathbb{R}$ we have $\mathrm{sgn}(x) = \mathrm{us}(x) - \mathrm{us}(-x)$. ◁

EXERCISE 4.2.20. For all $x \in \mathbb{R}$ we have $\mathrm{us}(x) = 1 - \dfrac{1}{2}\bigl(\mathrm{sgn}(x) - 1\bigr)\bigl(\mathrm{sgn}(x)\bigr)$. ◁

EXERCISE 4.2.21. For all $x \in \mathbb{R}$ we have $|x| = x\,\bigl(2\,\mathrm{us}(x) - 1\bigr)$. ◁

4.3. Intervals

4.3.1. Nine kinds of intervals. In Exercise 4.1.12 we proved that for arbitrary real numbers $a, b \in \mathbb{R}$ such that $a < b$ there exists $c \in \mathbb{R}$ such that $a < c < b$. This fact proves that all the sets defined below are nonempty.

Definition 4.3.1. Let a and b be real numbers such that $a < b$. We will use the following notation and terminology:

$$[a, b] \stackrel{\text{def}}{=} \{x \in \mathbb{R} : a \leq x \ \wedge \ x \leq b\} \quad \text{is called} \quad \textit{a closed interval},$$
$$(a, b) \stackrel{\text{def}}{=} \{x \in \mathbb{R} : a < x \ \wedge \ x < b\} \quad \text{is called} \quad \textit{an open interval},$$
$$[a, b) \stackrel{\text{def}}{=} \{x \in \mathbb{R} : a \leq x \ \wedge \ x < b\} \quad \text{is called} \quad \textit{a half-open interval},$$
$$(a, b] \stackrel{\text{def}}{=} \{x \in \mathbb{R} : a < x \ \wedge \ x \leq b\} \quad \text{is called} \quad \textit{a half-open interval}.$$

The above four types of intervals are *bounded* intervals. We also define four types of *unbounded* intervals:

$$\mathbb{R}_{\geq a} \stackrel{\text{def}}{=} \{x \in \mathbb{R} : x \geq a\} \quad \text{is called} \quad \textit{a closed unbounded interval},$$
$$\mathbb{R}_{> a} \stackrel{\text{def}}{=} \{x \in \mathbb{R} : x > a\} \quad \text{is called} \quad \textit{an open unbounded interval}$$
$$\mathbb{R}_{\leq a} \stackrel{\text{def}}{=} \{x \in \mathbb{R} : x \leq a\} \quad \text{is called} \quad \textit{an unbounded closed interval},$$
$$\mathbb{R}_{< a} \stackrel{\text{def}}{=} \{x \in \mathbb{R} : x < a\} \quad \text{is called} \quad \textit{an unbounded open interval}.$$

The entire set of real numbers $\mathbb{R}$ is itself a special unbounded interval.

Geometric illustrations of these intervals are given in Figures 4.3.1 through 4.3.8.

Fig. 4.3.1. A bounded closed interval

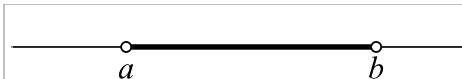


Fig. 4.3.2. An bounded open interval

Fig. 4.3.3. A bounded half-open int.

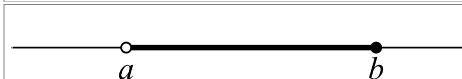


Fig. 4.3.4. A bounded half-open int.

Fig. 4.3.5. An closed unbounded int.

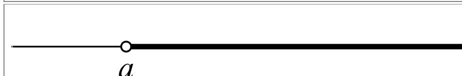


Fig. 4.3.6. An open unbounded int.

Fig. 4.3.7. An unbounded closed int.

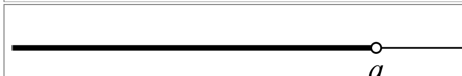


Fig. 4.3.8. An unbounded open int.

Remark 4.3.2. (A Long Remark on Notation) Common notation for unbounded intervals is as follows:

$$\mathbb{R}_{\leq a} = (-\infty, a], \quad \mathbb{R}_{< a} = (-\infty, a), \quad \mathbb{R} = (-\infty, \infty),$$
$$\mathbb{R}_{\geq a} = [a, \infty), \qquad \mathbb{R}_{> a} = (a, \infty).$$

Although the notation using $\mathbb{R}$ with a subscript is not that common, I prefer it. One reason for this preference is that a superficial reader seeing the infinity symbol, ∞, next to a real number a might mistakenly deduce that infinity is also a special kind of real number.

It might be too much to expect from a superficial reader to explore how treating ∞ as a real number would lead to contradictions in our axiomatic system.

What property must ∞ have if it were considered a real number in $\mathbb{R}$? Certainly, the property

$$\forall x \in \mathbb{R} \qquad x \leq \infty, \tag{4.3.1}$$

is essential if ∞ were to be considered a real number. By Axiom AE we would have $\infty + 1 \in \mathbb{R}$. Recall that in Theorem 4.1.8(v) we proved that $0 < 1$. Assuming $\infty \in \mathbb{R}$, by Axiom OA, applied to $0 < 1$, it would follow that $\infty + 0 < \infty + 1$, thus contradicting (4.3.1).

The infinity symbols $-\infty$ and ∞ are used in calculus to indicate that the set is unbounded in the negative ($-\infty$) or positive (∞ or $+\infty$) direction of the real number line. These symbols are just symbols; they are not real numbers. $\triangleleft$

EXERCISE 4.3.3. Let $a \in \mathbb{R}$ and $\epsilon > 0$. The following statements hold:

(i) $\forall x \in \mathbb{R} \qquad |x - a| < \epsilon \quad \Leftrightarrow \quad a - \epsilon < x < a + \epsilon.$

(ii) $\big\{ x \in \mathbb{R} : |x - a| < \epsilon \big\} = (a - \epsilon, a + \epsilon).$ $\triangleleft$

REMARK 4.3.4. (A MNEMONIC TOOL: BBB-PRINCIPLE) The content of Exercise 4.3.3 I call the BBB-principle. The reason for this nickname comes from the local geography. This class takes place in Bellingham, South of Bellingham on I-5 is Burlington and North of Bellingham on I-5 is Blaine. It just happens that both Burlington and Blaine are around 20 miles from Bellingham. In the inequality $|x - a| < \epsilon$, the symbol a stands for Bellingham, ϵ stands for 20 miles, and the inequality is asking: Where is a point x on I-5 which is at the distance 20 miles from Bellingham? A common sense answer is: That point is between Burlington and Blaine. That is exactly what the inequality $a - \epsilon < x < a + \epsilon$ is saying; it is saying that x is North of Burlington $a - \epsilon$ and South of Blaine $a + \epsilon$. This is just a mnemonic tool to internalize this important equivalence. $\triangleleft$

4.3.2. Intersections and unions of infinite families of intervals. In this subsection we apply definitions presented in Subsection 3.1.4 to infinite families of intervals. Recall that when I talk about sets whose members are sets, then I use 'family' as a synonym for set. So, 'infinite families of intervals' means 'infinite sets of intervals'.

EXERCISE 4.3.5. Let $a \in \mathbb{R}$. Prove that

$$\bigcap \big\{ (a - u, a + u) \ : \ u > 0 \big\} = \{a\}. \qquad \triangleleft$$

SOLUTION. (A very detailed solution.) What is the meaning of the expression $\bigcap \big\{ (a - u, a + u) : u > 0 \big\}$? Here $\big\{ (a - u, a + u) : u > 0 \big\}$ stands for a family of open intervals. For every $u > 0$ we are given an interval $(a - u, a + u)$. This is a set of sets, which we often call a family of sets.

Recall the definition of the intersection of a family of sets $\mathcal{A}$. The meaning of the expression $\bigcap \big\{ A : A \in \mathcal{A} \big\}$ is as follows

$$x \in \bigcap \big\{ A : A \in \mathcal{A} \big\} \quad \Leftrightarrow \quad \forall A \in \mathcal{A} \quad x \in A.$$

Thus, to prove

$$\bigcap \big\{ (a - u, a + u) : u > 0 \big\} = \{a\},$$

we have to prove the following equivalence

$$x \in \bigcap \big\{ (a - u, a + u) : u > 0 \big\} \quad \Leftrightarrow \quad x = a.$$

In other words, we have to prove

$$\forall u > 0 \quad x \in (a-u, a+u) \quad \Leftrightarrow \quad x = a.$$

Let us prove the preceding equivalence in two steps. First we prove $\Leftarrow$, then we prove $\Rightarrow$.

Assume that $x = a$. Let $u > 0$ be arbitrary. By Axiom OA, adding a to both sides of $0 < u$ we get $a < a + u$. Adding $a - u$ to both sides of $0 < u$ we get $a - u < a$. Hence $x = a \in (a-u, a+u)$. Since $u > 0$ was arbitrary we proved $\Leftarrow$.

Now we prove $\Rightarrow$. Assume that

$$\forall u > 0 \qquad x \in (a-u, a+u).$$

Then

$$\forall u > 0 \qquad a - u < x \ \wedge \ x < a + u.$$

We can rewrite the last quantified statement as two quantified statements

$$\forall u > 0 \quad a - u < x \qquad \text{and} \qquad \forall u > 0 \quad x < a + u.$$

Applying Axiom OA, from the preceding two statements we obtain

$$\forall u > 0 \quad a - x < u \qquad \text{and} \qquad \forall u > 0 \quad x - a < u.$$

Now recall Exercise 4.1.20,

$$\forall u > 0 \quad \alpha \leq u \quad \Rightarrow \quad \alpha \leq 0.$$

Therefore

$$\forall u > 0 \quad a - x < u \quad \Rightarrow \quad a - x \leq 0$$

and

$$\forall u > 0 \quad x - a < u \quad \Rightarrow \quad x - a \leq 0.$$

Thus, we have proved $a \leq x$ and $x \leq a$. Consequently, $x = a$. In conclusion, we have proved

$$\forall u > 0 \quad x \in (a-u, a+u) \quad \Rightarrow \quad x = a. \qquad \square$$

Exercise 4.3.6. Let $a, b \in \mathbb{R}$ and $a < b$. Prove that

$$\bigcap \bigl\{(a, b+u) \,:\, u > 0\bigr\} = (a, b]. \qquad \triangleleft$$

Exercise 4.3.7. Let $a, b \in \mathbb{R}$ and $a < b$. Prove that

$$\bigcap \Bigl\{(a-u, b+u) \,:\, u > 0\Bigr\} = [a, b]. \qquad \triangleleft$$

Solution. Denote by A the intersection in the equality and assume $x \in A$. Then, by the definition of intersection, $x \in (a-u, b+u)$ for all $u > 0$. By the definition of an open interval, $a - u < x$ and $x < b + u$ for all $u > 0$. Hence, $a - x < u$ and $x - b < u$ for all $u > 0$. By Exercise 4.1.20 we have $a - x \leq 0$ and $x - b \leq 0$, that is, $a \leq x$ and $x \leq b$. By the definition of a closed interval $x \in [a, b]$. This proves $A \subseteq [a, b]$.

Now assume that $x \in [a, b]$. Then, $a - x \leq 0$ and $x - b \leq 0$. Let $u > 0$ be arbitrary. By the transitivity of the order in $\mathbb{R}$, $a - x < u$ and $x - b < u$ for all $u > 0$. Hence, $a - u < x$ and $x < b + u$ for all $u > 0$. Consequently, $x \in (a-u, b+u)$ for all $u > 0$. Therefore, $x \in A$. This proves $[a, b] \subseteq A$.

Since we proved both $A \subseteq [a, b]$ and $[a, b] \subseteq A$, the equality $A = [a, b]$ is proved. $\square$

EXERCISE 4.3.8. Let $a, b \in \mathbb{R}$ and $a < b$. Prove that

$$\bigcup \Big\{ [a+u, b) \,:\, 0 < u < b-a \Big\} = (a, b). \qquad \triangleleft$$

EXERCISE 4.3.9. Let $a, b \in \mathbb{R}$ and $a < b$. Prove that

$$\bigcup \Big\{ [a+u, b-u] \,:\, 0 < u < \frac{b-a}{2} \Big\} = (a, b). \qquad \triangleleft$$

4.3.3. Cardinality of intervals. The goal here is to prove that any two intervals have the same cardinality.

THEOREM 4.3.10. *Let I and J be any two of the nine types of intervals defined in Definition 4.3.1. Then there exists a bijection $f : I \to J$.*

We start with three specific bijections in the following two lemmas and a corollary. In the lemmas, we use the set of positive integers $\mathbb{N}$ and the set of integers $\mathbb{Z}$ which will be formally introduced in Chapter 5.

LEMMA 4.3.11. *The function $\phi : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$, defined by the following piecewise formula, is a bijection:*

$$\forall x \in \mathbb{R}_{>0} \qquad \phi(x) = \begin{cases} x & \textit{if} \quad x \in \mathbb{R}_{>0} \setminus \mathbb{N}, \\ x - 1 & \textit{if} \quad x \in \mathbb{N}. \end{cases}$$

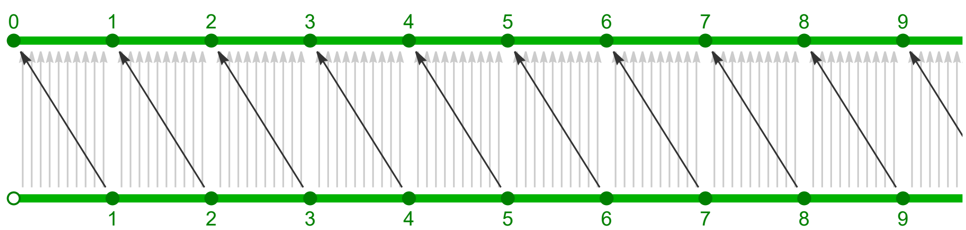


FIG. 4.3.9. The bijection $\phi : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$

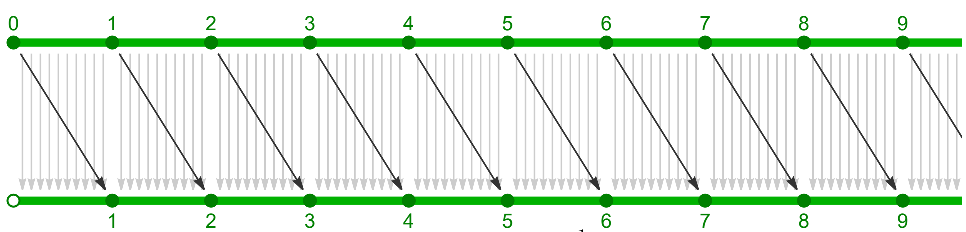


FIG. 4.3.10. The bijection $\phi^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$

PROOF. The inverse of the function $\phi : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ defined in the lemma is

$$\forall x \in \mathbb{R}_{\geq 0} \qquad \phi^{-1}(x) = \begin{cases} x & \text{if} \quad x \in \mathbb{R}_{>0} \setminus \mathbb{N}, \\ x + 1 & \text{if} \quad x \in \mathbb{N} \cup \{0\}, \end{cases}$$

which is a bijection with domain $\mathbb{R}_{\geq 0}$ and range $\mathbb{R}_{>0}$. This claim is proved by straightforward algebra. For visualisations of the function $\phi : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ and its inverse $\phi^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$ Figures 4.3.9, 4.3.10, Figures 4.3.11, and 4.3.12.

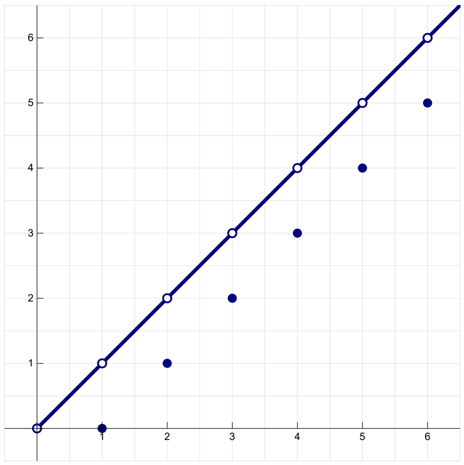


FIG. 4.3.11. $\phi : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$

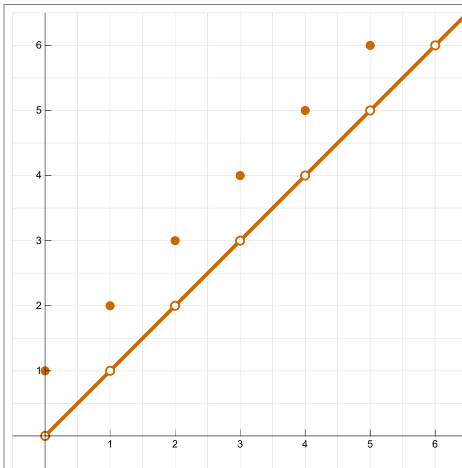


FIG. 4.3.12. $\phi^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$

It is helpful to notice that the sets $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ can be represented as disjoint unions as follows:

$$\mathbb{R}_{>0} = \big(\mathbb{R}_{>0} \setminus \mathbb{N}\big) \cup \mathbb{N},$$
$$\mathbb{R}_{\geq 0} = \big(\mathbb{R}_{>0} \setminus \mathbb{N}\big) \cup \big(\mathbb{N} \cup \{0\}\big).$$

The function ϕ on the set $\mathbb{R}_{>0} \setminus \mathbb{N}$, that is

$$\phi : \mathbb{R}_{>0} \setminus \mathbb{N} \;\to\; \mathbb{R}_{>0} \setminus \mathbb{N},$$

acts as the identity, so it is a bijection. The function ϕ restricted to the domain $\mathbb{N}$, that is

$$\phi : \mathbb{N} \;\to\; \{0\} \cup \mathbb{N},$$

is also a bijection, so called 'backward shift'. □

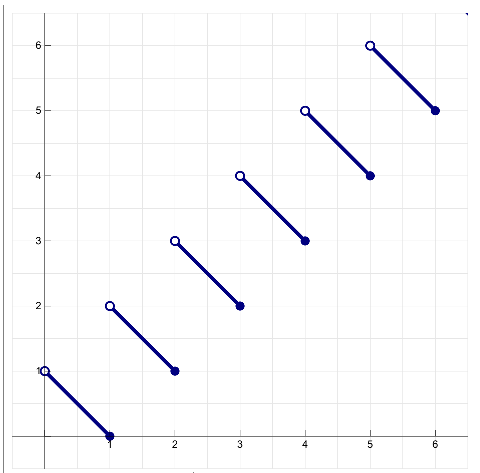


FIG. 4.3.13. A bijection $\mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$

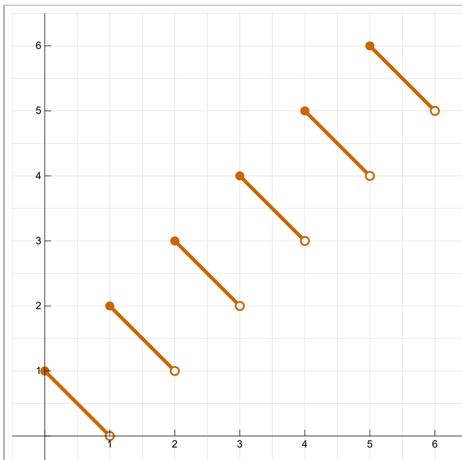


FIG. 4.3.14. A bijection $\mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$

REMARK 4.3.12. I can not move on to the next lemma without recording another beautiful formula for a bijection with domain $\mathbb{R}_{>0}$ and range $\mathbb{R}_{\geq 0}$, given by the formula:

$$\forall x \in \mathbb{R}_{>0} \qquad x \mapsto 2\lceil x \rceil - 1 - x,$$

and its inverse with domain $\mathbb{R}_{\geq 0}$ and range $\mathbb{R}_{>0}$ is given by the formula:

$$\forall x \in \mathbb{R}_{\geq 0} \qquad x \mapsto 2\lfloor x \rfloor + 1 - x.$$

We will talk about the functions floor, denoted by $\lfloor x \rfloor$, and ceiling, denoted by $\lceil x \rceil$, in Section 6.2. The claim that the above functions are inverses of each other follows from the properties of the floor and ceiling. The beauty of these functions is revealed in their graphs, see Figures 4.3.13 and 4.3.14. ◁

LEMMA 4.3.13. *The function $\psi : \mathbb{R} \to [-1, 1]$ defined by the following formula is a bijection:*

$$\forall x \in \mathbb{R} \quad \psi(x) = \frac{\operatorname{sgn}(x)}{\phi(|x|) + 1} = \begin{cases} 0 & \text{if } \ x = 0, \\ \dfrac{1}{x + \operatorname{sgn}(x)} & \text{if } \ x = \mathbb{R} \setminus \mathbb{Z}, \\ \dfrac{1}{x} & \text{if } \ x \in \mathbb{Z} \setminus \{0\}. \end{cases}$$

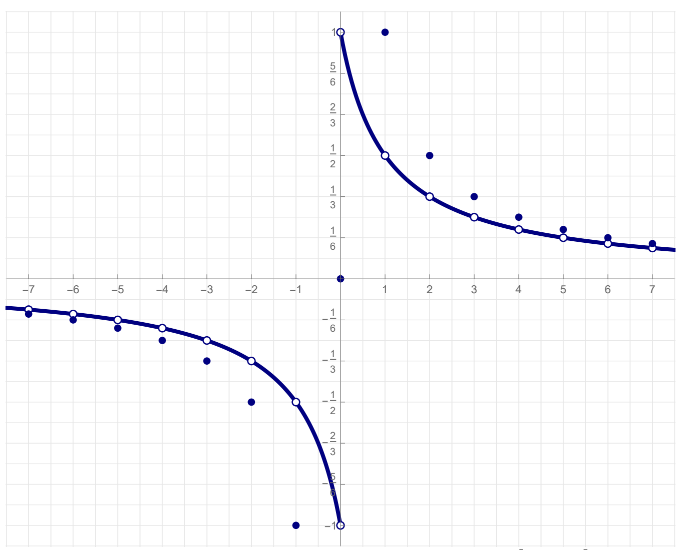


FIG. 4.3.15. The bijection $\psi : \mathbb{R} \to [-1, 1]$

PROOF. We construct the bijection $\psi : \mathbb{R} \to [-1, 1]$ in two steps. We use the following representations of $\mathbb{R}$ and $[-1, 1]$ as disjoint unions:

$$\mathbb{R} = \mathbb{R}_{<0} \cup \{0\} \cup \mathbb{R}_{>0},$$
$$[-1, 1] = [-1, 0) \cup \{0\} \cup (0, 1].$$

Next we construct a bijection with domain $\mathbb{R}_{>0}$ and $(0, 1]$. This is done in two steps:

$$\mathbb{R}_{>0} \xrightarrow{\ x \mapsto \phi(x)\ } \mathbb{R}_{\geq 0} \xrightarrow{\ x \mapsto \frac{1}{x+1}\ } (0, 1].$$

Hence the composition is

$$\forall x \in \mathbb{R}_{>0} \qquad x \mapsto \frac{1}{\phi(x)+1},$$

which is a bijection with domain $\mathbb{R}_{>0}$ and range $(0,1]$.

Next we construct a bijection with domain $\mathbb{R}_{<0}$ and $[-1,0)$. This is done in two steps:

$$\mathbb{R}_{<0} \xrightarrow{\;x \mapsto -\phi(-x)\;} \mathbb{R}_{\leq 0} \xrightarrow{\;x \mapsto \frac{1}{x-1}\;} [-1,0).$$

Hence the composition is

$$\forall x \in \mathbb{R}_{<0} \quad x \mapsto \frac{1}{-\phi(-x)-1} = \frac{-1}{\phi(-x)+1},$$

which is a bijection with domain $\mathbb{R}_{<0}$ and range $[-1,0)$.

These two formulas can be unified as follows: The function defined as

$$\forall x \in \mathbb{R} \quad \psi(x) = \frac{\operatorname{sgn}(x)}{\phi(|x|)+1} = \begin{cases} 0 & \text{if } \; x = 0, \\ \dfrac{1}{x+\operatorname{sgn}(x)} & \text{if } \; x = \mathbb{R}\setminus\mathbb{Z}, \\ \dfrac{1}{x} & \text{if } \; x \in \mathbb{Z}\setminus\{0\}, \end{cases}$$

is a bijection with domain $\mathbb{R}$ and range $[-1,1]$. Its inverse can be calculated to be

$$\forall x \in [-1,1] \quad \psi^{-1}(x) = \operatorname{sgn}(x)\,\phi^{-1}\!\left(\frac{1}{|x|}-1\right) = \begin{cases} 0 & \text{if } \; x = 0, \\ \dfrac{1}{x} - \operatorname{sgn}(x) & \text{if } \; x = \mathbb{R}\setminus\mathbb{Z}, \\ x & \text{if } \; \dfrac{1}{x} \in \mathbb{Z}\setminus\{0\}, \end{cases}$$

which is a bijection with domain $[-1,1]$ and range $\mathbb{R}$. In the first definition of Φ^{-1} the expression involving ϕ^{-1} is not defined for $x=0$, but, for $x=0$ we have $\operatorname{sgn}(0)=0$, that is the undefined quantity is multiplied by 0. In the definition of Φ^{-1}, for simplicity, we adopt a common convention that if an undefined expression is multiplied by 0, the entire expression assumes the value 0. □

Corollary 4.3.14. *The restriction of the bijection ψ from Lemma 4.3.13 to $\mathbb{R}_{\geq 0}$, see Figure 4.3.16, is a bijection with domain $\mathbb{R}_{\geq 0}$ and range $[0,1]$. The function $\psi_r : \mathbb{R}_{\geq 0} \to [0,1]$ is defined by the following piecewise formula:*

$$\forall x \in \mathbb{R}_{\geq 0} \qquad \psi_r(x) = \begin{cases} 0 & \textit{if } \; x = 0, \\ \dfrac{1}{x+1} & \textit{if } \; x = \mathbb{R}_{>0}\setminus\mathbb{N}, \\ \dfrac{1}{x} & \textit{if } \; x \in \mathbb{N}. \end{cases}$$

Proof of Theorem 4.3.10. **Part I.** In this part we consider closed intervals. Let $a,b,c,d \in \mathbb{R}$ be such that $a<b$ and $c<d$. The function

$$\forall x \in [a,b] \qquad x \mapsto \frac{d-c}{b-a}(x-a)+c,$$

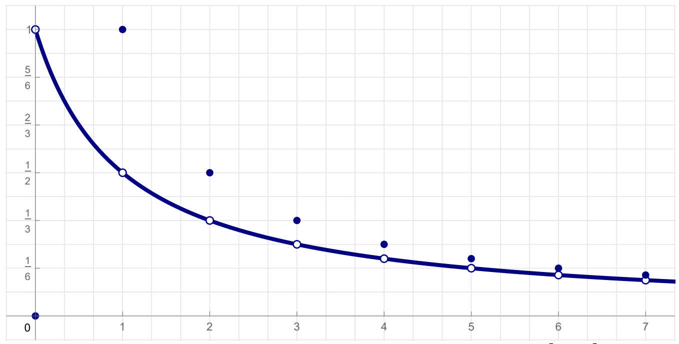


FIG. 4.3.16. The bijection $\psi_r : \mathbb{R}_{\geq 0} \to [0, 1]$

is a bijection with domain $[a, b]$ and range $[c, d]$. Its inverse is

$$\forall x \in [c, d] \qquad x \mapsto \frac{b-a}{d-c}(x-c)+a,$$

which is a bijection with domain $[c, d]$ and range $[a, b]$. These claims are proved by straightforward algebra.

In this way, we have constructed a bijection between any two closed intervals.

Part II. In this part we consider open intervals. Let $a, b, c, d \in \mathbb{R}$ be such that $a < b$ and $c < d$. There are four kinds of open intervals:

(i) The open interval of its own kind is $\mathbb{R}$.

(ii) Finite open intervals, like (a, b) and (c, d). The formulas given in Part I, with adjusted domains provide bijections between any two finite open intervals.

(iii) Right-infinite open intervals, like $\mathbb{R}_{>a}$ and $\mathbb{R}_{>c}$. The function

$$\forall x \in \mathbb{R}_{>a} \qquad x \mapsto (x-a)+c,$$

is a bijection with domain $\mathbb{R}_{>a}$ and range $\mathbb{R}_{>c}$. Its inverse is

$$\forall x \in \mathbb{R}_{>c} \qquad x \mapsto (x-c)+a,$$

which is a bijection with domain $\mathbb{R}_{>c}$ and range $\mathbb{R}_{>a}$. These claims are proved by straightforward algebra.

In this way, we have constructed a bijection between any two right-infinite open intervals.

(iv) Left-infinite open intervals, like $\mathbb{R}_{<a}$ and $\mathbb{R}_{<c}$. Straightforward modifications of the bijections in (iii) provide bijections between any two left-infinite open intervals.

Next we provide bijections between selected typical representatives of intervals in each of the kinds presented in (i), (ii), (iii), (iv). The typical representative in (i) is the only one, $\mathbb{R}$. We select the typical representative in (ii) to be $(-1, 1)$, in (iii) to be $\mathbb{R}_{>0}$, and in (iv) to be $\mathbb{R}_{<0}$.

The promised bijections are as follows:

(i)-(ii). A bijection with domain $\mathbb{R}$ and range $(-1, 1)$, see Figure 4.3.17, is given by

$$\forall x \in \mathbb{R} \qquad x \mapsto \frac{x}{1+|x|}.$$

Its inverse is

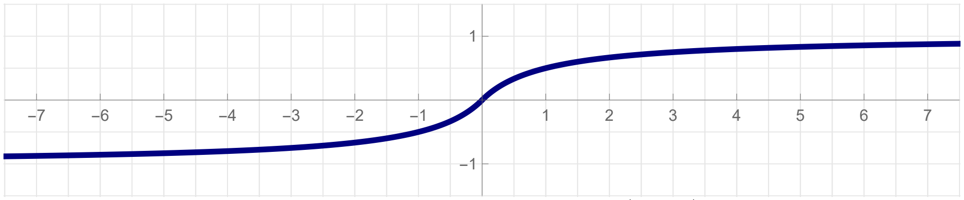


FIG. 4.3.17. A bijection $\mathbb{R} \to (-1, 1)$

$$\forall x \in (-1,1) \qquad x \mapsto \frac{x}{1-|x|},$$

which is a bijection with domain $(-1, 1)$ and range $\mathbb{R}$. These claims are proved by straightforward algebra.

(iii)-(ii). A bijection with domain $\mathbb{R}_{>0}$ and range $(-1, 1)$, see Figure 4.3.18, is given by

$$\forall x \in \mathbb{R}_{>0} \qquad x \mapsto \frac{x-1}{x+1}.$$

Its inverse is

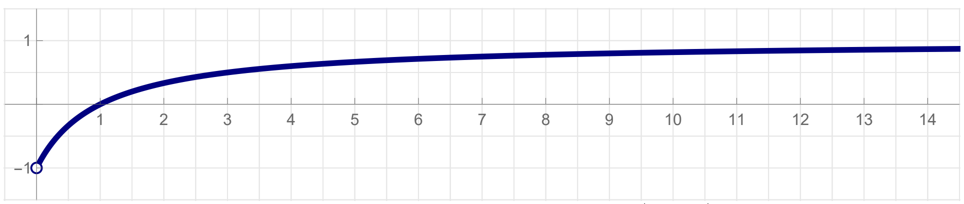


FIG. 4.3.18. A bijection $\mathbb{R}_{>0} \to (-1, 1)$

$$\forall x \in (-1,1) \qquad x \mapsto \frac{1+x}{1-x},$$

which is a bijection with domain $(-1, 1)$ and range $\mathbb{R}_{>0}$. These claims are proved by straightforward algebra.

(iii)-(iv). A bijection with domain $\mathbb{R}_{>0}$ and range $\mathbb{R}_{<0}$ is given by

$$\forall x \in \mathbb{R}_{>0} \qquad x \mapsto -x.$$

Its inverse is

$$\forall x \in \mathbb{R}_{<0} \qquad x \mapsto -x,$$

which is a bijection with domain $\mathbb{R}_{<0}$ and range $\mathbb{R}_{>0}$. These claims are proved by straightforward algebra.

Forming compositions of bijections presented in **Part II**, we can construct a bijection between any two open intervals.

Part III. In this part we consider half-open intervals. Let $a, b, c, d \in \mathbb{R}$ be such that $a < b$ and $c < d$. There are four kinds of half-open intervals:

(a) Finite, left-open intervals, like $(a, b]$ and $(c, d]$. The formulas given in Part I, with adjusted domains, provide bijections between these intervals.
(b) Finite, right-open intervals, like $[a, b)$ and $[c, d)$. The formulas given in Part I, with adjusted domains, provide bijections between these intervals.
(c) Right-infinite half-open intervals, like $\mathbb{R}_{\geq a}$ and $\mathbb{R}_{\geq c}$. The formulas given in Part II item (iii), with adjusted domains, provide bijections between these intervals.

(d) Left-infinite half-open intervals, like $\mathbb{R}_{\leq a}$ and $\mathbb{R}_{\leq c}$. The formulas given in Part II item (iv), with adjusted domains, provide bijections between these intervals.

Next we provide bijections between selected typical representatives of intervals in each of the kinds presented in (a), (b), (c), (d). We select a typical representative in (a) to be $(-1, 1]$, in (b) to be $[-1, 1)$, in (c) to be $\mathbb{R}_{\geq 0}$, and in (d) is $\mathbb{R}_{\leq 0}$.

The promised bijections are as follows:

(a)-(b). A bijection with domain $(-1, 1]$ and range $[-1, 1)$ is given by the formula provided in Part II item (iii)-(iv), with adjusted domains, of this proof.

(c)-(b). A bijection with domain $\mathbb{R}_{\geq 0}$ and range $[-1, 1)$ is given by the formula provided in Part II item (iii)-(ii), with adjusted domains, of this proof.

(c)-(d). A bijection between with domain $\mathbb{R}_{\geq 0}$ and range $\mathbb{R}_{\leq 0}$ is given by the formula in Part II item (iii)-(iv), with adjusted domains, of this proof.

Forming compositions of bijections presented in **Part III**, we can construct a bijection between any two half-open intervals.

Conclusion. Let I and J be arbitrary intervals. If I and J are both closed, then a bijection between them is given in Part I. If I and J are both open, then a bijection between them is given in Part II. If I and J are both half-open, then a bijection between them is given in Part III.

If I is open and J is closed, then Part II provides a bijection

$$g : I \to \mathbb{R},$$

and Part I provides a bijection

$$h : [-1, 1] \to J.$$

With the bijection $\psi : \mathbb{R} \to [-1, 1]$ from Lemma 4.3.13, consider the composition

$$I \xrightarrow{g} \mathbb{R} \xrightarrow{\psi} [-1, 1] \xrightarrow{h} J.$$

By Theorem 3.2.18, the above composition, that is the function

$$f = h \circ \psi \circ g : I \to J,$$

is a desired bijection.

If I is open and J is half-open, then Part II provides a bijection

$$g : I \to \mathbb{R}_{>0},$$

and Part III provides a bijection

$$h : \mathbb{R}_{\geq 0} \to J.$$

With the bijection $\phi : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ from Lemma 4.3.11, consider the composition

$$I \xrightarrow{g} \mathbb{R}_{>0} \xrightarrow{\phi} \mathbb{R}_{\geq 0} \xrightarrow{h} J.$$

By Theorem 3.2.18, the above composition, that is the function

$$f = h \circ \phi \circ g : I \to J,$$

is a desired bijection.

If I is half-open and J is closed, then Part III provides a bijection

$$g : I \to \mathbb{R}_{\geq 0},$$

and Part I provides a bijection

$$h : [0,1] \to J.$$

With the bijection $\psi_r : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ from Corollary 4.3.14, consider the composition

$$I \xrightarrow{\ g\ } \mathbb{R}_{\geq 0} \xrightarrow{\ \psi_r\ } [0,1] \xrightarrow{\ h\ } J.$$

By Theorem 3.2.18, the above composition, that is the function

$$f = h \circ \psi_r \circ g : I \to J,$$

is a desired bijection.

For any combination of kinds of intervals I and J we constructed a bijection with domain I and range J. Thus, the theorem is proved. □

REMARK 4.3.15. In this remark, I present two bijections that I considered for the proof of Theorem 4.3.10, but they were not needed in the end.

A specific bijection with domain $\mathbb{R}$ and range $\mathbb{R}_{>0}$ in Part II of the proof of Theorem 4.3.10 is the following function, see Figure 4.3.19:

$$\forall x \in \mathbb{R} \qquad x \mapsto \frac{1+\operatorname{rlu}(x)}{1+\operatorname{rlu}(-x)} = \begin{cases} \dfrac{1}{1-x} & \text{if} \quad x \in \mathbb{R}_{<0}, \\ 1+x & \text{if} \quad x \in \mathbb{R}_{\geq 0}, \end{cases}$$

with domain $\mathbb{R}$ and range $\mathbb{R}_{>0}$. Its inverse is

$$\forall x \in \mathbb{R}_{>0} \qquad x \mapsto \frac{1-x}{\operatorname{rlu}(1-x)-1} = \begin{cases} 1-\dfrac{1}{x} & \text{if} \quad x \in (0,1], \\ x-1 & \text{if} \quad x \in \mathbb{R}_{>1}, \end{cases}$$

which is a bijection with domain $\mathbb{R}_{>0}$ and range $\mathbb{R}$. It is desirable to avoid using piecewise formulas. The only way I found to express the above two functions without using piecewise formulas was by using Rectified Linear Unit, or ReLU, function, denoted here as rlu.

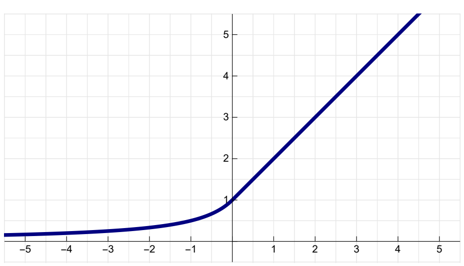


FIG. 4.3.19. A bijection $\mathbb{R} \to \mathbb{R}_{>0}$

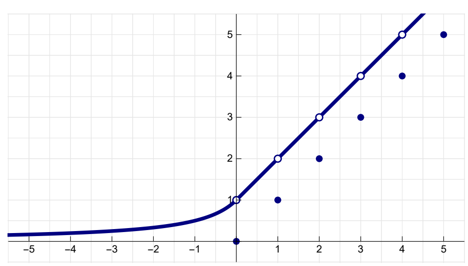


FIG. 4.3.20. A bijection $\mathbb{R} \to \mathbb{R}_{\geq 0}$

The composition of the preceding bijection $\mathbb{R} \to \mathbb{R}_{>0}$ and the bijection $\phi : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ from Lemma 4.3.11, that is the function

$$\mathbb{R} \xrightarrow{\ x \mapsto \frac{1+\operatorname{rlu}(x)}{1+\operatorname{rlu}(-x)}\ } \mathbb{R}_{>0} \xrightarrow{\ x \mapsto \phi(x)\ } \mathbb{R}_{\geq 0},$$

is the following function, see Figure 4.3.20:

$$\forall x \in \mathbb{R} \qquad x \mapsto= \phi\left(\frac{1+\mathrm{rlu}(x)}{1+\mathrm{rlu}(-x)}\right) = \begin{cases} \dfrac{1}{1-x} & \text{if} \quad x \in \mathbb{R}_{<0}, \\ x & \text{if} \quad x \in \{0\} \cup \mathbb{N}, \\ 1+x & \text{if} \quad x \in \mathbb{R}_{>0} \setminus \mathbb{N}. \end{cases} \qquad \triangleleft$$

4.4. Minimums, maximums, and boundedness of sets in $\mathbb{R}$

4.4.1. Minimums, maximums. In Definition 4.2.7, we introduced the concepts of the minimum and the maximum for subsets of $\mathbb{R}$ of the form $\{a, b\} \subset \mathbb{R}$. For such subsets, the minimum and the maximum always exist and are explicitly defined in Definition 4.2.7.

We now extend these concepts to arbitrary nonempty subsets of $\mathbb{R}$. As we will see, it is common for such sets to lack a minimum, a maximum, or both.

DEFINITION 4.4.1. Let A be a nonempty subset of $\mathbb{R}$. A number $a \in \mathbb{R}$ is a *minimum* of A if it satisfies the following two conditions:

(i) For all $x \in A$, we have $a \leq x$.
(ii) $a \in A$.

The minimum of A, if it exists, is denoted by $\min(A)$.

DEFINITION 4.4.2. Let A be a nonempty subset of $\mathbb{R}$. A number $b \in \mathbb{R}$ is a *maximum* of A if it satisfies the following two conditions:

(i) For all $x \in A$, we have $x \leq b$.
(ii) $b \in A$.

The maximum of A, if it exists, is denoted by $\max(A)$.

REMARK 4.4.3. (ALWAYS CONSIDER THE NEGATIONS) What does it mean for a nonempty subset of $\mathbb{R}$ not to have a minimum? To answer this question, we first restate Definition 4.4.1 as follows: a nonempty set A has a minimum if and only if

$$\exists\, a \in A \quad \text{such that} \quad \forall\, x \in A \quad \text{we have} \quad x \geq a. \tag{4.4.1}$$

The negation of (4.4.1) is:

$$\forall\, a \in A \quad \exists\, x \in A \quad \text{such that} \quad x < a. \tag{4.4.2}$$

Notice that the number x in (4.4.2) depends on a. It is sometimes helpful to emphasize this dependence by writing $x(a)$. A more precise version of the negation is:

$$\forall\, a \in A \quad \exists\, x(a) \in A \quad \text{such that} \quad x(a) < a.$$

What does it mean for a nonempty subset of $\mathbb{R}$ not to have a maximum? To answer this, we restate Definition 4.4.2: a nonempty set A has a maximum if and only if

$$\exists\, b \in A \quad \text{such that} \quad \forall\, x \in A \quad \text{we have} \quad x \leq b. \tag{4.4.3}$$

The negation of (4.4.3) is:

$$\forall b \in A \quad \exists x \in A \quad \text{such that} \quad b < x. \tag{4.4.4}$$

Again, the number x depends on b, so we may write $x(b)$. A more precise version is:

$$\forall b \in A \quad \exists x(b) \in A \quad \text{such that} \quad b < x(b). \qquad \triangleleft$$

Exercise 4.4.4. Prove that the set of all positive real numbers has neither a minimum nor a maximum. ◁

Exercise 4.4.5. Give examples of subsets of $\mathbb{R}$ such that:

(a) The set has neither a minimum nor a maximum.
(b) The set has a minimum but no maximum.
(c) The set has both a minimum and a maximum. ◁

Exercise 4.4.6. Let A be a nonempty subset of $\mathbb{R}$, and let $a \in \mathbb{R}$. Prove the following equivalence: $a = \min A$ and $a = \max A$ if and only if $A = \{a\}$. That is, prove

$$(a = \min A) \wedge (a = \max A) \quad \Leftrightarrow \quad A = \{a\}. \qquad \triangleleft$$

4.4.2. Boundedness. It is common for a nonempty subset of $\mathbb{R}$ to have neither a minimum nor a maximum. Are there any surrogates? That question is answered in this subsection.

Definition 4.4.7. Let A be a nonempty subset of $\mathbb{R}$. If there exists $a \in \mathbb{R}$ such that

$$\forall x \in A \quad \text{we have} \quad a \leq x, \tag{4.4.5}$$

then A is said to be *bounded below*. A number a satisfying (4.4.5) is called a *lower bound* of A.

Definition 4.4.8. Let A be a nonempty subset of $\mathbb{R}$. If there exists $b \in \mathbb{R}$ such that

$$\forall x \in A \quad \text{we have} \quad x \leq b, \tag{4.4.6}$$

then A is said to be *bounded above*. A number b satisfying (4.4.6) is called an *upper bound* of A.

Definition 4.4.9. A nonempty subset of $\mathbb{R}$ that is both bounded above and bounded below is said to be *bounded*.

Exercise 4.4.10. Let A be a nonempty subset of $\mathbb{R}$. Prove that A is bounded if and only if there exists $K > 0$ such that $-K \leq x \leq K$ for all $x \in A$. ◁

Solution. Let A be a nonempty subset of $\mathbb{R}$. Assume that A is bounded. Then there exist $a, b \in \mathbb{R}$ such that for all $x \in A$ we have $a \leq x \leq b$. Set $K = 1 + \max\{|a|, |b|\}$. Then $K \geq 1 > 0$. Also, $K > |b|$ and $-K < -|a|$. Since $b \leq |b|$ and $-|a| \leq a$, for all $x \in A$ we have

$$-K < -|a| \leq a \leq x \leq b \leq |b| < K.$$

Hence, $K > 0$ has the desired property. The converse is straightforward. □

Exercise 4.4.11. Let A be a nonempty subset of $\mathbb{R}$. Prove that A is bounded if and only if there exist $a, b \in \mathbb{R}$, such that $a < b$ and $A \subseteq [a, b]$. ◁

EXERCISE 4.4.12. Prove that $\left\{x \in \mathbb{R} : x^2 < 2\right\}$ is a bounded set. $\triangleleft$

EXERCISE 4.4.13. Let A and B be bounded above subsets of $\mathbb{R}$. Prove that $A \cup B$ is bounded above. $\triangleleft$

CHAPTER 5

The subsets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ of $\mathbb{R}$

5.1. The set $\mathbb{N}$

We mentioned natural numbers and integers informally in the course of our discussion of the fundamental properties of $\mathbb{R}$. Notice again that the only numbers that are specifically mentioned in Axioms 1 through 15 are 0 and 1. But, in Section 4.1.2 Exercise 4.1.11 we proved that there are other numbers in $\mathbb{R}$, and we defined the numbers $2, 3, 4, 5, 6, 7, 8$ and 9. The reason that we stopped at 9 is the fact that the number $9 + 1$ plays a special role in our culture. We could continue this process further, but it would not lead to a rigorous definition of the set of natural numbers. Therefore we chose a different route.

5.1.1. The definition of the set $\mathbb{N}$. Consider the following two properties of a subset S of $\mathbb{R}$:

$$1 \in S, \tag{5.1.1}$$

$$n \in S \quad \Rightarrow \quad n + 1 \in S. \tag{5.1.2}$$

There are many subsets of $\mathbb{R}$ that have these two properties. For example, one such set is the set of positive real numbers, that is the open infinite interval,

$$\mathbb{R}_{>0}.$$

Another such set is the closed infinite interval

$$\mathbb{R}_{\geq 1},$$

and also the union

$$\{1\} \cup \mathbb{R}_{\geq 2}.$$

There are many such sets. Next we form the family of <u>all</u> subsets of $\mathbb{R}$ with the properties (5.1.1) and (5.1.2):

$$\mathcal{N} \stackrel{\text{def}}{=} \Bigl\{ S \subseteq \mathbb{R} \, : \, 1 \in S \;\text{ and }\; n \in S \;\Rightarrow\; n + 1 \in S \Bigr\}$$

Intuitively, the set of natural numbers is the smallest set in $\mathcal{N}$.

DEFINITION 5.1.1. We define $\mathbb{N}$ to be the intersection of the family $\mathcal{N}$:

$$\mathbb{N} \stackrel{\text{def}}{=} \bigcap \bigl\{ S \, : \, S \in \mathcal{N} \bigr\}.$$

That is, $k \in \mathbb{N}$ if and only if $k \in S$ for all $S \in \mathcal{N}$. The elements of the set $\mathbb{N}$ are called *natural numbers* or *positive integers*.

5.1.2. Basic properties of the set $\mathbb{N}$. With this definition and Axioms 1 through 15 we should be able to prove all familiar properties of natural numbers.

THEOREM 5.1.2. *The following five statements hold:*

(N 1) $1 \in \mathbb{N}$.
(N 2) *The formula* $\sigma(n) = n + 1$ *defines a function* $\sigma : \mathbb{N} \to \mathbb{N}$.
(N 3) *If* $\sigma(m) = \sigma(n)$, *then* $n = m$; *that is* σ, *is an injection.*
(N 4) *For all* $n \in \mathbb{N}$, $\sigma(n) \neq 1$.
(N 5) *If* $K \subseteq \mathbb{N}$ *has the following two properties*

$$1 \in K,$$
$$\forall\, n \in \mathbb{N} \qquad n \in K \quad \Rightarrow \quad n + 1 \in K,$$

then $K = \mathbb{N}$.

PROOF. Since $1 \in S$ for all $S \in \mathcal{N}$, we have $1 \in \mathbb{N}$. This proves (N 1). To prove (N 2), let $n \in \mathbb{N}$ be arbitrary. Then $n \in S$ for all $S \in \mathcal{N}$. Since (5.1.2) holds for each $S \in \mathcal{N}$, we conclude that $n + 1 \in S$ for all $S \in \mathcal{N}$. Hence $n + 1 \in \mathbb{N}$ for all $n \in \mathbb{N}$. Property (N 3) follows from Theorem 4.1.4 (ii). To prove (N 5) assume that $K \subseteq \mathbb{N}$ and K has properties (5.1.1) and (5.1.2). Then $K \in \mathcal{N}$. Consequently, $\mathbb{N} = \bigcap\{S : S \in \mathcal{N}\} \subseteq K$. Thus, $K = \mathbb{N}$. □

REMARK 5.1.3. The five properties of $\mathbb{N}$ proved in Theorem 5.1.2 are known as Peano axioms. Italian mathematician Giuseppe Peano (1858-1932) used these five properties for an axiomatic foundation of natural numbers. All other familiar properties of the positive integers can be proved using these axioms. The theory of natural numbers developed from Peano's axioms is called Peano's arithmetic. ◁

An important consequence of the property (N 5) in Theorem 5.1.2 is the *Principle of Mathematical Induction*. It is stated and proved in the next theorem. This principle is the main tool in dealing with statements involving natural numbers.

THEOREM 5.1.4. *Let* $P(n)$, $n \in \mathbb{N}$, *be a family of statements such that*

(I) $P(1)$ *is true,*
(II) *For all* $n \in \mathbb{N}$ *we have* $P(n)$ *implies* $P(n + 1)$.

Then the statement $P(n)$ *is true for each* $n \in \mathbb{N}$.

PROOF. Consider the set

$$S = \bigl\{ n \in \mathbb{N} : P(n) \ \text{ is true } \bigr\}.$$

By (I), $1 \in S$. By (II), for all $n \in \mathbb{N}$, if $n \in S$, then $n + 1 \in S$. Hence, S has both properties from Theorem 5.1.2 (N 5). Consequently, $S = \mathbb{N}$. This means that for all $n \in \mathbb{N}$ the statement $P(n)$ is true. □

REMARK 5.1.5. (ON PROOFS BY MATHEMATICAL INDUCTION) Step (II) of Mathematical Induction requires us to conclude that $P(n + 1)$ is true, assuming that $P(n)$ is true. That is, we must prove the implication

$$\forall n \in \mathbb{N} \qquad P(n) \Rightarrow P(n + 1).$$

Therefore, a proof of Step (II) should always begin as follows:

Let $n \in \mathbb{N}$ be arbitrary. Assume $P(n)$.

The proof then proceeds by deducing $P(n+1)$, using all green steps that are ideally stated in the **Background Knowledge**. ◁

The following theorem can be proved using the properties from Theorem 5.1.2 and the Principle of Mathematical Induction.

THEOREM 5.1.6. *The following statements hold:*

(i) $1 = \min \mathbb{N}$*; that is,* $1 \in \mathbb{N}$ *and* $1 \leq n$ *for all* $n \in \mathbb{N}$*.*
(ii) *For every* $n \in \mathbb{N}\backslash\{1\}$*, we have* $n - 1 \in \mathbb{N}$*.*
(iii) *For all* $m, n \in \mathbb{N}$*, we have* $m + n \in \mathbb{N}$*.*
(iv) *For all* $m, n \in \mathbb{N}$ *we have* $mn \in \mathbb{N}$*.*
(v) *For all* $m, n \in \mathbb{N}$ *such that* $m < n$*, we have* $n - m \in \mathbb{N}$*.*
(vi) *If* $m, n \in \mathbb{N}$ *and* $m < n$*, then* $m + 1 \leq n$*.*

PROOF. (i) As we mentioned before, the closed infinite interval $\mathbb{R}_{\geq 1}$ belongs to the family $\mathcal{N}$. Therefore $\mathbb{N} \subseteq \mathbb{R}_{\geq 1}$. Thus $n \geq 1$ for all $n \in \mathbb{N}$. Since $1 \in \mathbb{N}$ was proved in Theorem 5.1.2, the statement in (i) is proved.

(ii) Consider the following set $S = \{1\} \cup \big\{m \in \mathbb{N} : m - 1 \in \mathbb{N}\big\}$. Clearly $S \subseteq \mathbb{N}$ and $1 \in S$. Notice also that $2 \in S$, since $2 - 1 = 1 \in \mathbb{N}$. Next we will prove

$$n \in S \quad \Rightarrow \quad n + 1 \in S. \tag{5.1.3}$$

Assume $n \in S$. We distinguish two cases: $n = 1$ and $n \in \big\{m \in \mathbb{N} : m - 1 \in \mathbb{N}\big\}$. If $n = 1$, then $n + 1 = 2 \in S$. Hence (5.1.3) holds in this case. If $n \in \big\{m \in \mathbb{N} : m - 1 \in \mathbb{N}\big\}$, then $n \in \mathbb{N}$ and $n - 1 \in \mathbb{N}$. By Theorem 5.1.2 (N 2), $n + 1 \in \mathbb{N}$ and, obviously, $(n + 1) - 1 = n \in \mathbb{N}$. Therefore $n + 1 \in \big\{m \in \mathbb{N} : m - 1 \in \mathbb{N}\big\}$. Hence $n + 1 \in S$. Thus (5.1.3) holds. Now, by Theorem 5.1.2 (N 5), $S = \mathbb{N}$. This proves $\mathbb{N}\backslash\{1\} = \big\{m \in \mathbb{N} : m - 1 \in \mathbb{N}\big\}$.

The remaining properties are proved similarly. □

5.1.3. Sequences.

DEFINITION 5.1.7. A function whose domain is $\mathbb{N}$ and whose codomain is a nonempty set S is called a *sequence in* S.

REMARK 5.1.8. (ON NOTATION). Traditionally, if $f : A \to B$ is a function and $x \in A$, then the value of f at x is denoted by $f(x)$. In addition to this standard notation, for a sequence $f : \mathbb{N} \to S$, we will often write f_n instead of $f(n)$ for $n \in \mathbb{N}$. When convenient, we will use both notations for the same sequence. The reason for this is purely typographical. For example, if $n = \dfrac{m(m+1)}{2} + 1$, then writing $f_{\frac{m(m+1)}{2}+1}$ is visually awkward. In such cases, the expression $f\big(\frac{m(m+1)}{2}+1\big)$ is preferable, as it is easier to read and understand. ◁

EXAMPLE 5.1.9. Here we give examples of sequences given by a formula. In each formula below $n \in \mathbb{N}$.

$$\text{(a)} \quad n \mapsto n, \quad \text{(b)} \quad n \mapsto n^2, \quad \text{(c)} \quad n \mapsto \frac{1}{n}, \quad \text{(d)} \quad n \mapsto \frac{1}{n^2}, \quad \text{(e)} \quad n \mapsto \frac{n}{1+n}.$$

◁

A special, probably the most often encountered kind of sequences is introduced in the next definition.

DEFINITION 5.1.10. Let S be a nonempty set and $k \in \mathbb{N}$. A sequence $\varphi : \mathbb{N} \to S$ is said to be *defined recursively of order* m if there exist m-tuple in S^m and a function $F\colon (\mathbb{N}_{>m}) \times S^m \to S$, that is,

$$\exists \big(s_1, \ldots, s_m\big) \in S^m \quad \text{and} \quad F\colon (\mathbb{N}_{>m}) \times S^m \to S,$$

such that

(I) $\forall k \in \{1, \ldots, m\} \quad \varphi(k) = s_k$,
(II) $\forall n \in \mathbb{N}_{>m} \quad \varphi(n) = F\big(n, \underbrace{\varphi(n-m), \ldots, \varphi(n-1)}_{m\text{-tuple in } S^m}\big)$.

THEOREM 5.1.11 (Recursion Theorem). *Let S be a nonempty set, $m \in \mathbb{N}$, and let*

$$\big(s_1, \ldots, s_m\big) \in S^m \quad \text{and} \quad F\colon (\mathbb{N}_{>m}) \times S^m \to S,$$

There exists a unique sequence $\varphi : \mathbb{N} \to S$ such that

(I) $\forall k \in \{1, \ldots, m\} \quad \varphi(k) = s_k$,
(II) $\forall n \in \mathbb{N}_{>m} \quad \varphi(n) = F\big(n, \underbrace{\varphi(n-m), \ldots, \varphi(n-1)}_{m\text{-tuple in } S^m}\big)$.

PROOF. The proof is technical, so we skip it. □

EXAMPLE 5.1.12. In this example we give several recursively defined sequences.

(i) $\varphi(1) = 1 \quad \text{and} \quad \forall n \in \mathbb{N}\setminus\{1\} \quad \varphi(n) = -\dfrac{\varphi(n-1)}{2}$,

(ii) $\varphi(1) = 1 \quad \text{and} \quad \forall n \in \mathbb{N}\setminus\{1\} \quad \varphi(n) = 1 + \dfrac{\varphi(n-1)}{4}$,

(iii) $\varphi(1) = 2 \quad \text{and} \quad \forall n \in \mathbb{N}\setminus\{1\} \quad \varphi(n) = \dfrac{\varphi(n-1)}{2} + \dfrac{1}{\varphi(n-1)}$,

(iv) $\varphi(1) = 0 \quad \text{and} \quad \forall n \in \mathbb{N}\setminus\{1\} \quad \varphi(n) = \dfrac{9 + \varphi(n-1)}{1+9}$,

(v) $\varphi(0) = 0,\ \varphi(1) = 1 \quad \text{and} \quad \forall n \in \mathbb{N}\setminus\{1\} \quad \varphi(n) = \varphi(n-1) + \varphi(n-2)$. This sequence is a recursively defined sequence of order 2. The domain of this sequence is $\mathbb{N} \cup \{0\}$.

For a recursively defined sequence it is useful to evaluate the values of the first few terms to get an idea how sequence behaves. For example, the first eighteen terms of the sequence in (v) are

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ 233,\ 377,\ 610,\ 987,\ 1597, \ldots$$

◁

5.2. Examples and exercises related to $\mathbb{N}$

The following two examples deal with two familiar functions: the factorial and the power function. Let $n \in \mathbb{N}$. The factorial is informally "defined" as

$$n! = 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n.$$

Let $a \in \mathbb{R}$. The n-th power of a is informally expressed as

$$a^n = \underbrace{a \cdot a \cdot \ldots \cdot a \cdot a}_{n \text{ times}}.$$

Next we give the rigorous definitions of the factorial and the power function as examples of recursive definitions.

Example 5.2.1. The sequence $f : \mathbb{N} \to \mathbb{N}$ defined by

(i) $f(1) = 1$,
(ii) $\forall\, n \in \mathbb{N} \quad f(n+1) = (n+1)\, f(n)$,

is called the *factorial*.

The standard notation for the factorial is $f(n) = n!$. The definition of factorial is extended to 0 by setting $0! = 1$. ◁

Example 5.2.2. Let $a \in \mathbb{R}$. Define the sequence $g : \mathbb{N} \to \mathbb{R}$ by

(i) $g(1) = a$,
(ii) $\forall\, n \in \mathbb{N} \quad g(n+1) = a\, g(n)$.

The standard notation for the function g is $g(n) = a^n$. The expression a^n is called the n-th *power* of a. For $a \neq 0$, the definition of the power is extended to 0 by setting $a^0 = 1$. The expression 0^0 is not defined. ◁

Exercise 5.2.3. Let $a, b \in \mathbb{R}$ be such that $a, b \geq 0$. Let $n \in \mathbb{N}$. Prove that $a < b$ if and only if $a^n < b^n$. ◁

Solution. □

Use the Principle of Mathematical Induction to do the following exercises.

Exercise 5.2.4. Consider the sequence $f : \mathbb{N} \to \mathbb{N}$ defined by

(i) $f(1) = 1$,
(ii) $\forall\, n \in \mathbb{N} \quad f(n+1) = f(n) + (2n+1)$.

Evaluate the values $f(2), f(3), f(4), f(5)$. Based on the numbers that you get, guess a simple formula for $f(n)$ and prove it. ◁

Exercise 5.2.5. Consider the sequence $T : \mathbb{N} \to \mathbb{N}$ defined by

(i) $T(1) = 1$,
(ii) $\forall\, n \in \mathbb{N} \quad T(n+1) = T(n) + (n+1)$.

Evaluate the values $T(2), T(3), T(4), T(5), T(6)$. Based on these numbers guess a simple formula for $T(n)$ in terms of n and prove it. ◁

Remark 5.2.6. The numbers $T(n)$, with $n \in \mathbb{N}$, are called the *triangular numbers*. For $n \in \mathbb{N}$, the triangular number

$$T(n) = 1 + 2 + \cdots + (n-1) + n$$

is the additive analog of the factorial (see Example 5.2.1)

$$n! = 1 \cdot 2 \cdot \cdots \cdot (n-1) \cdot n.$$

For completeness we set $T(0) = 0$. ◁

EXERCISE 5.2.7. Let $a \in \mathbb{R}$ and $x \in \mathbb{R} \setminus \{0\}$. Consider the sequence $g : \mathbb{N} \to \mathbb{R}$ defined by

(i) $g(1) = a$,
(ii) $\forall n \in \mathbb{N} \quad g(n+1) = g(n) + a\,x^n$.

Another way of writing $g(n)$ is

$$g(n) = \sum_{k=0}^{n-1} a\,x^k.$$

Informally this sum is sometimes written as

$$g(n) = a + a\,x + \cdots + a\,x^{n-1}.$$

This sum is called the *geometric sum*.

Prove that

$$g(n) = \begin{cases} a\,\dfrac{1-x^n}{1-x} & \text{if } \ x \neq 1, \\ n\,a & \text{if } \ x = 1. \end{cases}$$ ◁

EXERCISE 5.2.8. (Bernoulli's inequality) For all $n \in \mathbb{N}$ and all $x \in \mathbb{R}_{>-1}$ we have

$$(1+x)^n \geq 1 + nx.$$ ◁

SOLUTION. Let $x \in \mathbb{R}_{>-1}$ and let $n \in \mathbb{N}$ be arbitrary. The statement that we are proving is

$$P(n): \qquad (1+x)^n \geq 1 + nx.$$

We want to prove

$$\forall n \in \mathbb{N} \quad P(n).$$

We proceed by Mathematical Induction. Verify

$$P(1): \qquad (1+x)^1 \geq 1 + 1x.$$

Since

$$(1+x)^1 = 1 + x = 1 + 1x,$$

the statement $P(1)$ is true. □

EXERCISE 5.2.9. For all $n \in \mathbb{N}$ and all $x \in [0,1]$ we have

$$(1+x)^n \leq 1 + \big(2^n - 1\big)x.$$ ◁

EXERCISE 5.2.10. For all $n \in \mathbb{N}$ and all $x \in (0,1)$ we have

$$(1-x)^n < 1 - nx + \frac{1}{2}(nx)^2.$$ ◁

EXERCISE 5.2.11. For all $n \in \mathbb{N}$ and all $a \in [-1,1]$ we have

$$|a|^n \leq \frac{|a|}{n(1-|a|)+|a|}.$$ ◁

EXERCISE 5.2.12. (Binomial theorem) For all $n \in \mathbb{N}$ and all $x, y \in \mathbb{R}$ we have

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} \, y^k.$$

Here $\binom{n}{k}$ denotes the *binomial coefficient* which is defined by

$$\binom{n}{k} \stackrel{\text{def}}{=} \frac{n!}{k! \, (n-k)!} \quad \text{where} \quad n \in \mathbb{N}, \quad k \in \{0, 1, \ldots, n\}.$$

The most important property of binomial coefficients is given by the following equality: For all $n \in \mathbb{N}$ and all $k \in \{0, 1, \ldots, n\}$ we have

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$$

This formula is proved by using the definition of the binomial coefficients and the rules for adding fractions. ◁

EXERCISE 5.2.13. (a) Prove that for all $n \in \mathbb{N}$ we have

$$\sum_{k=1}^{n} \frac{1}{k^2} \leq 2 - \frac{1}{n}$$

(b) Prove that the set

$$A = \left\{ \sum_{k=1}^{n} \frac{1}{k^2} \, : \, n \in \mathbb{N} \right\}$$

is bounded.

(c) Does A have a minimum? Does A have a maximum? ◁

5.3. Finite sets and infinite sets

One of the most important applications of the natural numbers is counting. The following special subsets of $\mathbb{N}$ are used for counting:

$$[\![1, n]\!] \stackrel{\text{def}}{=} \big\{ k \in \mathbb{N} \, : \, k \leq n \big\}, \quad n \in \mathbb{N}.$$

For $n \in \mathbb{N}$, this set is called the *n-element counting set.*

For example

$$\begin{aligned}
[\![1, 1]\!] &= \{1\} \\
[\![1, 2]\!] &= \{1, 2\} \\
[\![1, 3]\!] &= \{1, 2, 3\} \\
[\![1, 4]\!] &= \{1, 2, 3, 4\} \\
[\![1, 5]\!] &= \{1, 2, 3, 4, 5\} \\
[\![1, 6]\!] &= \{1, 2, 3, 4, 5, 6\} \\
&\ \ \vdots
\end{aligned}$$

Next we give a formal mathematical definition of the counting process.

DEFINITION 5.3.1. A nonempty set A is *finite* if there exists $n \in \mathbb{N}$ and a bijection $\phi : [\![1, n]\!] \to A$ with domain $[\![1, n]\!]$ and range A. In this case we say that A *has* n *elements* and write $\operatorname{card}(A) = n$.

EXERCISE 5.3.2. If A is finite and $b \notin A$, then $A \cup \{b\}$ is finite and $\operatorname{card}\big(A \cup \{b\}\big) = 1 + \operatorname{card}(A)$. ◁

EXERCISE 5.3.3. A nonempty set A is finite if and only if there exists $n \in \mathbb{N}$ and a surjection $f : [\![1, n]\!] \to A$. In this case $\operatorname{card}(A) \leq n$. HINT: The "only if" direction is straightforward. The other direction is a statement involving a natural number n: If $f : [\![1, n]\!] \to A$ is a surjection, then A is finite and $\operatorname{card}(A) \leq n$. This can be proved by mathematical induction. The statement is true for $n = 1$. Assume that it is true for n and $f : [\![1, n+1]\!] \to A$ is a surjection. Set $B = f\big([\![1, n]\!]\big) \subseteq A$. Then by the inductive hypothesis B is finite and $\operatorname{card}(B) \leq n$. There are two cases $B = A$ and $B \subsetneq A$. If $B = A$, then, by the inductive hypothesis, A finite and $\operatorname{card}(A) \leq n \leq n + 1$. If $B \subsetneq A$, then $A = B \cup \{f(n+1)\}$. By Exercise 5.3.2 A is finite and $\operatorname{card}(A) = 1 + \operatorname{card}(B) \leq n + 1$. ◁

Finite sets are often informally written as $A = \{a_1, a_2, \ldots, a_n\}$. However, this way of writing does not imply that the mapping $k \mapsto a_k$, $k \in [\![1, n]\!]$, is a bijection, but it does imply that this mapping is a surjection.

EXERCISE 5.3.4. A nonempty subset B of a finite set A is finite and we have $\operatorname{card}(B) \leq \operatorname{card}(A)$. HINT: Let $f : [\![1, n]\!] \to A$ be a bijection. Since B is not empty there exists $b \in B$. Define the function $g : A \to B$ by $g(x) = \begin{cases} x & \text{if } x \in B \\ b & \text{if } x \in A \setminus B \end{cases}$. Clearly the range of g is B. Therefore $g \circ f : [\![1, n]\!] \to B$ is a surjection. By Exercise 5.3.3, B is finite. ◁

EXERCISE 5.3.5. Let B is a nonempty proper subset of a finite set A. Then B is finite and $\operatorname{card}(B) < \operatorname{card}(A)$. HINT: Let $c \in A \setminus B$. Then

$$\operatorname{card}(B) \leq \operatorname{card}\big(A \setminus \{c\}\big) = \operatorname{card}(A) - 1 < \operatorname{card}(A). \qquad ◁$$

EXERCISE 5.3.6. If A is a finite subset of $\mathbb{R}$, then A has a minimum and a maximum. ◁

DEFINITION 5.3.7. A nonempty set which is not finite is said to be *infinite.*

REMARK 5.3.8. The previous definition of an infinite set is certainly logically correct, but it is not 'constructive'.

Therefore it is desirable to give a formal negation of the definition of finite set. Before doing that I will restate the definition of a finite set as:

"A set A is finite if (and only if) there exists $n \in \mathbb{N}$ and there exists a bijection $f : [\![1, n]\!] \to A$."

The negation of the last statement (and thus a characterization of an infinite set) is the following

"For every $n \in \mathbb{N}$ and for every $f : [\![1, n]\!] \to A$ we have that f is not a bijection." ◁

Remark 5.3.9. The importance of Exercise 5.3.6 is twofold. First, it states the most important property of finite sets of real numbers. Second, its contrapositive provides a simple way of proving that a set is infinite: If a nonempty subset of $\mathbb{R}$ does not have a minimum or it does not have a maximum, then it is infinite.

The fact that infinite sets might not have a minimum and/or maximum makes dealing with such sets more difficult. The following proposition states that not having a minimum and/or maximum is to some extend a universal property of an infinite subset of $\mathbb{R}$. $\lhd$

Theorem 5.3.10. *Let $S \subseteq \mathbb{R}$. If S is infinite, then there exists a nonempty subset A of S such that A does not have a minimum or there exists a nonempty subset B of S such that B does not have a maximum.*

Proof. We will prove the equivalent implication: If S is an infinite subset of $\mathbb{R}$ and each nonempty subset of S has a minimum, then there exist a nonempty subset B of S such that B does not have a maximum.

So, assume that S is an infinite subset of $\mathbb{R}$ and each nonempty subset of S has a minimum. Then, in particular, $\min(S)$ exists. Let W be the set of all minimums of infinite subsets of S. Formally,

$$W = \Bigl\{ x \in S \; : \; x = \min(E) \;\; \text{where} \;\; E \subseteq S \;\; \text{and} \;\; E \text{ is infinite} \Bigr\}.$$

Clearly $\min(S)$ is an element in W. Hence $W \neq \emptyset$.

Next we will prove that W does not have a maximum. Let $y \in W$ be arbitrary. Then there exists an infinite subset F of S such that $y = \min(F)$. Since F is infinite, the set $F \setminus \{y\}$ is also infinite. Since $F \setminus \{y\} \subset S$, by the assumption $z = \min\bigl(F \setminus \{y\}\bigr)$ exists. Therefore, $z \in W$. Since $z \in F \setminus \{y\}$, we have $z \neq y$. Since $z \in F$ and $y = \min(F)$, we have $z \geq y$. Hence $z > y$. Thus, for each $y \in W$ there exists $z \in W$ such that $z > y$. This proves that W is a nonempty subset of S which does not have a maximum. $\square$

Remark 5.3.11. The above theorem is an implication of the form:

$$p \;\Rightarrow\; q \vee r.$$

The displayed implication is equivalent to the implication

$$p \wedge (\neg q) \;\Rightarrow\; r.$$

One way to see this is to consider the negations of both displayed implications. The negation of $p \Rightarrow q \vee r$ is $p \wedge (\neg q) \wedge (\neg r)$. The negation of $p \wedge (\neg q) \Rightarrow r$ is is $p \wedge (\neg q) \wedge (\neg r)$. Since the negations are clearly equivalent, the implications are also equivalent. $\lhd$

Exercise 5.3.12. Prove that $\mathbb{N}$ does not have a maximum. $\lhd$

Exercise 5.3.13. Prove that a nonempty subset of $\mathbb{N}$ is finite if and only if it has a maximum. $\lhd$

Exercise 5.3.14. Prove that the set $\mathbb{N}$ is infinite. $\lhd$

THEOREM 5.3.15 (Well Ordering Property). *If A is a nonempty subset of $\mathbb{N}$, then A has a minimum.*

REMARK 5.3.16. Subsets of $\mathbb{N}$ can be infinite. As I mentioned in Remark 5.3.9, a problem with infinite sets is a possible absence of minimum and maximum. Theorem 5.3.15 tells us that a subset of natural numbers must at least have a minimum. Consequently, infinite subsets of $\mathbb{N}$ are not as bad as infinite subsets of $\mathbb{R}$. ◁

PROOF OF THEOREM 5.3.15. This proof uses the following two facts:

(i) Each finite set has a minimum. (Proved in Exercise 5.3.6.)
(ii) For each $n \in \mathbb{N}$ each subset of the set $\{1, 2, \ldots, n\} = [\![1, n]\!]$ is finite. (Proved in Exercise 5.3.4.)

Since $A \neq \emptyset$, there exists $n \in A$. Consider the set $B = \{x \in A : x \leq n\}$. Then $B \subseteq [\![1, n]\!]$. By fact (ii) B is finite. Now, by fact (i) B has a minimum; denote it by $m = \min B$. Then m is also the minimum of A. (Here is a proof: If $a \in A$, then either $a \leq n$, or $n < a$. In the first case $a \in B$, and therefore $m \leq a$. If $n < a$, then $m \leq n < a$, and therefore $m \leq a$ for each $a \in A$.) □

5.4. Countable sets

Our intuition, gained from everyday life, has been helpful for understanding the mathematical aspects of finite sets. In fact, many statements proved in the preceding section were intuitively clear and almost did not require proofs. We proved them to emphasize our focus on rigor.

However, as we proceed with our studies beyond finite sets, our intuition from everyday life becomes less reliable, and we need to build a mathematical intuition for infinite sets.

One of the goals of this section is to understand that the set of natural numbers, $\mathbb{N}$, is the smallest infinite set.

DEFINITION 5.4.1. A nonempty set A is *countable* if there exists a bijection $\phi : \mathbb{N} \to A$.

The following three theorems provide powerful tools for proving that a particular set is countable, especially when constructing a specific bijection between the set and $\mathbb{N}$ is challenging.

THEOREM 5.4.2. *If S is an infinite subset of $\mathbb{N}$, then S is countable. That is, there exists a bijection with domain $\mathbb{N}$ and range S.*

PROOF. Let S be an infinite subset of $\mathbb{N}$. Let $s \in S$ be arbitrary. Then the set $S \cap [\![1, s]\!]$ is finite, since it is a nonempty subset of the finite set $[\![1, s]\!]$. Define the function $f : S \to \mathbb{N}$ by setting

$$\forall s \in S \qquad \phi(s) = \operatorname{card}\big(S \cap [\![1, s]\!]\big).$$

The function $\phi : S \to \mathbb{N}$ has the following three properties:

(I) If $s, t \in S$ and $s < t$, then $\phi(s) < \phi(t)$.

(II) If $s = \min S$, then $\phi(s) = 1$.
(III) If $s \in S$ and $t = \min\bigl(S \setminus [\![1, s]\!]\bigr)$, then $\phi(t) = \phi(s) + 1$.

Property (I) follows from Exercise 5.3.5. Property (II) follows from the fact that, $s = \min(S)$ implies $S \cap [\![1, s]\!] = \{s\}$. Property (III) follows from Exercise 5.3.2.

Property (I) implies that ϕ is an injection. Properties (II) and (III) imply that the range of ϕ, call it K, has the following properties: $1 \in K$ and $n \in K \Rightarrow n+1 \in K$. Since $K \subseteq \mathbb{N}$, this, by Theorem 5.1.2 (N 5), implies $K = \mathbb{N}$. Thus ϕ is a surjection. Hence, $\phi : S \to \mathbb{N}$ is a bijection. Its inverse $\phi^{-1} : \mathbb{N} \to S$ is the desired bijection. □

COROLLARY 5.4.3. *An infinite subset of a countable set is countable.*

REMARK 5.4.4. (THE SPECIAL ROLE OF COUNTABLE SETS). Let A be an infinite set. Recall that by Definition 3.3.4, we have $\operatorname{card}(A) \leq \operatorname{card}(\mathbb{N})$ if there exists a subset $S \subseteq \mathbb{N}$ and a bijection $g : A \to S$. Since A is infinite and $g : A \to S$ is a bijection, S is also infinite. By Theorem 5.4.2, S is countable. Given that $g : A \to S$ is a bijection, A is countable as well. Hence, the cardinality of a countable set represents the smallest infinite cardinality. ◁

THEOREM 5.4.5. *If A is an infinite set and $f : A \to \mathbb{N}$ is an injection, then A is countable. That is, there exists a bijection with domain $\mathbb{N}$ and range A.*

PROOF. Assume that A is an infinite set and $f : A \to \mathbb{N}$ is an injection. Denote by $S \subseteq \mathbb{N}$ the range of $f : A \to \mathbb{N}$. Since f is injective, the function $f : A \to S$ is a bijection. Since A is infinite and there is a bijection between A and S, the set S is infinite as well.

By Theorem 5.4.2, there exists a bijection $\phi : S \to \mathbb{N}$. Now, consider the composition $\psi = \phi \circ f$. Since $f : A \to S$ is a bijection and $\phi : S \to \mathbb{N}$ is a bijection, their composition $\psi : A \to \mathbb{N}$ is also a bijection, by Theorem 3.2.18. Therefore, $\psi^{-1} : \mathbb{N} \to A$ is the desired bijection with domain $\mathbb{N}$ and range A, proving that A is countable. □

THEOREM 5.4.6. *If A is an infinite set and $g : \mathbb{N} \to A$ is a surjection, then A is countable. That is, there exists a bijection with domain $\mathbb{N}$ and range A.*

PROOF. Assume that A is an infinite set and $g : \mathbb{N} \to A$ is a surjection. Since g is a surjection, for every $a \in A$, the set $\{k \in \mathbb{N} : g(k) = a\}$ is nonempty. By Theorem 5.3.15, this set has a minimum. Define the function $h : A \to \mathbb{N}$ by

$$\forall a \in A \quad h(a) = \min\{k \in \mathbb{N} : g(k) = a\}.$$

Let us show that h is injective. Suppose, for some $a, b \in A$,

$$h(a) = h(b) = l \in \mathbb{N}$$

By definition of $h(a)$ we have $g(l) = a$ and by definition of $h(b)$ we have $g(l) = b$. Since $g : \mathbb{N} \to A$ is a function (unival), we have $a = b$, thus $h : A \to \mathbb{N}$ is an

$$\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & \dots \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & \dots \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & \dots \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & \dots \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

TABLE 5.4.1. $\mathbb{N}\times\mathbb{N}$

injection. By Theorem 5.4.5, A is countable, that is a bijection with domain $\mathbb{N}$ and range A. □

It will be proved in Section 5.5 that the set of integers and the set of rational numbers are countable sets.

The following exercise deals with the Cartesian square of the set $\mathbb{N}$; that is the set $\mathbb{N} \times \mathbb{N}$. Recall that this is the set of all ordered pairs of positive integers:

$$\mathbb{N} \times \mathbb{N} \stackrel{\text{def}}{=} \big\{(s,t) \,:\, s,t \in \mathbb{N}\big\}.$$

A part of the set $\mathbb{N}\times\mathbb{N}$ is presented in Table 5.4.1.

THEOREM 5.4.7. *The Cartesian product* $\mathbb{N}\times\mathbb{N}$ *is countable.*

PROOF. The function $A : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by

$$\forall (s,t) \in \mathbb{N}\times\mathbb{N} \quad A(s,t) = \frac{(s+t-1)(s+t)}{2} - t + 1,$$

is a bijection. This bijection is inspired by Table 5.4.3. Subsection 5.4.1 is devoted to the proof that it is a bijection. □

EXERCISE 5.4.8. Let $\mathcal{A}$ be countable family of sets. Assume that each set in $\mathcal{A}$ is countable. Prove that $\bigcup\big\{A \,:\, A \in \mathcal{A}\big\}$ is also countable. ◁

Recall from Definition 3.2.4 that $\{0,1\}^{\mathbb{N}}$ denotes the set of all functions with domain $\mathbb{N}$ and codomain $\{0,1\}$; in other words, it is the set of all sequences taking values in $\{0,1\}$.

THEOREM 5.4.9. *The set* $\{0,1\}^{\mathbb{N}}$ *is uncountable.*

PROOF. To prove that $\{0,1\}^{\mathbb{N}}$ is uncountable, we will prove that an arbitrary function $\Phi : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ is <u>not</u> a surjection. This will be proved by constructing a specific function $f \in \{0,1\}^{\mathbb{N}}$, that is a sequence $f : \mathbb{N} \to \{0,1\}$, such that $f \neq \Phi_n$ for all $n \in \mathbb{N}$. To construct f let us analyze Φ_n, $n \in \mathbb{N}$. Clearly $\Phi_1 : \mathbb{N} \to \{0,1\}$, that is $\Phi_1(n) \in \{0,1\}$ for all $n \in \mathbb{N}$. We can indicate the action of Φ_1 on $\mathbb{N}$ by listing its first 7 values:

$$\Phi_1(1) \qquad \Phi_1(2) \qquad \Phi_1(3) \qquad \Phi_1(4) \qquad \Phi_1(5) \qquad \Phi_1(6) \qquad \Phi_1(7) \qquad \dots$$

$$\begin{array}{cccccccc}
\boxed{\Phi_1(1)} & \Phi_1(2) & \Phi_1(3) & \Phi_1(4) & \Phi_1(5) & \Phi_1(6) & \Phi_1(7) & \cdots \\
\Phi_2(1) & \boxed{\Phi_2(2)} & \Phi_2(3) & \Phi_2(4) & \Phi_2(5) & \Phi_2(6) & \Phi_2(7) & \cdots \\
\Phi_3(1) & \Phi_3(2) & \boxed{\Phi_3(3)} & \Phi_3(4) & \Phi_3(5) & \Phi_3(6) & \Phi_3(7) & \cdots \\
\Phi_4(1) & \Phi_4(2) & \Phi_4(3) & \boxed{\Phi_4(4)} & \Phi_4(5) & \Phi_4(6) & \Phi_4(7) & \cdots \\
\Phi_5(1) & \Phi_5(2) & \Phi_5(3) & \Phi_5(4) & \boxed{\Phi_5(5)} & \Phi_5(6) & \Phi_5(7) & \cdots \\
\Phi_6(1) & \Phi_6(2) & \Phi_6(3) & \Phi_6(4) & \Phi_6(5) & \boxed{\Phi_6(6)} & \Phi_6(7) & \cdots \\
\Phi_7(1) & \Phi_7(2) & \Phi_7(3) & \Phi_7(4) & \Phi_7(5) & \Phi_7(6) & \boxed{\Phi_7(7)} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

TABLE 5.4.2. Cantor's diagonal argument

We can do the same for $\Phi_2, \Phi_3, \Phi_4, \Phi_5, \Phi_6, \ldots$ to get the infinite table in Table 5.4.2. The reason why the diagonal elements $\Phi_n(n), n \in \mathbb{N}$, in Table 5.4.2 are boxed will be clear shortly.

Our goal is to define $f : \mathbb{N} \to \{0,1\}$ which will differ from Φ_n for each $n \in \mathbb{N}$. The easiest way to achieve this is to define $f(n)$ to have a different value from $\Phi_n(n)$, the boxed value in Table 5.4.2. Thus we define:

$$\forall n \in \mathbb{N} \qquad f(n) = 1 - \Phi_n(n).$$

For all $n \in \mathbb{N}$ we have $\Phi_n(n) \in \{0,1\}$. Therefore, if $\Phi_n(n) = 0$, we have $f(n) = 1$, and, if $\Phi_n(n) = 1$, we have $f(n) = 0$. In each cace $f(n) \in \{0,1\}$ and $f(n) \neq \Phi_n(n)$.

In particular, $f(1) \neq \Phi_1(1)$. Therefore, $f \neq \Phi_1$. Also, $f(2) \neq \Phi_2(2)$; so $f \neq \Phi_2$. In general, for all $n \in \mathbb{N}$ we have $f(n) \neq \Phi_n(n)$. Therefore,

$$\forall n \in \mathbb{N} \qquad f \neq \Phi_n.$$

Hence $\Phi : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ is not a surjection. Since $\Phi : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ was an arbitrary function, we conclude that there does not exist a bijection between $\mathbb{N}$ and $\{0,1\}^{\mathbb{N}}$. □

EXERCISE 5.4.10. Prove that the set of all prime numbers is countable. ◁

5.4.1. An explicit bijection between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$. We conclude this section by establishing an explicit formula for a bijection between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$. A natural way to proceed to find this bijection is to rearrange Table 5.4.1 in a triangular shape and label the ordered pairs in $\mathbb{N} \times \mathbb{N}$ by elements of $\mathbb{N}$ placed in a circle next to the corresponding ordered pair. As a result we get Table 5.4.3.

First, notice that in each column in Table 5.4.3 the sum of the elements of the ordered pairs is constant. Second, notice that the label at the bottom of each column is a **triangular number**, see Exercise 5.2.5. The label for the pair of the form $(s, 1)$ is the triangular number $(s)(s+1)/2$. The labels for the pairs in the same column are obtained by subtracting the value of $t - 1$. Thus the bijection $A : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is given by

$$A(s,t) = \frac{(s+t-1)(s+t)}{2} - t + 1, \qquad s, t \in \mathbb{N}.$$

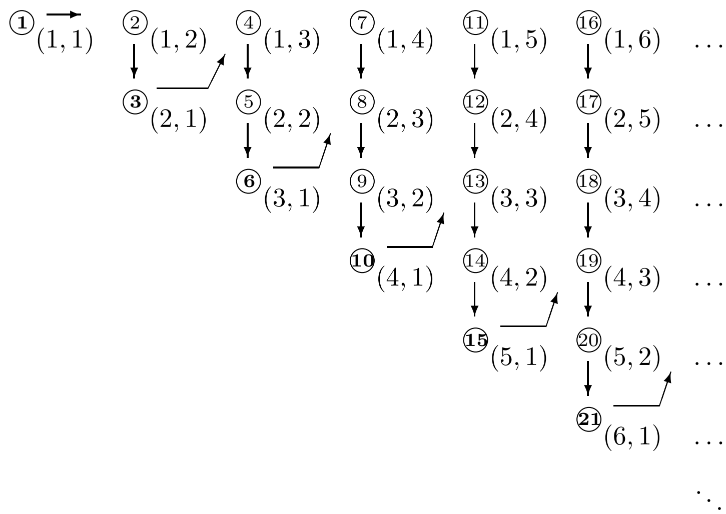


TABLE 5.4.3. Labeled $\mathbb{N} \times \mathbb{N}$

The easiest way to show that A is a bijection is to find its inverse $B : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$. We want to figure out which pair is associated with the label $n \in \mathbb{N}$. Looking at Table 5.4.3 we see that for the label, say 19, the first component of the corresponding pair $(4,3)$ is the distance of 19 to the previous triangular number 15 and the second component is the distance of 19 to the next triangular number 21 plus 1. So, to construct B we need to need to figure out a way of capturing each positive integer n between consecutive triangular numbers. That is, for each $n \in \mathbb{N}$ we need to find m such that

$$\frac{(m-1)m}{2} + 1 \leq n \leq \frac{m(m+1)}{2}, \tag{5.4.1}$$

or, equivalently, in the notation of Exercise 5.2.5,

$$T_{m-1} < n \leq T_m. \tag{5.4.2}$$

Such m is uniquely determined by n and we set $m = R_n$. It follows from (5.4.2) that

$$R_n = \min\{k \in \mathbb{N} : n \leq T_k\}, \qquad n \in \mathbb{N}.$$

Clearly $n(n+1) \geq 2n$ for all $n \in \mathbb{N}$. Therefore, $T_n \geq n$ for all $n \in \mathbb{N}$. Consequently the set in the definition of R_n is not empty. By Theorem 5.3.15 every nonempty subset of $\mathbb{N}$ has a minimum, so R_n is well defined.

The following table illustrates the relationship between the sequences $\{T_n\}$ and $\{R_n\}$. The **triangular numbers** are in bold face.

	T_1		T_2			T_3				T_4					T_5					T_6	
n	**1**	2	**3**	4	5	**6**	7	8	9	**10**	11	12	13	14	**15**	16	17	18	19	20	**21**	22
R_n	1	2	2	3	3	3	4	4	4	4	5	5	5	5	5	6	6	6	6	6	6	7

REMARK 5.4.11. There are several other formulas for the sequence R. For example for $n \in \mathbb{N}$,

$$R_n = \left\lfloor \frac{1}{2} + \sqrt{2n} \right\rfloor \quad \text{or} \quad R_n = \left\lceil -\frac{1}{2} + \sqrt{2n} \right\rceil.$$

Here $\lfloor \cdot \rfloor$ is the floor function, $\lceil \cdot \rceil$ is the ceiling function and $\sqrt{\cdot}$ is the square root function. These functions will be introduced in Section 6.2.

An alternative way to define the sequence R is the following recursive definition.

(i) $R_1 = 1$,
(ii) $\forall n \in \mathbb{N} \quad R_{n+1} = 1 + R_{(n+1-R_n)}$. ◁

With the sequence $\{R_n\}$ at our disposal we can define $B : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$:

$$B(n) = \Biggl(\underbrace{n - \frac{(R_n - 1)R_n}{2}}_{\substack{\text{distance to the preced-}\\ \text{ing triangular number}}}, \underbrace{\frac{R_n(R_n+1)}{2} - n}_{\substack{\text{distance to the follow-}\\ \text{ing triangular number}}} + 1 \Biggr), \quad n \in \mathbb{N}.$$

EXERCISE 5.4.12. Prove that

$$B\bigl(A(s,t)\bigr) = (s,t) \quad \text{and} \quad A\bigl(B(n)\bigr) = n, \quad \text{for all} \quad s,t,n \in \mathbb{N}.$$

HINT: Notice that by Exercise 5.2.5 the formulas for A and B can be written as

$$A(s,t) = T_{(s+t-1)} - t + 1, \qquad s,t \in \mathbb{N},$$
$$B(n) = \Bigl(n - T_{R_n - 1}\, , \, -n + 1 + T_{R_n} \Bigr), \quad n \in \mathbb{N}.$$

Let $s,t \in \mathbb{N}$. To evaluate $B\bigl(A(s,t)\bigr)$ you will need to evaluate $R_{A(s,t)}$ first. The following straightforward inequalities

$$T_{(s+t-2)} = T_{(s+t-1)} - (s+t-1) < A(s,t) = T_{(s+t-1)} - t + 1 \leq T_{(s+t-1)}$$

imply that $R_{A(s,t)} = s + t - 1$. With this identity calculating $B\bigl(A(s,t)\bigr)$ is just a careful application of definitions and algebra. Calculating $A(B(n))$ is also just an application of definitions and algebra. In this calculation use $T_k - T_{k-1} = k$. This is the end of HINT. ◁

5.5. The sets ℤ and ℚ

DEFINITION 5.5.1. A real number x is said to be an *integer* if one of the following three conditions is satisfied $x \in \mathbb{N}$ or $x = 0$ or $-x \in \mathbb{N}$. The set of all integers is denoted by $\mathbb{Z}$. That is,

$$\mathbb{Z} = \bigl\{ x \in \mathbb{R} : \ x \in \mathbb{N} \ \ \vee \ \ x = 0 \ \ \vee \ \ -x \in \mathbb{N} \bigr\}.$$

EXERCISE 5.5.2. Prove that $\mathbb{Z}$ is countable. ◁

SOLUTION. One bijection is the function $f : \mathbb{Z} \to \mathbb{N}$ defined by

$$\forall z \in \mathbb{Z} \quad f(z) = 2|z| + \operatorname{us}(z) = 2\left| z + \frac{1}{4} \right| + \frac{1}{2}.$$

Here, one needs to prove that $f(z) \in \mathbb{N}$ for all $z \in \mathbb{Z}$. The inverse of f is the function $g : \mathbb{N} \to \mathbb{Z}$ defined by

$$\forall n \in \mathbb{N} \quad g(n) = \frac{(-1)^{n+1}}{4}(2n-1) - \frac{1}{4}.$$

Here, one needs to prove that $g(n) \in \mathbb{Z}$ for all $n \in \mathbb{N}$ and

$$\forall n \in \mathbb{N} \quad f\bigl(g(n)\bigr) = n \qquad \text{and} \qquad \forall z \in \mathbb{Z} \quad g\bigl(f(z)\bigr) = z. \qquad \square$$

EXERCISE 5.5.3. Prove that $\mathbb{Z} \times \mathbb{N}$ is countable. ◁

DEFINITION 5.5.4. A real number x is said to be a *rational number* if there exist $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ such that $x = m \cdot \frac{1}{n}$. That is $x = m/n = \frac{m}{n}$. The set of all rational numbers we denote by $\mathbb{Q}$, that is

$$\mathbb{Q} = \left\{ x \in \mathbb{R} \ : \ \exists m \in \mathbb{Z} \ \wedge \ \exists n \in \mathbb{N} \ \text{ such that } \ x = \frac{m}{n} \right\}.$$

REMARK 5.5.5. (MULTIPLE REPRESENTATIONS OF RATIONAL NUMBERS) By Theorem 4.1.4(xiii), we have

$$\forall m \in \mathbb{Z} \ \ \forall k, n \in \mathbb{N} \quad \frac{km}{kn} = \frac{m}{n}.$$

That is, each rational number admits multiple representations as a fraction of two integers. In particular, the mapping with the domain $\mathbb{Z}\times\mathbb{N}$ and codomain $\mathbb{Q}$ defined by

$$\forall (m,n) \in \mathbb{Z}\times\mathbb{N} \quad (m,n) \mapsto \frac{m}{n},$$

is a surjection, but not an injection. ◁

EXERCISE 5.5.6. Denote by $\mathbb{Q}_+$ the set of all positive rational numbers. Prove that $\mathbb{Q}_+$ is countable. ◁

EXERCISE 5.5.7. Prove that there exists a bijection between $\mathbb{Q}_+$ and $\mathbb{Q}$. ◁

EXERCISE 5.5.8. Prove that the set $\mathbb{Q}$ is countable. ◁

THEOREM 5.5.9. *For all $r \in \mathbb{Q}$ we have $r^2 \neq 2$.*

PROOF. In this proof we will use the concepts of even and odd integers. We will use the fact that the square of an even integer is even and the square of an odd integer is odd. We will also use the fact that the sets of even and odd integers are disjoint.

We will first prove that for each $r \in \mathbb{Q}$ there exist $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ which are not both even such that $r = p/q$. Let $r \in \mathbb{Q}$ be arbitrary. Set

$$S = \left\{ n \in \mathbb{N} : \exists\ m \in \mathbb{Z} \ \text{ such that } \ r = \frac{m}{n} \right\}.$$

Since $r \in \mathbb{Q}$ the set S is not empty. Since S is a nonempty subset of $\mathbb{N}$, by the Well Ordering Property (Theorem 5.3.15), S has a minimum. Set $q = \min S$. Since $q \in S$, there exists $p \in \mathbb{Z}$ such that $r = p/q$. Next we will prove that p and q are not both even. That is we will prove the following implication: If $q = \min S$, $p \in \mathbb{Z}$, and $r = p/q$, then p and q are not both even.

It is easier to prove a partial contrapositive of the last implication. If $n \in S$, $m \in \mathbb{Z}$, $r = m/n$ and both m and n are even, then n is not a minimum of S. So, assume $n \in S$, $m \in \mathbb{Z}$ are both even and $r = m/n$. Then there exist $k \in \mathbb{Z}$ and $j \in \mathbb{N}$ such that $m = 2k$ and $n = 2j$. Clearly $j < n$. Also,

$$r = \frac{m}{n} = \frac{2\,k}{2\,j} = \frac{k}{j}.$$

Hence $j \in S$ and therefore n is not a minimum of S.

Let $r = p/q$ where p and q are not both even. Consider two cases: Case 1: p is odd, and Case 2: p is even.
Case 1. Assume that p is odd. By the background knowledge, p^2 is odd. Since $2q^2$ is even and p^2 is odd, we have $p^2 \neq 2q^2$.
Case 2. Assume that p is even. Then there exists an integer k such that $p = 2k$. Since not both q and p are even, q must be odd. Then q^2 is odd as well. Since $2k^2$ is even we have $2k^2 \neq q^2$. Consequently $4k^2 \neq 2q^2$. Since $4k^2 = (2k)^2 = p^2$, it follows that $p^2 \neq 2q^2$. Thus, $r^2 \neq 2$. □

In Theorem 5.5.9 we proved that there is no rational number r such that $r^2 = 2$. Since the set $\mathbb{Q}$ of rational numbers satisfies all Axioms 1 through 15, we cannot expect that based only on Axioms 1 through 15 we can prove that there exists a real number α such that $\alpha^2 = 2$. This is a proof that an additional axiom is needed for the set of real numbers; an axiom that will not be satisfied by the set of rational numbers. This is the **Completeness Axiom**.

There are additional properties of $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ that we cannot prove without the Completeness Axiom. For example:

- The set $\mathbb{N}$ is not bounded.
- For every $a, b \in \mathbb{R}$ such that $a < b$ there exists $r \in \mathbb{Q}$ such that $a < r < b$.
- The set $\mathbb{Q}$ is a proper subset of $\mathbb{R}$.
- There exists a positive real number α such that $\alpha^2 = 2$.
- The function $x \mapsto x^2$ with domain and codomain $\mathbb{R}_{\geq 0}$ is a bijection.
- The set $\mathbb{R}$ is not countable.

5.6. The Quadruplicity of Sets

In this subsection, we summarize the classification of sets established in this chapter. The fundamental dichotomy of sets divides all sets into two categories: the empty set and nonempty sets. This division leads to a trichotomy when we further classify nonempty sets into finite and infinite sets:

- The empty set.
- Finite sets.
- Infinite sets.

The classification extends by partitioning the infinite sets further into countable and uncountable sets. This results in the following quadruplicity:

- The empty set.
- Finite sets.
- Countable sets.
- Uncountable sets.

I would like to point out that this quadruplicity is not universally accepted. I value it for its clarity. However, when consulting other resources, it is crucial to verify the specific terminology used. For instance, many textbooks classify the empty set as finite, a classification I find confusing. The essence of a finite set lies in the existence of a bijection, and I am uncomfortable considering functions whose domain or codomain is the empty set. Additionally, many sources categorize finite sets as countable, distinguishing between finite countable sets and infinite countable sets. I find this distinction cumbersome. Therefore, I advocate for the quadruplicity outlined above due to its clarity.

CHAPTER 6

The Completeness Axiom

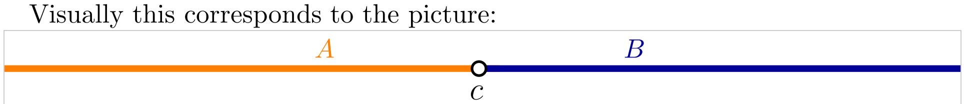
6.1. The Completeness Axiom

Recall Exercise 4.1.12: If a and b are real numbers such that $a < b$, then there exists a real number c such that $a < c < b$.

This statement assures that there are no big holes in $\mathbb{R}$; between any two real numbers there is another real number. A natural question to ask is whether the same is true for sets. If we are given two sets which are in some sense separated, does there exists a real number between them? Somewhat surprisingly this has to be postulated as the last axiom of the real numbers:

Axiom 16 (**CA**: Completeness Axiom). If A and B are nonempty subsets of $\mathbb{R}$ such that for all $a \in A$ and for all $b \in B$ we have $a \leq b$, then there exists $c \in \mathbb{R}$ such that for all $a \in A$ and for all $b \in B$ we have $a \leq c \leq b$.

Visually this corresponds to the picture:



However, the reader should realize that the pictured setting with A being the orange points and B being the blue points is very special. The set A could have been represented by a dust cloud of orange points, while the set B could have been represented by a dust cloud of blue points, the only requirement being that all the blue points are to the right of all the orange points.

Since we perceive the real number line to have no holes, the place marked by the open circle must correspond to a real number c.

Now we have 16 axioms of $\mathbb{R}$. It is remarkable that all statements about real numbers that are studied in beginning mathematical analysis courses can be deduced from these sixteen axioms and basic properties of sets.

The formulation of the Completeness Axiom given as Axiom 16 above is not standard. I adopted this version from the book *Mathematical Analysis* by Vladimir Zorich (Springer, 2004). The standard formulation of the Completeness Axiom appears in Section 6.4, where we prove that Zorichs version is equivalent to the standard one.

It is perhaps surprising that the Completeness Axiom has implications for the sets $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$, and how they are situated within $\mathbb{R}$. This relationship is explored in Section 6.2. In Section 6.3, we prove that $\mathbb{R}$ is uncountable. We emphasize the role of the Completeness Axiom in this proof and aim to minimize the use of other background knowledge.

6.2. The Completeness Axiom and the sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$

The Archimedean Property states that the set $\mathbb{N}$ is not bounded above in $\mathbb{R}$.

THEOREM 6.2.1 (Archimedean Property). *For every $b \in \mathbb{R}$ there exists $n \in \mathbb{N}$ such that $b < n$.*

PROOF. We will prove the statement by contradiction. Assume that the negation of the statement is true. That is, assume that there exists $\beta \in \mathbb{R}$ such that $\beta \geq n$ for all $n \in \mathbb{N}$. Set

$$A = \mathbb{N} \quad \text{and} \quad B = \bigl\{b \in \mathbb{R} : b \geq n \ \ \forall n \in \mathbb{N}\bigr\}.$$

Since $1 \in \mathbb{N}$ and $\beta \in B$, the sets A and B are nonempty subsets of $\mathbb{R}$. By the definition of the set B we have $a \leq b$ for all $a \in A$ and all $b \in B$. By Completeness Axiom there exists $c \in \mathbb{R}$ such that $a \leq c \leq b$ for all $a \in A$ and all $b \in B$. In other words, $n \leq c \leq b$ for all $n \in \mathbb{N}$ and all $b \in B$. Since $c \leq b$ for all $b \in B$, we conclude that $c - 1/2 \notin B$. Thus, there exists $m \in \mathbb{N}$ such that $c - 1/2 < m$. Since $n \leq c$ for all $n \in \mathbb{N}$ we conclude that $m + 1 \leq c$. Hence,

$$c - 1/2 < m < m + 1 \leq c.$$

Using the above inequalities we get

$$1 = (m + 1) - m < c - \bigl(c - 1/2\bigr) = 1/2,$$

that is $2 < 1$. Wrong! (by Exercise 4.1.11) This proves the statement. □

As a corollary we obtain a generalized version of the Archimedean Property.

COROLLARY 6.2.2 (Archimedean Property). *For every $b \in \mathbb{R}$ and every $a \in \mathbb{R}_{>0}$ there exists $n \in \mathbb{N}$ such that $b < na$.*

EXERCISE 6.2.3. For every $a \in \mathbb{R}_{>0}$ there exists $n \in \mathbb{N}$ such that $\dfrac{1}{n} < a$. ◁

EXERCISE 6.2.4. For every $c \in \mathbb{R}$ the following implication holds: If for all $n \in \mathbb{N}$ we have $-\dfrac{1}{n} < c < \dfrac{1}{n}$, then $c = 0$. ◁

The following theorem highlights the distribution of integers among the real numbers.

THEOREM 6.2.5. *Let $\alpha, \beta \in \mathbb{R}$ be such that $\beta - \alpha > 1$. Then there exists $m \in \mathbb{Z}$ such that $\alpha < m < \beta$. That is, $(\alpha, \beta) \cap \mathbb{Z} \neq \emptyset$.*

PROOF. Assume $\beta - \alpha > 1$. Then $\alpha < \alpha + 1 < \beta$.

Case 1. $\alpha \geq 0$. Consider the set $A = \{k \in \mathbb{N} : \alpha < k\}$. By Theorem 6.2.1, the set A is nonempty. By definition, $A \subseteq \mathbb{N}$. By Theorem 5.3.15 (Well Ordering Property), the set A has a minimum. Let $m = \min(A)$. Since $m \in A$, we have $m \in \mathbb{N}$, so either $m - 1 = 0$ or $m - 1 \in \mathbb{N}$. In either case, $m - 1 \leq \alpha$. Therefore,

$$m \leq \alpha + 1 < \alpha + (\beta - \alpha) = \beta.$$

Since $m \in A$, we also have $\alpha < m$, proving the proposition in this case. In fact, here $m \in \mathbb{N}$.

Case 2. $\alpha < 0 < \beta$. In this case, we take $m = 0$.

Case 3. $\alpha < \beta \le 0$. By Theorem 4.1.8(i), we have $0 \le -\beta < -\alpha$. Note that

$$(-\alpha) - (-\beta) = \beta - \alpha > 1.$$

Thus, we can apply the result proved in Case 1 for the numbers $-\beta$ and $-\alpha$. Therefore, there exists $m \in \mathbb{N}$ such that $-\beta < m < -\alpha$. Multiplying this inequality by $-1 < 0$, Theorem 4.1.8(ii) yields $\alpha < -m < \beta$. Since $-m \in \mathbb{Z}$, the proposition holds in this case as well. □

REMARK 6.2.6. (PAY ATTENTION TO A COMMON PROOF STRATEGY). In Theorem 6.2.5, the objective is to prove the existence of an integer with a certain property. Specifically, given α and β, we aim to construct an integer m satisfying the desired condition. Pay close attention to how this is accomplished in Case 1 of the preceding proof:

First, we form a set that must contain the desired integer. In that proof, the set is denoted by A; it consists of all candidates for the solution m.

Second, we observe that this set of candidates is well-behaved in the sense that it possesses an extreme element, in this case, a minimum. A minimum element has useful properties that allow us to prove it satisfies the required condition.

The proof strategy used in Case 1 is both common and powerful. It demonstrates a beautiful and systematic approach to constructing numbers with specific properties. ◁

The following two statements highlight the distribution of rational numbers among the real numbers.

THEOREM 6.2.7. *Let $a, b \in \mathbb{R}$ be such that $a < b$. Then there exists $r \in \mathbb{Q}$ such that $a < r < b$. That is, $(a, b) \cap \mathbb{Q} \neq \emptyset$.*

PROOF. Assume $a, b \in \mathbb{R}$ are such that $a < b$. Then $b - a > 0$. By Corollary 6.2.2 there exists $n \in \mathbb{N}$ such that $n(b - a) > 1$. By Axiom DL, $nb - na > 1$. Now, Theorem 6.2.5 yields that there exists $m \in \mathbb{Z}$ such that $na < m < nb$. Since, $n > 0$, Theorem 4.1.8(vii) yields $1/n > 0$. Multiplying $na < m < nb$ by $1/n > 0$, Axiom OM yields

$$a < \frac{m}{n} < b.$$

By the definition of $\mathbb{Q}$ we have $m/n \in \mathbb{Q}$. □

COROLLARY 6.2.8. *Let $a, b \in \mathbb{R}$ and $a < b$. Then $(a, b) \cap \mathbb{Q}$ is an infinite set.*

PROOF. Let

$$S = (a, b) \cap \mathbb{Q} = \{x \in \mathbb{Q} : a < x < b\}.$$

By Theorem 6.2.7 the set S is not empty. First we will prove that S does not have a minimum. Let $x \in S$ be arbitrary. Then $a < x < b$. Since $a, x \in \mathbb{R}$ and $a < x$, by Theorem 6.2.7 there exists $r \in \mathbb{Q}$ such that $a < r < x$. Since $r \in \mathbb{Q}$ and

$a < r < x < b$, we have $r \in S$. Therefore, for every $x \in S$ there exists $r \in S$ such that $r < x$. Since S does not have a minimum, it is infinite. $\square$

THEOREM 6.2.9. *For every $x \in \mathbb{R}$ we have that:*

(i) *A nonempty bounded below subset of $\mathbb{Z}$ has a minimum.*
(ii) *A nonempty bounded above subset of $\mathbb{Z}$ has a maximum.*

PROOF. (i) Let A be a nonempty subset of $\mathbb{Z}$ and let $a \in \mathbb{R}$ be such that for all $k \in A$ we have $a \leq k$. By the Archimedean Property, Theorem 6.2.1, there exists $n \in \mathbb{N}$ such that $-a < n$, and consequently, $-n < a$. Hence, for all $k \in A$ we have $-n < k$. Therefore, for all $k \in A$ we have $0 < k + n$. Thus, the set $\{k + n : k \in A\}$ is a nonempty subset of $\mathbb{N}$. By the Well Ordering Property (Theorem 5.3.15), the following minimum exists

$$l = \min\{k + n : k \in A\}.$$

It is straightforward to verify that $m = l - n = \min A$.

(ii) Let B be a nonempty bounded above subset of $\mathbb{Z}$. Then the set $C = \{-k : k \in B\}$ is a nonempty bounded below subset of $\mathbb{Z}$. By (i), C has a minimum, say $m = \min C$. It is straightforward to verify that $-m = \max B$. $\square$

COROLLARY 6.2.10. *For every $x \in \mathbb{R}$ we have that:*

(i) *The set $\{k \in \mathbb{Z} : x \leq k\}$ has a minimum.*
(ii) *The set $\{k \in \mathbb{Z} : k \leq x\}$ has a maximum.*

PROOF. The Archimedean Property (Theorem 6.2.1) yields that the set in (i) is nonempty. The set in (i) is bounded below by x. Hence, the claim follows from Theorem 6.2.9(i). The claim in (ii) follows from (i) applied to $-x$. $\square$

Corollary 6.2.10 justifies the definitions of the floor and the ceiling functions, which associate integers with real numbers.

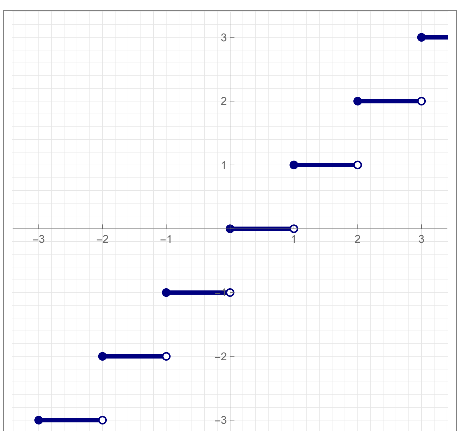


FIG. 6.2.1. The floor function

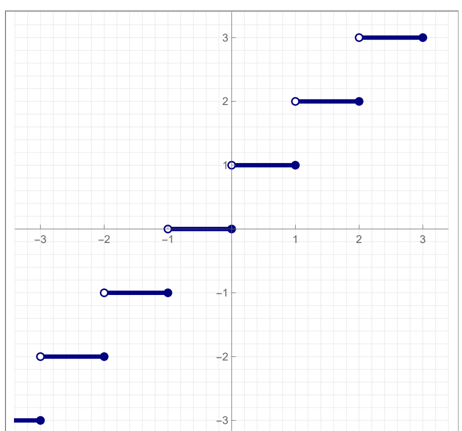


FIG. 6.2.2. The ceiling function

DEFINITION 6.2.11. The *floor* function is defined by

$$\forall x \in \mathbb{R} \qquad \operatorname{flr}(x) = \lfloor x \rfloor \stackrel{\text{def}}{=} \max\bigl\{k \in \mathbb{Z} : k \leq x\bigr\}.$$

The *ceiling* function is defined by

$$\forall x \in \mathbb{R} \qquad \operatorname{clg}(x) = \lceil x \rceil \stackrel{\text{def}}{=} \min\bigl\{k \in \mathbb{Z} : x \leq k\bigr\}.$$

THEOREM 6.2.12. (I) *For all $x \in \mathbb{R}$ the following equalities hold:*
(i) $\lfloor x \rfloor = -\lceil -x \rceil$,
(ii) $\lceil x \rceil = -\lfloor -x \rfloor$,
(II) *Let $x \in \mathbb{R}$ and $m \in \mathbb{Z}$. The following statements are equivalent:*
(i) $m = \lfloor x \rfloor$.
(ii) $m \leq x < m + 1$.
(iii) $x - 1 < m \leq x$.
(III) *Let $x \in \mathbb{R}$ and $n \in \mathbb{Z}$. The following statements are equivalent:*
(i) $n = \lceil x \rceil$.
(ii) $n - 1 < x \leq n$.
(iii) $x \leq n < x + 1$.

PROOF. □

EXERCISE 6.2.13. Let $a, b \in \mathbb{R}$ be such that $a < b$. Then $(a,b) \cap \mathbb{Z}$ is finite and

$$\operatorname{card}\bigl((a,b) \cap \mathbb{Z}\bigr) = \lceil b \rceil - \lfloor a \rfloor - 1. \qquad \triangleleft$$

EXERCISE 6.2.14. Prove that for all $x \in \mathbb{R} \setminus \{0\}$ we have

$$1 - \operatorname{rlu}(-x) \leq x \left\lceil \frac{1}{x} \right\rceil \leq 1 + \operatorname{rlu}(x),$$

where $x \mapsto \operatorname{rlu}(x)$ is the Rectified Linear Unit function introduced in Definition 4.2.5. HINT: Consider $x < 0$ and $x > 0$ separately. ◁

The next exercise offers an alternative proof for Theorem 6.2.7.

EXERCISE 6.2.15. Let $a, b \in \mathbb{R}$.
(i) If $b - a \geq 1$, then

$$a < \frac{\lceil a \rceil + \lfloor b \rfloor}{2} < b.$$

(ii) If $a < b$, then

$$a < \frac{\left\lceil \left\lceil \frac{1}{b-a} \right\rceil a \right\rceil + \left\lfloor \left\lceil \frac{1}{b-a} \right\rceil b \right\rfloor}{2 \left\lceil \frac{1}{b-a} \right\rceil} < b. \qquad \triangleleft$$

SOLUTION. (i) Let $a, b \in \mathbb{R}$ and assume $b - a \geq 1$. Then $a + 1 \leq b$ and $a + 1 \leq b$. The following inequalities hold, the first on being a consequence of $a \leq b - 1$,

$$2a \leq a + b - 1$$

$$\boxed{\text{Since } a \leq \lceil a \rceil \text{ and } b - 1 < \lfloor b \rfloor} < \lceil a \rceil + \lfloor b \rfloor$$

$$\boxed{\text{Since } \lceil a \rceil < a+1 \text{ and } \lfloor b \rfloor \le b} < a+b+1$$
$$\boxed{\text{Since } a+1 \le b} \le 2b.$$

(ii) Let $a, b \in \mathbb{R}$ and assume $b-a>0$. Then $\dfrac{1}{b-a} > 0$. Since $\left\lceil \dfrac{1}{b-a} \right\rceil \ge \dfrac{1}{b-a}$ and $b-a>0$, we deduce $\left\lceil \dfrac{1}{b-a} \right\rceil (b-a) \ge 1$. Hence, applying the result in part (i) to the numbers

$$\left\lceil \frac{1}{b-a} \right\rceil a \quad \text{and} \quad \left\lceil \frac{1}{b-a} \right\rceil b$$

we obtain

$$\left\lceil \frac{1}{b-a} \right\rceil a < \frac{\left\lceil \left\lceil \frac{1}{b-a} \right\rceil a \right\rceil + \left\lfloor \left\lceil \frac{1}{b-a} \right\rceil b \right\rfloor}{2} < \left\lceil \frac{1}{b-a} \right\rceil b.$$

The claim in (ii) follows from the last two inequalities. ☐

In Theorem 5.5.9, we proved that for all $r \in \mathbb{Q}$, we have $r^2 \neq 2$. In the next theorem, we prove what is commonly taught in middle school: that there exists a positive real number whose square is 2. Of course, this result relies on the **Completeness Axiom**.

While much more general statements are true, we present this specific result separately to clearly illustrate how the **Completeness Axiom** is used in the proof.

THEOREM 6.2.16. *There exists a unique positive real number α such that $\alpha^2 = 2$.*

PROOF. **Step 1.** Define two sets of positive real numbers as follows:

$$A = \{x \in \mathbb{R} : x > 0 \quad \wedge \quad x^2 < 2\},$$

$$B = \{y \in \mathbb{R} : y > 0 \quad \wedge \quad 2 < y^2\}.$$

Since $1 \in A$ and $2 \in B$, A and B are nonempty subsets of $\mathbb{R}$. If $x \in A$ and $y \in B$, then $x, y \in \mathbb{R}_{>0}$ and $x^2 < 2 < y^2$. Exercise 4.1.16(ii) yields $x < y$. Hence, for all $x \in A$ and all $y \in B$ we have $x < y$. The Completeness Axiom implies that there exists $c \in \mathbb{R}$ such that $x \le c \le y$ for all $x \in A$ and all $y \in B$. Since $1 \in A$ we have that $1 \le c$. In particular, $c > 0$.

Step 2. The set A has no maximum. To prove this, we show that for every $x \in A$, there exists $\epsilon > 0$ such that $x + \delta \in A$. This amounts to finding a formula for $\delta > 0$, depending on x, for which one can verify that $(x+\delta)^2 < 2$. We leave the details as an exercise.

Step 3. The set B has no minimum. To prove this, we show that for every $y \in B$, there exists $\delta > 0$ such that $y - \delta \in B$. This amounts to finding a formula for $\delta > 0$, depending on y, for which one can verify that $(y-\delta)^2 > 2$. We leave the details as an exercise.

Step 4. In **Step 1** we proved that there exists $c \in \mathbb{R}$ such that $c > 0$ and for all $x \in A$, we have $x \le c$. Thus, c is an upper bound for A. Since A has no maximum, we cannot have $c \in A$, so $c^2 < 2$ is not true. Therefore, $c^2 \ge 2$.

Similarly, in **Step 1** we proved that there exists $c \in \mathbb{R}$ such that $c > 0$ and for all $y \in B$, we have $c \leq y$. Thus, c is a lower bound for B. Since B has no minimum, we cannot have $c \in B$, so $c^2 > 2$ is not true. Therefore, $c^2 \leq 2$.

Hence, we have shown that

$$c^2 \geq 2 \qquad \text{and} \qquad c^2 \leq 2,$$

and thus $c^2 = 2$. This completes the proof of existence. Uniqueness follows from Exercise 4.1.16(i). □

Theorem 6.2.16 justifies the following definition:

DEFINITION 6.2.17. The unique positive real number α such that $\alpha^2 = 2$ is denoted by $\sqrt{2}$ and is called the *square root* of 2.

THEOREM 6.2.18. *The set of rational numbers is a proper subset of the set of real numbers; that is,* $\mathbb{Q} \subsetneq \mathbb{R}$.

PROOF. In Theorem 5.5.9, we proved that $r^2 \neq 2$ for all $r \in \mathbb{Q}$. Since $(\sqrt{2})^2 = 2$, it follows that $\sqrt{2} \notin \mathbb{Q}$. By Theorem 6.2.16, $\sqrt{2} \in \mathbb{R}$, so we have $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$. Since $\mathbb{Q} \subseteq \mathbb{R}$, this shows that $\mathbb{Q} \subsetneq \mathbb{R}$. □

6.3. $\mathbb{R}$ is uncountable

THEOREM 6.3.1 (Nested Intervals Theorem). *For each* $k \in \mathbb{N}$, *let* $I_k \subset \mathbb{R}$ *be a nonempty closed and bounded interval such that* $I_{k+1} \subseteq I_k$ *for all* $k \in \mathbb{N}$. *Then*

$$\bigcap_{k \in \mathbb{N}} I_k \neq \emptyset.$$

That is, there exists $c \in \mathbb{R}$ *such that* $c \in I_k$ *for all* $k \in \mathbb{N}$.

PROOF. By the definition of a finite closed interval, for every $k \in \mathbb{N}$ there exist $a_k, b_k \in \mathbb{R}$ such that $a_k < b_k$ and $I_k = [a_k, b_k]$. Assume that for all $k \in \mathbb{N}$ we have $I_{k+1} \subseteq I_k$. Then, for all $k \in \mathbb{N}$ we have $a_{k+1}, b_{k+1} \in [a_k, b_k]$, and consequently

$$a_k \leq a_{k+1} < b_{k+1} \leq b_k. \tag{6.3.1}$$

Next we will prove that for all $k, m \in \mathbb{N}$ such that $k < m$ we have

$$a_k \leq a_m < b_m \leq b_k. \tag{6.3.2}$$

Let $k \in \mathbb{N}$ be arbitrary and write $m \in \mathbb{N}$ such that $k < m$ as $m = k + l$ with $l \in \mathbb{N}$. We will prove that for all $l \in \mathbb{N}$ we have

$$a_k \leq a_{k+l} < b_{k+l} \leq b_k. \tag{6.3.3}$$

Denote the statement in (6.3.3) by $P(l)$ and prove it by Mathematical Induction. The statement $P(1)$ is proved earlier, see (6.3.1). Let $l \in \mathbb{N}$ be arbitrary and assume that $P(l)$ is true, that is assume (6.3.3). Applying (6.3.1) with k replaced by $k + l$ yields

$$a_{k+l} \leq a_{k+l+1} < b_{k+l+1} \leq b_{k+l}. \tag{6.3.4}$$

Now (6.3.3) and (6.3.4) and Axiom OT yield

$$a_k \le a_{k+l+1} < b_{k+l+1} \le b_k.$$

Thus, $P(l+1)$ holds. This proves (6.3.2) for all $k, m \in \mathbb{N}$ such that $k < m$.

It follows from (6.3.2) that we have

$$\forall m \in \mathbb{N} \ \ \forall n \in \mathbb{N} \qquad a_m < b_n. \tag{6.3.5}$$

Set

$$A = \{a_m : m \in \mathbb{N}\} \qquad \text{and} \qquad B = \{b_n : n \in \mathbb{N}\}.$$

Clearly A and B are nonempty subsets of $\mathbb{R}$ and (6.3.5) shows that we can apply the Completeness Axiom to the sets A and B. By the Completeness Axiom we deduce that there exists $c \in \mathbb{R}$ such that

$$\forall m \in \mathbb{N} \ \ \forall n \in \mathbb{N} \qquad a_m \le c \le b_n.$$

In particular

$$\forall n \in \mathbb{N} \qquad a_n \le c \le b_n.$$

Thus $c \in [a_n, b_n] = I_n$ for all $n \in \mathbb{N}$. □

The inequalities proved in the following lemma are illustrated in Figure 6.3.1 on the right.

LEMMA 6.3.2. *Let $u, v \in \mathbb{R}$ be such that $u < v$. Then*

$$u < \tfrac{2}{3}u + \tfrac{1}{3}v < \tfrac{1}{2}u + \tfrac{1}{2}v < \tfrac{1}{3}u + \tfrac{2}{3}v < v.$$

PROOF. Assume $u < v$. Since $2 > 0$, Axiom OM yields $2u < 2v$, and Axioms OA, and DL imply $6u < 4u + 2v$. Axioms OA, and DL and $u < v$ imply $4u + 2v < 3u + 3v$. Axioms OA, and DL and $u < v$ also imply $3u + 3v < 2u + 4v$. Finally, $2u < 2v$, and Axioms OA, and DL imply $2u + 4v < 6v$. Now, Axiom OT yields

$$6u < 4u + 2v < 3u + 3v < 2u + 4v < 6v.$$

Since $1/6 > 0$, Axioms OM and DL yield the desired inequalities. □

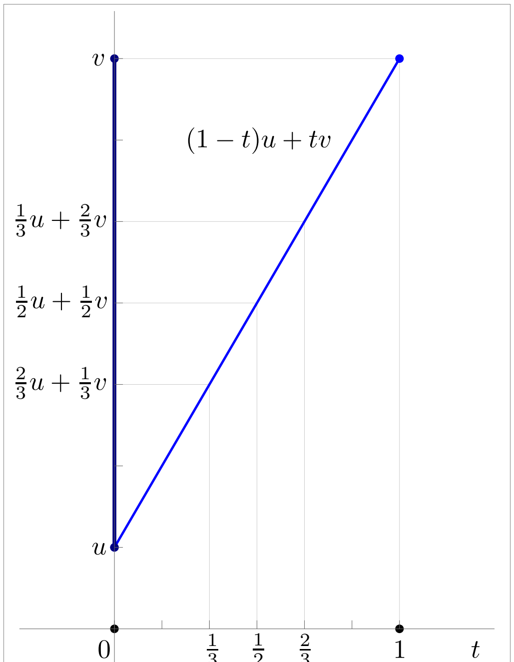


FIG. 6.3.1. $u < \frac{2}{3}u + \frac{1}{3}v < \frac{1}{2}u + \frac{1}{2}v < \frac{1}{3}u + \frac{2}{3}v < v$

In the next theorem, we prove that the set of real numbers is uncountable.

THEOREM 6.3.3. *Let $f : \mathbb{N} \to \mathbb{R}$ be a sequence of real numbers. Then, for every interval $I \subseteq \mathbb{R}$, there exists $c \in I$ such that c is not in the range of f. In particular, f is not a surjection.*

PROOF. Let $f : \mathbb{N} \to \mathbb{R}$ be an arbitrary sequence of real numbers. Let I be an arbitrary interval.

We recursively define a sequence of nested closed intervals $I_k = [a_k, b_k]$ where $k \in \mathbb{N}$, $a_k, b_k \in [a_0, b_0]$ and $a_k < b_k$. This sequence of intervals will be defined so that it has a special relation to the sequence $f : \mathbb{N} \to \mathbb{R}$:

$$\forall k \in \mathbb{N} \quad f_k \notin [a_k, b_k]. \tag{6.3.6}$$

Base step: For $k = 0$, let $a_0, b_0 \in I$ be such that $a_0 < b_0$. Then, $[a_0, b_0] \subseteq I$.

Recursive step: For each $k \in \mathbb{N}$ consider two cases based on the placement of f_k relative to the midpoint of the preceding interval $[a_{k-1}, b_{k-1}]$:

$$f_k \leq \tfrac{1}{2}(a_{k-1} + b_{k-1}) \quad \text{or} \quad f_k > \tfrac{1}{2}(a_{k-1} + b_{k-1}).$$

Define

$$\big[a_k, b_k\big] = \begin{cases} \Big[\tfrac{1}{3}a_{k-1} + \tfrac{2}{3}b_{k-1}, b_{k-1}\Big] & \text{if} \quad f_k \leq \tfrac{1}{2}(a_{k-1} + b_{k-1}), \\ \Big[a_{k-1}, \tfrac{2}{3}a_{k-1} + \tfrac{1}{3}b_{k-1}\Big] & \text{if} \quad f_k > \tfrac{1}{2}(a_{k-1} + b_{k-1}). \end{cases}$$

The cases considered in this definition are illustrated in Figures 6.3.2 and 6.3.3.

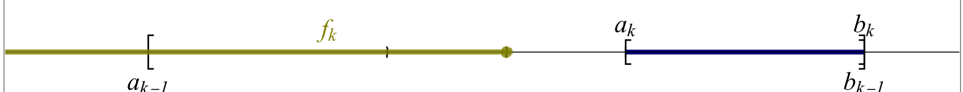


FIG. 6.3.2. If f_k is in the olive interval, choose the navy interval as $\big[a_k, b_k\big]$

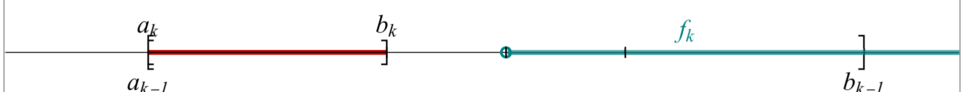


FIG. 6.3.3. If f_k is in the teal interval, choose the marron interval as $\big[a_k, b_k\big]$

If $f_k \leq \frac{1}{2}(a_{k-1} + b_{k-1})$, by Lemma 6.3.2, the first line in the definition of $\big[a_k, b_k\big]$ yields:

$$a_{k-1} < a_k = \tfrac{1}{3}a_{k-1} + \tfrac{2}{3}b_{k-1} < b_{k-1} = b_k,$$

and

$$f_k \leq \tfrac{1}{2}(a_{k-1} + b_{k-1}) < a_k = \tfrac{1}{3}a_{k-1} + \tfrac{2}{3}b_{k-1}.$$

Hence,

$$a_k < b_k, \quad \big[a_k, b_k\big] \subset \big[a_{k-1}, b_{k-1}\big], \quad \text{and} \quad f_k \notin \big[a_k, b_k\big]. \tag{6.3.7}$$

Similarly, if $f_k > \frac{1}{2}(a_{k-1}+b_{k-1})$, by Lemma 6.3.2, the second line in the definition of $\big[a_k, b_k\big]$ yields

$$a_k = a_{k-1} < \tfrac{2}{3}a_{k-1} + \tfrac{1}{3}b_{k-1} = b_k < b_{k-1},$$

and

$$b_k = \tfrac{2}{3}a_{k-1} + \tfrac{1}{3}b_{k-1} < \tfrac{1}{2}(a_{k-1} + b_{k-1}) < f_k.$$

Thus, (6.3.7) holds in this case as well. Since $k \in \mathbb{N}$ was arbitrary, (6.3.7) holds for all $k \in \mathbb{N}$.

In conclusion, for every $k \in \mathbb{N}$, the interval $I_k = \big[a_k, b_k\big]$ is a bounded closed interval such that $I_k \subseteq I_{k-1}$, meeting all assumptions of the Nested Intervals Theorem, Theorem 6.3.1. Consequently, there exists $c \in \mathbb{R}$ such that $c \in I_k$ for all $k \in \mathbb{N}$. Together with the last statement in (6.3.7), we have:

$$\forall k \in \mathbb{N} \qquad c \in I_k \quad \text{and} \quad f_k \notin I_k.$$

Therefore, for all $k \in \mathbb{N}$, $c \neq f_k$. Since

$$c \in I_1 \subset [a_0, b_0] \subseteq I,$$

the theorem is proved. □

6.4. Infimum and supremum

One of the main features of finite sets is that each finite set has a minimum and a maximum. As we saw in Theorem 5.3.10, every infinite set has a subset that does not have a minimum or a subset that does not have a maximum. This is a problem with infinite sets. To some extent, not having a maximum or minimum is inevitable for an infinite set. Therefore, we have to seek surrogates for the minimum and maximum. That is what the infimum and supremum are. They are surrogates for minimums and maximums.

6.4.1. Definitions and basic properties.

Definition 6.4.1. For a nonempty subset A of $\mathbb{R}$, a number $\alpha \in \mathbb{R}$ is a *infimum* (or a *greatest lower bound*) of A if α has the following two properties:

(i) For all $x \in A$ we have $\alpha \leq x$.
(That is, α is a lower bound for A.)

(ii) For every $\epsilon > 0$ there exists $x \in A$ such that $x < \alpha + \epsilon$.
(That is, $\alpha + \epsilon$ is NOT a lower bound for A.)

Definition 6.4.2. For a nonempty subset A of $\mathbb{R}$, a number $\beta \in \mathbb{R}$ is a *supremum* (or a *least upper bound*) of A if β has the following two properties:

(i) For all $x \in A$ we have $x \leq \beta$.
(That is, β is an upper bound for A.)

(ii) For every $\epsilon > 0$ there exists $x \in A$ such that $\beta - \epsilon < x$.
(That is, $\beta - \epsilon$ is NOT an upper bound for A.)

If α and β are as in Definitions 6.4.1 and 6.4.2, we write

$$\alpha = \inf A \,(\, = \operatorname{glb} A) \qquad \text{and} \qquad \beta = \sup A \ (\, = \operatorname{lub} A\ ).$$

Theorem 6.4.3. *Let A be a nonempty subset of $\mathbb{R}$. If $\inf A$ exists, it is unique. If $\sup A$ exists, it is unique.*

Proof. Assume that $\alpha \in \mathbb{R}$ satisfy (i) and (ii) in Definition 6.4.1. We will prove: If $\alpha' \in \mathbb{R} \setminus \{\alpha\}$, then α' is not an infimum of A. Assume $\alpha' \in \mathbb{R} \setminus \{\alpha\}$.

Then, By Axiom OE, we have two cases $\alpha' < \alpha$ and $\alpha < \alpha'$. If $\alpha' < \alpha$, then for $\epsilon = (\alpha - \alpha')/2 > 0$, by (i) in Definition 6.4.1, we have

$$\alpha' + \epsilon = \frac{\alpha' + \alpha}{2} < \alpha \leq x$$

for all $x \in A$. Thus, α' does not satisfy (ii) in Definition 6.4.1.

If $\alpha < \alpha'$, then, setting $\epsilon = \alpha' - \alpha > 0$, by (ii) in Definition 6.4.1, we have that there exists $x \in A$ such that

$$x < \alpha + \epsilon = \alpha'.$$

That is, α' does not satisfy (i) in Definition 6.4.1.

The proof of uniqueness of the supremum is similar. □

The following theorem gives the standard form of the **Completeness Axiom**.

THEOREM 6.4.4. *Every nonempty subset of* $\mathbb{R}$ *that is bounded above has a supremum.*

PROOF. Let $A \subseteq \mathbb{R}$ be a nonempty set and assume that A is bounded above. That is, there exists $b \in \mathbb{R}$ such that for all $x \in A$ we have $x \leq b$. Denote by B the set of all upper bounds of A. That is,

$$B = \{ y \in \mathbb{R} : \forall x \in A \ \ x \leq y \}.$$

Clearly, $b \in B$, so, B is a nonempty subset of $\mathbb{R}$. By the definition of B we have

$$\forall x \in A \ \ \forall y \in B \qquad x \leq y.$$

We proved that the sets A and B satisfy all the assumptions of the Completeness Axiom. The application of the Completeness Axiom yields that there exists $c \in \mathbb{R}$ such that

$$\forall x \in A \ \ \forall y \in B \qquad x \leq c \leq y.$$

Hence, $c \in \mathbb{R}$ satisfies (i) in Definition 6.4.2.

Since for all $x \in A$ we have $x \leq c$, by the definition of B, we conclude $c \in B$. Since for all $y \in B$ we have $c \leq y$ we conclude $c = \min(B)$. Thus, $c \in \mathbb{R}$ is the minimum of the set of all upper bounds of A. Consequently, for an arbitrary $\epsilon > 0$ we have that $c - \epsilon < c$ is not an upper bound for A. Therefore, there exists $x \in A$ such that $c - \epsilon < x$. Hence, $c \in \mathbb{R}$ satisfies (ii) in Definition 6.4.2, proving that

$$c = \sup(A).$$

This proves the existence of a supremum of a nonempty subset of $\mathbb{R}$ that is bounded above. □

THEOREM 6.4.5. *Every nonempty subset of* $\mathbb{R}$ *that is bounded below has an infimum.*

PROOF. Let $B \subseteq \mathbb{R}$ be a nonempty set and assume that B is bounded below. That is, there exists $a \in \mathbb{R}$ such that for all $y \in B$ we have $a \leq y$. Denote by A the set of all lower bounds of B. That is,

$$A = \{ x \in \mathbb{R} : \forall y \in B \ \ x \leq y \}.$$

Clearly, $a \in A$, so, A is a nonempty subset of $\mathbb{R}$. By the definition of A we have

$$\forall x \in A \ \ \forall y \in B \qquad x \leq y.$$

We proved that the sets A and B satisfy all the assumptions of the Completeness Axiom. The application of the Completeness Axiom yields that there exists $c \in \mathbb{R}$ such that

$$\forall x \in A \ \ \forall y \in B \qquad x \leq c \leq y.$$

Hence, $c \in \mathbb{R}$ satisfies (i) in Definition 6.4.1.

Since for all $y \in B$ we have $c \leq y$, by the definition of A, we conclude $c \in A$. Since for all $x \in A$ we have $x \leq c$ we conclude $c = \max(A)$. Thus, $c \in \mathbb{R}$ is the maximum of the set of all lower bounds of B. Consequently, for an arbitrary $\epsilon > 0$ we have that $c + \epsilon < c$ is not a lower bound for B. Therefore, there exists $y \in B$ such that $y < c + \epsilon$. Hence, $c \in \mathbb{R}$ satisfies (ii) in Definition 6.4.1, proving that

$$c = \inf(B).$$

This proves the existence of an infimim of a nonempty subset of $\mathbb{R}$ that is bounded below. □

Exercise 6.4.6. Let A be a nonempty subset of $\mathbb{R}$ that is bounded above. Then the set of all upper bounds of A is the closed unbounded interval, see Figure 6.4.1,

$$\mathbb{R}_{\geq \sup(A)} = \bigl\{ y \in \mathbb{R} : \sup(A) \leq y \bigr\}. \qquad \triangleleft$$

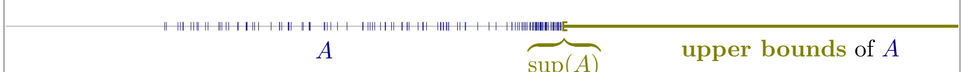


Fig. 6.4.1. The **upper bounds** and the **supremum** of the **set** A

Exercise 6.4.7. Let A be a nonempty subset of $\mathbb{R}$ that is bounded below. Then the set of all lower bounds of A is the closed unbounded interval, see Figure 6.4.2,

$$\mathbb{R}_{\leq \inf(A)} = \bigl\{ x \in \mathbb{R} : x \leq \inf(A) \bigr\}. \qquad \triangleleft$$

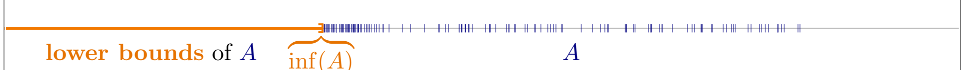


Fig. 6.4.2. The **lower bounds** and the **infimum** of the **set** A

Remark 6.4.8. Here I describe the setting in Figures 6.4.1 and 6.4.2. The gray line represents the real number line. The navy blue set A is a set of real numbers, in general infinite. The numbers in the set A are represented by navy blue ticks. Thus, the numbers in A are indicated by points where the ticks intersect the gray line. Keep in mind that it is not possible to depict infinitely many numbers in a picture. In Figure 6.4.1, the upper bounds of A are colored **olive**. The minimum of these upper bounds, the supremum, is marked by a small **olive** square bracket. In Figure 6.4.2, the lower bounds of A are colored **orange**. The maximum of these lower bounds, the infimum, is marked by a small **orange** square bracket. $\triangleleft$

Exercise 6.4.9. (i) Let A be a nonempty subset of $\mathbb{R}$ that is bounded above. Then $\sup(A) \in A$ if and only if $\sup(A) = \max(A)$.

(ii) Let A be a nonempty subset of $\mathbb{R}$ that is bounded below. Then $\inf(A) \in A$ if and only if $\inf(A) = \min(A)$. $\lhd$

SOLUTION. (i) Assume $\beta = \sup(A) \in A$. Then, for all $x \in A$, we have $x \le \beta$ and $\beta \in A$. Hence, by the definition of maximum, $\beta = \max(A)$. Conversely, if $\beta = \sup(A) = \max(A)$, then by the definition of $\max(A)$, we have $\beta \in A$. Hence, $\sup(A) = \beta \in A$.

A proof of (ii) is similar. $\square$

EXERCISE 6.4.10. Find the supremum and infimum of the sets

$$A = \left\{ \frac{1}{n} : n \in \mathbb{N} \right\} \qquad \text{and} \qquad B = \left\{ \frac{n}{n+1} : n \in \mathbb{N} \right\}.$$

Formal proofs are required. (By a formal proof, I mean a rigorous mathematical proof of properties (i) and (ii) in Definitions 6.4.2 and 6.4.1.) $\lhd$

EXERCISE 6.4.11. Find the supremum and infimum of the set $\left\{ \dfrac{n^{(-1)^n}}{n+1} : n \in \mathbb{N} \right\}$. $\lhd$

EXERCISE 6.4.12. Let A be a nonempty and bounded above subset of $\mathbb{R}$. If B is a nonempty subset of A, then B is bounded above and $\sup(B) \le \sup(A)$. Formulate the corresponding statement for the infimums. $\lhd$

EXERCISE 6.4.13. Let A and B be nonempty bounded above subsets of $\mathbb{R}$. Prove

$$\sup(A \cup B) = \max\bigl\{\sup(A), \sup(B)\bigr\}. \qquad \lhd$$

EXERCISE 6.4.14. Let A and B be nonempty subsets of $\mathbb{R}$ such that for all $x \in A$ and for all $y \in B$ we have $x \le y$. Prove that $\sup(A) \le \inf(B)$.

If the condition $x \le y$ is replaced by the condition $x < y$, can we conclude that $\sup(A) < \inf(B)$? $\lhd$

EXERCISE 6.4.15. Suppose that A and B are nonempty subsets of $\mathbb{R}$ such that for all $x \in A$ and for all $y \in B$ we have $x \le y$. Prove that $\sup(A) = \inf(B)$ if and only if for each $\delta > 0$ there exist $x \in A$ and $y \in B$ such that $x + \delta > y$. $\lhd$

EXERCISE 6.4.16. Let A be a nonempty and bounded above subset of $\mathbb{R}$, and let F be a finite subset of A. If $(\sup A) \notin A$, then $\sup(A \setminus F) = \sup(A)$.

State and prove an analogous statement for $\inf(A)$? $\lhd$

6.4.2. The Completeness Axiom in action.

THEOREM 6.4.17. *Let $a, b \in \mathbb{R}$ with $a < b$. Let $f : [a,b] \to \mathbb{R}$ be a function, and let $K > 0$ be such that*

$$\forall\, x, y \in [a,b] \qquad \bigl|f(x) - f(y)\bigr| \le K|x-y|. \tag{6.4.1}$$

If $f(a) < 0$ and $f(b) > 0$, then there exists $\alpha \in (a,b)$ such that $f(\alpha) = 0$.

PROOF. Define the set

$$A = \bigl\{ x \in [a,b] : f(x) < 0 \bigr\}.$$

Since $a \in A$ we have $A \neq \emptyset$. Since $A \subset [a,b)$, A is bounded above. Therefore $\sup A$ exists. Set $\alpha = \sup A$. Since b is an upper bound for A, we have $\alpha \le b$.

Step 1. We will prove that A does not have a maximum. Let $x \in A$ be arbitrary. Then $f(x) < 0$. Set

$$\delta = \min\left\{-\frac{f(x)}{2K}, \frac{b-x}{2}\right\}.$$

Since $f(x) < 0, K > 0$, we have $-\frac{f(x)}{2K} > 0$, and since $x < b$, $\frac{b-x}{2} > 0$. Therefore $\delta > 0$. Since $\delta < \frac{b-x}{2}$, we have $x + \delta \in (x, b)$. From (6.4.1), it follows

$$f(x+\delta) \leq f(x) + K\delta \leq f(x) - K\frac{f(x)}{2K} = f(x) - \frac{f(x)}{2} = \frac{f(x)}{2} < 0.$$

Hence, for every $x \in A$ there exists $\delta > 0$ such that $x + \delta \in a$, proving that A does not have a maximum. Consequently, $f(\alpha) \geq 0$.

Step 2. Since $\alpha = \sup A$, for every $\epsilon > 0$ there exists $x \in A$ such that $\alpha - \epsilon < x < \alpha$. From (6.4.1), we have

$$f(\alpha) - K|x - \alpha| \leq f(x) < 0.$$

Since

$$|x - \alpha| = \alpha - x < \epsilon$$

we have

$$f(\alpha) - K\epsilon < f(\alpha) - K|x - \alpha| < 0.$$

Hence

$$f(\alpha) < K\epsilon$$

or. equivalently

$$f(\alpha)/K < \epsilon$$

Since $\epsilon > 0$ was arbitrary, we have proved

$$\forall \epsilon > 0 \quad \text{we have} \quad f(\alpha)/K < \epsilon.$$

By Exercise 4.1.20, the last statement implies $f(\alpha)/K \leq 0$. SInce $K > 0$, it follows that $f(\alpha) \leq 0$.

In conclusion, since we proved $f(\alpha) \geq 0$ and $f(\alpha) \leq 0$, we deduce $f(\alpha) = 0$. □

Exercise 6.4.18. For all $x, y \in [0,1]$ and all $n \in \mathbb{N}$ we have

$$\left|x^n - y^n\right| \leq n|x - y|.$$ ◁

Exercise 6.4.19. Let $c \in \mathbb{R}_{>0}$ and $n \in \mathbb{N}$. Prove that there exists a unique $\alpha \in \mathbb{R}_{>0}$ such that $\alpha^n = c$.

Hint: First consider $c \in (0,1)$. Apply Theorem 6.4.17 to the function $f(x) = x^n - c$ defined for all $x \in [0,1]$. Exercise 6.4.18 is important here to establish (6.4.1). ◁

Theorem 6.4.20. *Let $n \in \mathbb{N}$. The function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined by*

$$\forall x \in \mathbb{R}_{\geq 0} \qquad f(x) = x^n$$

is a bijection.

Proof. □

DEFINITION 6.4.21. The inverse of the bijection $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ from Theorem 6.4.20 is called the *n-th root function*. For $x \in \mathbb{R}_{\geq 0}$ the value of the n-th root function at x is denoted by $\sqrt[n]{x}$ and it is called the *n-th root of* x.

In conclusion, the set $\mathbb{R}$ is completely described by sixteen axioms: five **Axioms** of addition **AE**, **AA**, **AC**, **AZ**, **AO**, five **Axioms** of multiplication **ME**, **MA**, **MC**, **MO**, **MR**, **Axiom DL**, four **Axioms** of order **OE**, **OT**, **OA**, **OM**, and in its full glory the **Completeness Axiom**. All claims about real numbers can be proved using these sixteen axioms. That is what we attempted to illustrate in Chapters 4, 5, and 6. As you probably noticed, in proofs we also use elementary properties of sets and operations with sets.

CHAPTER 7

The topology of $\mathbb{R}$

7.1. The topology of $\mathbb{R}$

The terminology that we introduce in the next definition provides the essential vocabulary of the modern analysis.

DEFINITION 7.1.1. All points in this definition are elements of $\mathbb{R}$ and all sets are subsets of $\mathbb{R}$.

(i) Let $a \in \mathbb{R}$ and $\epsilon > 0$. A *neighborhood* (or ϵ-*neighborhood*) of a point a is the set

$$N(a,\epsilon) = \bigl\{x \in \mathbb{R} : |x-a| < \epsilon\bigr\} = (a-\epsilon, a+\epsilon)$$

The number ϵ is called the *radius* of $N(a,\epsilon)$.

(ii) Let $a \in \mathbb{R}$ and $\epsilon > 0$. The neighbourhood $N(a,\epsilon)$ with the point a removed, that is the set $N(a,\epsilon) \setminus \{a\}$, is called a *punctured neighborhood* of a point a.

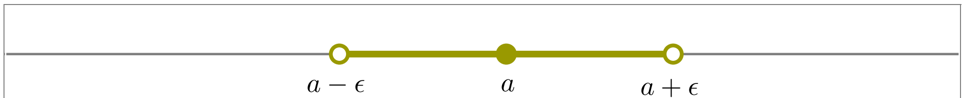


FIG. 7.1.1. A **neighborhood** of a

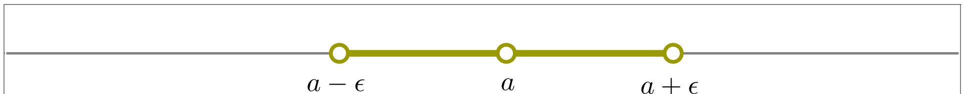


FIG. 7.1.2. A **punctured neighborhood** of a

DEFINITION 7.1.2. (i) A point $a \in \mathbb{R}$ is an *accumulation point* of a set $E \subseteq \mathbb{R}$ if every neighborhood of a contains a point $x \neq a$ such that $x \in E$. That is, a is an *accumulation point* of the set E if

$$\forall \epsilon > 0 \qquad E \cap \bigl(N(a,\epsilon) \setminus \{a\}\bigr) \neq \emptyset. \tag{7.1.1}$$

(ii) The set of all accumulation points of a set E is denoted by E'.

(iii) A set E is *closed* if it contains all its accumulation points. That is, E is closed if the inclusion $E' \subseteq E$ holds.

(iv) The *closure* of a set E is defined as the union of $E \cup E'$. It is denoted by $\overline{E}$.

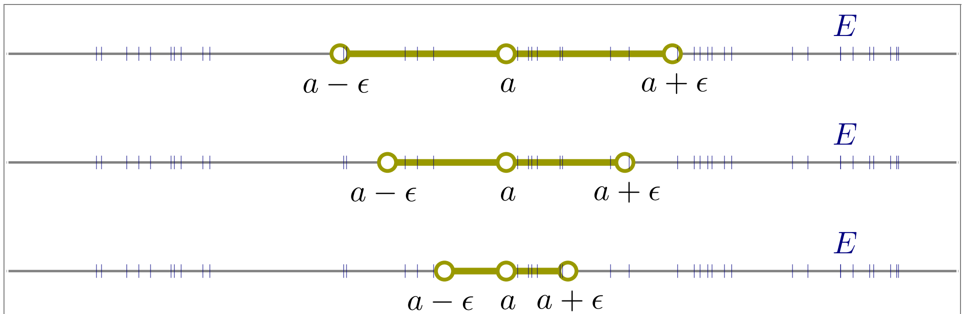


FIG. 7.1.3. Testing a point a to be an accumulation point of E

DEFINITION 7.1.3. (i) A point a is an *interior point* of the set E if there is a neighborhood of a that is a subset of E. That is, a is an *interior point* of E if there exists $\epsilon > 0$ such that $N(a, \epsilon) \subseteq E$.
(ii) The set of all the interior points of a set E is called the *interior* of E. It is denoted by E°.
(iii) A set E is *open* if every point of E is an interior point of E. That is, E is open if $E = E^\circ$.

DEFINITION 7.1.4. The *boundary* of a set E is the set $\overline{E} \setminus E^\circ$. It is denoted by ∂E.

DEFINITION 7.1.5. A set E is *compact* if every infinite subset of E has an accumulation point in E.

DEFINITION 7.1.6. Let $E \subseteq F$. A set E is *dense* in F if every neighborhood of every point in F contains a point of E.

The next proposition presents two important general settings in which accumulation points occur.

PROPOSITION 7.1.7. *Let A be a nonempty subset of* $\mathbb{R}$.
(i) *If A is bounded above and does not have a maximum, then* $\sup(A)$ *is an accumulation point of A.*
(ii) *If A is bounded below and does not have a minimum, then* $\inf(A)$ *is an accumulation point of A.*

PROOF. (i) Let A be a nonempty, bounded above subset of $\mathbb{R}$ and assume that A does not have a maximum. By Theorem 6.4.4, $\beta = \sup(A)$ exists. Since A does not have a maximum, by Exercise 6.4.9, $\beta \notin A$. To prove that β is an accumulation point of A we need to prove the statement in (7.1.1) with a replaced by β.

Let $\epsilon > 0$ be arbitrary. Since $\beta = \sup(A)$ is the minimum of the set of all upper bounds of A, we deduce that $\beta - \epsilon$ is not an upper bound for A. Therefore, there exists $x \in A$ such that $\beta - \epsilon < x$. Since β is an upper bound for A and $\beta \notin A$, we have $x < \beta$. Thus, $x \in A \cap (\beta - \epsilon, \beta)$. Since $(\beta - \epsilon, \beta) \subset N(\beta, \epsilon) \setminus \{\beta\}$, we have

$$x \in A \cap \bigl(N(\beta, \epsilon) \setminus \{\beta\}\bigr) \neq \emptyset.$$

Since $\epsilon > 0$ was arbitrary, we proved (7.1.1) for β. Hence, β is an accumulation point of A.

A proof of (ii) is similar. □

PROPOSITION 7.1.8. *Let $n \in \mathbb{N}$ and let A, B, $A_1, \ldots, A_n$ be subsets of $\mathbb{R}$. Then:*

(i) *If $A \subseteq B$, then $A' \subseteq B'$.*

(ii) $(A \cup B)' = A' \cup B'$.

(iii) $$\left(\bigcup_{k=1}^{n} A_k\right)' = \bigcup_{k=1}^{n} (A_k)'.$$

PROOF. □

EXERCISE 7.1.9. Consider the set $E = \{0, 1\}$. Find E', $\overline{E}$, E° and ∂E. ◁

EXERCISE 7.1.10. Find $\mathbb{Z}'$, $\overline{\mathbb{Z}}$, $\mathbb{Z}^\circ$ and $\partial \mathbb{Z}$. ◁

EXERCISE 7.1.11. Prove that for an arbitrary finite subset F of $\mathbb{R}$ we have $F' = \emptyset$. ◁

EXERCISE 7.1.12. Let $A \subset \mathbb{R}$ be a nonempty bounded set. Prove

$$\inf(A) \in \partial A \qquad \text{and} \qquad \sup(A) \in \partial A.$$ ◁

EXERCISE 7.1.13. Find all accumulation points of the set $\left\{ \dfrac{n^{(-1)^n}}{n+1} : n \in \mathbb{N} \right\}$. Provide formal proofs. ◁

EXERCISE 7.1.14. Find all accumulation points of $\left\{ \dfrac{4}{n} + \dfrac{n}{4} - \left\lfloor \dfrac{n}{4} \right\rfloor : n \in \mathbb{N} \right\}$. Provide formal proofs. ◁

EXERCISE 7.1.15. Let $E \subset \mathbb{R}$. Prove $\overline{E} = E \cup \partial E$. ◁

EXERCISE 7.1.16. Let $a < b$. Prove that the open interval (a, b) is an open set. Prove that the complement of (a, b), that is the set $\mathbb{R} \setminus (a, b)$ is closed. (HINT: State the contrapositive of the implication in the definition of a closed set. Simplify the contrapositive using the concept of an interior point.) ◁

EXERCISE 7.1.17. Let $a < b$. Prove that the closed interval $[a, b]$ is a closed set. Prove that the complement of $[a, b]$, that is the set $\mathbb{R} \setminus [a, b]$ is open. ◁

EXERCISE 7.1.18. Let $a < b$. Consider the interval $[a, b)$. Is this a closed set? Is it open? ◁

EXERCISE 7.1.19. Is $\mathbb{R}$ a closed set? Is it open? ◁

THEOREM 7.1.20. *Prove that $E \subseteq \mathbb{R}$ is closed if and only if $\mathbb{R} \setminus E$ is open.*

PROOF. If E is the empty set, then the equivalence in the theorem is true. Similarly, if $E = \mathbb{R}$, the equivalence is true as well.

In this proof, for $A \subseteq \mathbb{R}$ we use the notation $A^c = \mathbb{R} \setminus A$.Let $A, B \subseteq \mathbb{R}$. Recall that by Exercise 3.1.9 we have $A \subseteq B$ if and only if $B^c \subseteq A^c$.

Assume that $E \neq \emptyset$ and $E \subset \mathbb{R}$ (a proper subset of $\mathbb{R}$). The statement: E is closed is by definition equivalent to $E' \subseteq E$, which, in turn, is equivalent to $E^c \subseteq (E')^c$. The last inclusion is equivalent to

$$\forall a \in E^c \quad \text{we have} \quad a \notin E'.$$

The last displayed statement is equivalent to

$$\forall a \in E^c \quad \exists \epsilon > 0 \quad \text{such that} \quad E \cap \big(N(a,\epsilon) \setminus \{a\}\big) = \emptyset,$$

which, by Exercise 3.1.9, is equivalent to

$$\forall a \in E^c \quad \exists \epsilon > 0 \quad \text{such that} \quad N(a,\epsilon) \setminus \{a\} \subseteq E^c.$$

Since in the preceding displayed statement we consider $a \in a \in E^c$, this statement is equivalent to

$$\forall a \in E^c \quad \exists \epsilon > 0 \quad \text{such that} \quad N(a,\epsilon) \subseteq E^c,$$

which is equivalent to E^c being an open set. In conclusion, through six equivalences, we proved that E is closed if and only if E^c is an open set. □

THEOREM 7.1.21. *A bounded closed interval is a compact set.*

PROOF. Let I be a bounded closed interval. By the definition of a bounded closed interval, there exist $a, b \in \mathbb{R}$ such that $a < b$ and $I = [a, b]$. Let S be an arbitrary infinite subset of I. By Theorem 5.3.10 there exists a nonempty subset A of S such that A does not have a minimum or there exists a nonempty subset B of S such that B does not have a maximum. Consider two cases.

CASE 1. There exists a nonempty $A \subseteq S$ such that A does not have a minimum. Since $A \subseteq S \subseteq I = [a, b]$, we have

$$\forall x \in A \qquad a \leq x \leq b, \tag{7.1.2}$$

that is A is bounded. By Theorem 6.4.5, $\alpha = \inf(A)$ exists. Since α is the maximum lower bound of A, and by (7.1.2), a is a lower bound of A, we have $a \leq \alpha$. Since α is a lower bound of A, (7.1.2) yields, $\alpha \leq b$. Hence, $\alpha \in I$. Since A does not have a minimum, by Exercise 6.4.9, $\alpha \notin A$. By Proposition 7.1.7, α is an accumulation point of A. Since $A \subset S$, Proposition 7.1.8(i) yields that α is an accumulation point of S. Since $\alpha \in I$, S has an accumulation point that belongs to I.

CASE 2. There exists a nonempty $B \subseteq S$ such that B does not have a maximum. In this case, similarly as in CASE 1, we can prove that $\beta = \sup(A)$ exists, that β is an accumulation point of S and $\beta \in I$. That is S has an accumulation point that belongs to I.

Since S was an arbitrary infinite subset of I, and we proved that S has an accumulation point that belongs to I, this proves that I is compact. □

THE SECOND PROOF OF THEOREM 7.1.21. In this proof we use the Nested Intervals Theorem, Theorem 6.3.1. Let E be an arbitrary infinite subset of $[a,b]$. For every $n \in \mathbb{N} \cup \{0\}$, we will recursively define a closed intervals $[a_n, b_n]$. Set

$$[a_0, b_0] = [a, b],$$

and for all $n \in \mathbb{N}$, set

$$\big[a_n, b_n\big] = \begin{cases} \left[a_{n-1}, \dfrac{a_{n-1}+b_{n-1}}{2}\right] & \text{if } \ E \cap \left[a_{n-1}, \dfrac{a_{n-1}+b_{n-1}}{2}\right] \ \text{ is an infinite set,} \\ \\ \left[\dfrac{a_{n-1}+b_{n-1}}{2}, b_{n-1}\right] & \text{if } \ E \cap \left[a_{n-1}, \dfrac{a_{n-1}+b_{n-1}}{2}\right] \ \text{ is a finite set.} \end{cases}$$

(Please draw your own pictures corresponding to the above definition. Use Figures 6.3.2 and 6.3.3 on page 104 as a guide.)

The sequence of intervals that we just defined has the following three properties:

(I) For all $n \in \mathbb{N} \cup \{0\}$ the intersection $E \cap \big[a_n, b_n\big]$ is infinite.

(II) For all $n \in \mathbb{N}$ we have $\big[a_{n-1}, b_{n-1}\big] \subset \big[a_n, b_n\big]$.

(III) For all $n \in \mathbb{N} \cup \{0\}$ we have $b_n - a_n = \dfrac{1}{2^n}(b-1)$.

We use Mathematical Induction to prove (I). Since $E \subseteq [a,b] = \big[a_0, b_0\big]$, and E is infinite, (I) is true for $n = 0$. Let $n \in \mathbb{N}$ be arbitrary and assume that $E \cap \big[a_{n-1}, b_{n-1}\big]$ is infinite. We consider two cases. CASE 1. The set

$$E \cap \left[a_{n-1}, \frac{a_{n-1}+b_{n-1}}{2}\right] \ \text{ is infinite.}$$

CASE 2. The set

$$E \cap \left[a_{n-1}, \frac{a_{n-1}+b_{n-1}}{2}\right] \ \text{ is finite.}$$

In CASE 1, by the recursive definition of $\big[a_n, b_n\big]$, the set $E \cap \big[a_n, b_n\big]$ is infinite. In CASE 2, by the recursive definition of $\big[a_n, b_n\big]$ we have

$$E \cap \big[a_{n-1}, b_{n-1}\big] = \big(E \cap \big[a_n, b_n\big]\big) \bigcup \left(E \cap \left[a_{n-1}, \frac{a_{n-1}+b_{n-1}}{2}\right]\right).$$

Since in this case the second set in the union on the right-hand side of the preceding equality is finite and the entire union is infinite, the set $E \cap \big[a_n, b_n\big]$ is infinite. This completes the proof of (I).

Property (II) follows from the recursive definition of the sequence of intervals.

We use Mathematical Induction to prove (III). Since,

$$b_0 - a_0 = b - a = \frac{1}{2^0}(b-a),$$

(III) holds for $n = 0$. Let $n \in \mathbb{N}$ be arbitrary and assume that

$$b_{n-1} - a_{n-1} = \frac{1}{2^{n-1}}(b-a).$$

By the recursive definition of $[a_n, b_n]$ we have, either

$$b_n - a_n = \frac{a_{n-1}+b_{n-1}}{2} - a_{n-1} = \frac{b_{n-1}-a_{n-1}}{2} = \frac{1}{2^n}(b-a)$$

or

$$b_n - a_n = b_{n-1} - \frac{a_{n-1} + b_{n-1}}{2} = \frac{b_{n-1} - a_{n-1}}{2} = \frac{1}{2^n}(b-a).$$

In either case, (III) is proved.

By property (II), the Nested Intervals Theorem, Theorem 6.3.1, applies to the sequence of intervals $[a_n, b_n]$, with $n \in \mathbb{N} \cup \{0\}$. Consequently, there exists $c \in \mathbb{R}$ such that for all $n \in \mathbb{N} \cup \{0\}$ we have $c \in [a_n, b_n]$. In particular $c \in [a, b]$.

Before proceeding, recall that by Exercise 5.2.11, with $a = 1/2$, we have

$$\forall n \in \mathbb{N} \qquad \frac{1}{2^n} \leq \frac{1}{n+1}. \tag{7.1.3}$$

(The statement (7.1.3) can be proved directly by Mathematical Induction.)

Next, we will prove that c is an accumulation point of E. Let $\epsilon > 0$ be arbitrary. Set

$$m = \left\lceil \frac{b-a}{\epsilon} \right\rceil.$$

Since $(b-a)/\epsilon > 0$, we have $m \in \mathbb{N}$. Next we prove that for this specific $m \in \mathbb{N}$ we have $b_m - a_m < \epsilon$. We start by using the property (III) proven above:

$$\begin{aligned}
b_m - a_m &= \frac{b-a}{2^m} \\
\boxed{\text{by (7.1.3)}} &\leq \frac{b-a}{m+1} \\
\boxed{\text{by pizza-party}} &< \frac{b-a}{m} \\
\boxed{\text{by the definition } m} &= \frac{b-a}{\left\lceil \frac{b-a}{\epsilon} \right\rceil} \\
\boxed{\text{by pizza-party}} &\leq \frac{b-a}{\frac{b-a}{\epsilon}} \\
\boxed{\text{Remark ??}} &= (b-a)\frac{1}{(b-a)\frac{1}{\epsilon}} \\
\boxed{\text{Exercise 4.1.4(x)}} &= \left((b-a)\frac{1}{(b-a)}\right)\frac{1}{\frac{1}{\epsilon}} \\
\boxed{\text{Axioms MO, MR, Exercise 4.1.4(ix)}} &= \epsilon.
\end{aligned}$$

In the preceding eight displayed lines, I overexplained the following sequence of inequalities:

$$b_m - a_m = \frac{b-a}{2^m} \leq \frac{b-a}{m+1} < \frac{b-a}{\left\lceil \frac{b-a}{\epsilon} \right\rceil} \leq \frac{b-a}{\frac{b-a}{\epsilon}} = \epsilon.$$

Since $c \in [a_m, b_m]$, the just-proven inequality $b_m - a_m < \epsilon$ yields

$$c - \epsilon < c - (b_m - a_m) \leq a_m \leq c \leq b_m \leq b_m + c - a_m < c + \epsilon.$$

(Please overexplain these inequalities to yourself.)

Hence, $[a_m, b_m] \subset (c - \epsilon, c + \epsilon) = N(c, \epsilon)$. By property (I) proven above, the set $E \cap [a_m, b_m]$ is infinite. Therefore, the set

$$E \cap \big(N(c, \epsilon) \setminus \{c\}\big) \supset E \cap [a_m, b_m]$$

is infinite as well and thus nonempty. Since $\epsilon > 0$ was arbitrary, this proves that c is an accumulation point of E. Since $c \in [a_0, b_0] = [a, b]$ and $E \subseteq [a, b]$ was an arbitrary infinite set, we have proven that the bounded closed interval $[a, b]$ is a compact set. □

Definition 7.1.22. A family $\mathcal{G}$ of open sets is an *open cover* for a set E if

$$E \subseteq \bigcup \{G : G \in \mathcal{G}\}.$$

Definition 7.1.23. If every open cover of a set E has a finite subfamily that is also an open cover of E, than we say that E has the *Heine-Borel* property.

Theorem 7.1.24. *Let $a, b \in \mathbb{R}$ be such that $a < b$. The closed interval $[a, b]$ has the Heine-Borel property.*

Proof. Let $\mathcal{G}$ be an arbitrary open cover of $[a, b]$. That is, $\mathcal{G}$ is a family of open subsets of $\mathbb{R}$ such that

$$[a, b] \subseteq \bigcup \{G : G \in \mathcal{G}\}. \tag{7.1.4}$$

Consider the set

$$A = \left\{ x \in (a, b] \, : \, \exists \, n \in \mathbb{N} \ \wedge \ \exists \, G_1, \dots, G_n \in \mathcal{G} \text{ such that } [a, x] \subseteq \bigcup_{j=1}^{n} G_j \right\}.$$

The set A is nonempty. Since $A \subseteq [a, b]$, it is bounded above. By the Completeness Axiom, A has a supremum. Set $\alpha = \sup A$. Since b is an upper bound for A, we have $\alpha \leq b$. Thus

$$\alpha \in (a, b] \subseteq \bigcup \{G : G \in \mathcal{G}\}.$$

Therefore, there exists $G_0 \in \mathcal{G}$ such that $\alpha \in G_0$. Since G_0 is an open set, there exists an $\epsilon > 0$ such that

$$(\alpha - \epsilon, \alpha + \epsilon) \subseteq G_0.$$

Since $\alpha = \sup A$, (ii) in Definition 6.4.2 implies that there exists $x \in A$ such that $\alpha - \epsilon < x$. Since $x \in A$, there exist $G_1, \dots, G_n \in \mathcal{G}$ such that

$$[a, x] \subseteq \bigcup_{j=1}^{n} G_j.$$

Since

$$x \in \in (\alpha - \epsilon, \alpha + \epsilon) \subseteq G_0,$$

we have that

$$[x, \alpha] \subseteq G_0.$$

Therefore

$$[a, \alpha] = [a, x] \cup [x, \alpha] \subseteq \left(\bigcup_{j=1}^{n} G_j \right) \bigcup G_0.$$

By the definition of A, it follows that $\alpha \in A$. Thus, $\alpha = \max A$.

Next we prove the implication:

$$x \in A \ \wedge \ x < b \quad \Rightarrow \quad \exists y \in A \ \text{ such that } \ x < y.$$

That is we prove that the only possible number to be the maximum of A is b. □

7.1.1. The structure of open sets in $\mathbb{R}$.

DEFINITION 7.1.25. A subset $I \subseteq \mathbb{R}$ is an open interval if one of the following four conditions is satisfied

- ➢ $I = \mathbb{R}$.
- ➢ There exists $a \in \mathbb{R}$ such that $I = \mathbb{R}_{<a}$.
- ➢ There exists $b \in \mathbb{R}$ such that $I = \mathbb{R}_{>b}$.
- ➢ There exist $a, b \in \mathbb{R}$ such that $a < b$ and $I = (a, b)$.

EXERCISE 7.1.26. Let $\mathcal{I}$ be an infinite family of open mutually disjoint intervals. (Mutually disjoint means that if $I_1, I_2 \in \mathcal{I}$ and $I_1 \neq I_2$, then $I_1 \cap I_2 \neq \emptyset$.) Prove that $\mathcal{I}$ is countable. HINT: Construct a surjection $f : \mathbb{Q} \to \mathcal{I}$. ◁

EXERCISE 7.1.27. Let G be a nonempty open subset of $\mathbb{R}$. Assume that $\mathbb{R} \setminus G$ is neither bounded above nor below.

(i) Prove that for each $x \in G$ there exist $a, b \in \mathbb{R} \setminus G$ such that $a < b$, $x \in (a, b)$ and $(a, b) \subseteq G$. HINT: Consider the sets

$$\bigl\{\alpha : \alpha < x,\ (\alpha, x] \subset G\bigr\} \quad \text{and} \quad \bigl\{\beta : x < \beta,\ [x, \beta) \subset G\bigr\}.$$

Prove that both sets are nonempty and bounded, so a and b could be their infimum and supremum, respectively.

(ii) Let $a, b, c, d \in \mathbb{R} \setminus G$ be such that $a < b$, $c < d$ and such that the intervals (a, b), (c, d) are subsets of G. Then either $(a, b) = (c, d)$ or $(a, b) \cap (c, d) = \emptyset$. HINT: Assume $x \in (a, b) \cap (c, d)$. Then $a < x$ and $c < x$. Notice $c < a$, implies $a \in G$. Since $a \notin G$, $a \leq c$. Similarly $c \leq a$. Thus $a = c$. Similarly $b = d$.

(iii) Prove that there exists a finite or countable family of finite open mutually disjoint intervals whose union equals G. HINT: Use previous two parts and Exercise 7.1.26. ◁

EXERCISE 7.1.28. Let G be a nonempty open subset of $\mathbb{R}$. Then there exists a finite or countable family of open mutually disjoint intervals whose union equals G. HINT: Assume $\mathbb{R} \setminus G \neq \emptyset$. Prove: if $\mathbb{R} \setminus G$ bounded below, then there exists $a \in \mathbb{R} \setminus G$ such that $\mathbb{R}_{<a} \subseteq G$. Prove: if $\mathbb{R} \setminus G$ bounded above, then there exists $b \in \mathbb{R} \setminus G$ such that $\mathbb{R}_{>b} \subseteq G$. Then consider all possible cases for $\mathbb{R} \setminus G$. ◁

7.2. The topology of $\mathbb{R}^2$

Recall that $\mathbb{R}^2$ is the set of all ordered pairs (x, y) with $x, y \in \mathbb{R}$. The essential notion in Definition 7.1.1 is that of a neighborhood of a point. This notion is based on the idea of *distance* between real numbers x and y which is expressed simply as the absolute value of the difference $|x - y|$. Since we have a well established concept of distance in $\mathbb{R}^2$, the notions introduced in Section 7.1 extend to $\mathbb{R}^2$. Recall that the *Euclidean distance* between points $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$ in $\mathbb{R}^2$ is given as

$$d(\mathbf{a}, \mathbf{b}) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}.$$

DEFINITION 7.2.1. A *rectangle* is a subset of $\mathbb{R}^2$ which is a cross product of two intervals.

DEFINITION 7.2.2. A set $E \subset \mathbb{R}^2$ is *bounded* if there exists $a, b, c, d \in \mathbb{R}$ such that $a < b$ and $c < d$, and such that $E \subseteq [a, b] \times [c, d]$.

DEFINITION 7.2.3 (Topological notions in $\mathbb{R}^2$). All points in this definition are elements of $\mathbb{R}^2$ and all sets are subsets of $\mathbb{R}^2$.

(i) Let $\epsilon > 0$. A *neighborhood* (or ϵ*-neighborhood*) of a point $\mathbf{a}$ is the set
$$B(\mathbf{a}, \epsilon) = \left\{\mathbf{x} \in \mathbb{R}^2 : d(\mathbf{a}, \mathbf{x}) < \epsilon\right\}.$$
Commonly $B(\mathbf{a}, \epsilon)$ is called an *open ball*. The number ϵ is called the *radius* of $B(\mathbf{a}, \epsilon)$.

(ii) A point $\mathbf{a}$ is an *accumulation point* of a set E if every neighborhood of $\mathbf{a}$ contains a point $\mathbf{x} \neq \mathbf{a}$ such that $\mathbf{x} \in E$. That is, $\mathbf{a}$ is an *accumulation point* of the set E if
$$E \bigcap \big(B(\mathbf{a}, \epsilon) \setminus \{\mathbf{a}\}\big) \neq \emptyset \quad \text{for all} \quad \epsilon > 0.$$

(iii) The set of all accumulation points of a set E is denoted by E'.

(iv) A set E is *closed* if it contains all its accumulation points. That is, E is closed if the inclusion $E' \subset E$ holds.

(v) The *closure* of a set E is defined as the union of $E \cup E'$. It is denoted by $\overline{E}$.

(vi) A point $\mathbf{a}$ is an *interior point* of the set E if there is a neighborhood of $\mathbf{a}$ that is a subset of E. That is, $\mathbf{a}$ is an *interior point* of E if there exists $\epsilon > 0$ such that $B(\mathbf{a}, \epsilon) \subseteq E$.

(vii) The set of all the interior points of a set E is called the *interior* of E. It is denoted by E°.

(viii) A set E is *open* if every point of E is an interior point of E. That is, E is open if $E = E^\circ$.

(ix) The *boundary* of a set E is the set $\overline{E} \setminus E^\circ$. It is denoted by ∂E.

(x) A set E is *compact* if every infinite subset of E has an accumulation point in E.

(xi) Let $E \subseteq F$. A set E is *dense* in F if every neighborhood of every point in F contains a point of E.

EXERCISE 7.2.4. Prove that a subset of $\mathbb{R}^2$ is bounded if and only if it is contained in an open ball. ◁

EXERCISE 7.2.5. Consider the rectangle $R = [0, 1] \times [-1, 1]$ and define the set $E = R \cup \{(-1, 0)\}$. Find E', $\overline{E}$, E° and ∂E. ◁

EXERCISE 7.2.6. Prove the triangle inequality for the Euclidean distance: For every $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^2$ we have
$$d(\mathbf{a}, \mathbf{c}) \leq d(\mathbf{a}, \mathbf{b}) + d(\mathbf{b}, \mathbf{c}).$$
◁

Exercise 7.2.7. [Hausdorff Property] Let $\mathbf{a}$ and $\mathbf{b}$ be arbitrary distinct point in $\mathbb{R}^2$. Prove that there exist a neighborhood of $\mathbf{a}$ and a neighborhood of $\mathbf{b}$ which are disjoint sets. $\lhd$

Exercise 7.2.8. Let $E \subset \mathbb{R}^2$. Prove that E is open if and only if $\mathbb{R}^2 \setminus E$ is closed. $\lhd$

Exercise 7.2.9. Prove that the intersection and the union of two open sets is open. $\lhd$

Exercise 7.2.10. Does the previous exercise hold true for a countable family of open sets? $\lhd$

Exercise 7.2.11. Prove that the intersection and the union of two closed sets is closed. $\lhd$

Exercise 7.2.12. Does the previous exercise hold true for a countable family of closed sets? $\lhd$

Exercise 7.2.13. Let $a, b, c, d \in \mathbb{R}$ be such that $a < b$ and $c < d$. Prove that the rectangle $[a, b] \times [c, d]$ is a compact set. $\lhd$

CHAPTER 8

Sequences in $\mathbb{R}$

8.1. Definitions and examples

DEFINITION 8.1.1. A *sequence in* $\mathbb{R}$ is a function whose domain is $\mathbb{N}$ and whose range is in $\mathbb{R}$.

Let $s : \mathbb{N} \to \mathbb{R}$ be a sequence in $\mathbb{R}$. Then the values of s are

$$s(1),\ s(2),\ s(3),\ \dots,\ s(n),\ \dots\ .$$

It is customary to write s_n instead of $s(n)$ for the values of a sequence. Sometimes a sequence will be specified by listing its first few terms

$$s_1,\ s_2,\ s_3,\ s_4,\ \dots\ ,$$

and sometimes by listing of all its terms $\{s_n\}_{n=1}^{\infty}$ or $\{s_n\}$. One way of specifying a sequence is to give a formula, or a recursion formula for its $n-$th term s_n.

REMARK 8.1.2. In the above notation s is the "name" of the sequence and $n \in \mathbb{N}$ is the independent variable. $\triangleleft$

REMARK 8.1.3. Notice the difference between the following two expressions:

$\{s_n\}_{n=1}^{\infty}$ This expression denotes a function (sequence).

$\{s_n : n \in \mathbb{N}\}$ This expression denotes a set: The range of a sequence $\{s_n\}_{n=1}^{\infty}$.

For example $\{1-(-1)^n\}_{n=1}^{\infty}$ stands for the function $n \mapsto 1-(-1)^n$, $n \in \mathbb{N}$, while

$$\{1-(-1)^n : n \in \mathbb{N}\} = \{0, 2\}.$$ $\triangleleft$

EXAMPLE 8.1.4. Here we give examples of sequences given by a formula. In each formula below $n \in \mathbb{N}$.

(a) n, (b) n^2, (c) $\sqrt{n}$, (d) $(-1)^n$,

(e) $\dfrac{1}{n}$, (f) $\dfrac{1}{n^2}$, (g) $\dfrac{1}{\sqrt{n}}$, (h) $1-\dfrac{(-1)^n}{n}$,

(i) $\dfrac{1}{n!}$, (j) $2^{1/n}$, (k) $n^{1/n}$, (l) $n^{(-1)^n}$,

(m) $\dfrac{9^n}{n!}$, (n) $\dfrac{(-1)^{n+1}}{2n-1}$, (o) $\dfrac{n^{(-1)^n}}{n+1}$, (p) $\left(\dfrac{e}{n}\right)^n \dfrac{n!}{\sqrt{n}}$.

$\triangleleft$

EXAMPLE 8.1.5. Few more sequences given by a formula are

$$\text{(a)}\quad \left\{\sqrt{n^2+1}-n\right\}_{n=1}^{\infty}, \quad \text{(b)}\quad \left\{\sqrt{n^2+n}-n\right\}_{n=1}^{\infty}, \quad \text{(c)}\quad \left\{\sqrt{n+1}-\sqrt{n}\right\}_{n=1}^{\infty}.$$

◁

EXAMPLE 8.1.6. In this example we give several recursively defined sequences.

(a) $s_1 = 1 \quad \text{and} \quad \bigl(\forall n \in \mathbb{N}\bigr) \quad s_{n+1} = -\dfrac{s_n}{2}$,

(b) $x_1 = 1 \quad \text{and} \quad \bigl(\forall n \in \mathbb{N}\bigr) \quad x_{n+1} = 1 + \dfrac{x_n}{4}$,

(c) $x_1 = 2 \quad \text{and} \quad \bigl(\forall n \in \mathbb{N}\bigr) \quad x_{n+1} = \dfrac{x_n}{2} + \dfrac{1}{x_n}$,

(d) $a_1 = \sqrt{2} \quad \text{and} \quad \bigl(\forall n \in \mathbb{N}\bigr) \quad a_{n+1} = \sqrt{2 + a_n}$,

(e) $s_1 = 1 \quad \text{and} \quad \bigl(\forall n \in \mathbb{N}\bigr) \quad s_{n+1} = \sqrt{1 + s_n}$,

(f) $x_1 = 0 \quad \text{and} \quad \bigl(\forall n \in \mathbb{N}\bigr) \quad x_{n+1} = \dfrac{9 + x_n}{10}$.

For a recursively defined sequence it is useful to evaluate the values of the first few terms to get an idea how sequence behaves. ◁

EXAMPLE 8.1.7. The most important examples of sequences are listed below:

$$b_n = a, \ n \in \mathbb{N}, \quad \text{where} \quad a \in \mathbb{R}, \tag{8.1.1}$$

$$p_n = a^n, \ n \in \mathbb{N}, \quad \text{where} \quad -1 < a < 1, \tag{8.1.2}$$

$$E_n = \left(1 + \frac{1}{n}\right)^n, \quad n \in \mathbb{N}, \tag{8.1.3}$$

$$G_1 = a + ax \quad \text{and} \quad \bigl(\forall\, n \in \mathbb{N}\bigr)\ G_{n+1} = G_n + ax^{n+1}, \ \text{where} \ -1 < x < 1, \tag{8.1.4}$$

$$S_1 = 2 \quad \text{and} \quad \bigl(\forall\, n \in \mathbb{N}\bigr) \quad S_{n+1} = S_n + \frac{1}{(n+1)!}, \tag{8.1.5}$$

$$v_1 = 1 + a \quad \text{and} \quad \bigl(\forall\, n \in \mathbb{N}\bigr) \quad v_{n+1} = v_n + \frac{a^{n+1}}{(n+1)!}, \quad \text{where} \quad a \in \mathbb{R}. \tag{8.1.6}$$

◁

DEFINITION 8.1.8. Let $\{a_n\}$ be a sequence in $\mathbb{R}$. A sequence which is recursively defined by

$$S_1 = a_1 \quad \text{and} \quad \bigl(\forall\, n \in \mathbb{N}\bigr) \quad S_{n+1} = S_n + a_{n+1}, \tag{8.1.7}$$

is called a *sequence of partial sum* corresponding to $\{a_n\}$.

EXAMPLE 8.1.9. The sequences of partial sums associated with the sequences in Example 8.1.4 (e), (f) and (n) are important examples for Definition 8.1.8. Notice also that the sequences in (8.1.4), (8.1.5) and (8.1.6) are sequences of partial sums. All of these are very important. ◁

8.2. Bounded sequences

DEFINITION 8.2.1. Let $\{s_n\}$ be a sequence in $\mathbb{R}$.

(1) If a real number M satisfies

$$s_n \leq M \quad \text{for all} \quad n \in \mathbb{N}$$

then M is called an *upper bound* of $\{s_n\}$ and the sequence $\{s_n\}$ is said to be *bounded above*.

(2) If a real number m satisfies

$$m \leq s_n \quad \text{for all} \quad n \in \mathbb{N},$$

then m is called a *lower bound* of $\{s_n\}$ and the sequence $\{s_n\}$ is said to be *bounded below*.

(3) The sequence $\{s_n\}$ is said to be *bounded* if it is bounded above and bounded below.

REMARK 8.2.2. Clearly, a sequence $\{s_n\}$ is bounded above if and only if the set $\{s_n : n \in \mathbb{N}\}$ is bounded above. Similarly, a sequence $\{s_n\}$ is bounded below if and only if the set $\{s_n : n \in \mathbb{N}\}$ is bounded below. ◁

REMARK 8.2.3. The sequence $\{s_n\}$ is bounded if and only if there exists a real number $K > 0$ such that $|s_n| \leq K$ for all $n \in \mathbb{N}$. ◁

EXERCISE 8.2.4. There is a huge task here. For each sequence given in this section it is of interest to determine whether it is bounded or not. As usual, some of the proofs are easy, some are hard. It is important to do few easy proofs and observe their structure. This will provide the setting to appreciate proofs for hard examples. ◁

8.3. The definition of a convergent sequence

DEFINITION 8.3.1. A sequence $\{s_n\}$ is a *constant* sequence if there exists $L \in \mathbb{R}$ such that $s_n = L$ for all $n \in \mathbb{N}$.

EXERCISE 8.3.2. Prove that the sequence $s_n = \left\lfloor \dfrac{3n-1}{2n} \right\rfloor$, $n \in \mathbb{N}$, is a constant sequence. ◁

DEFINITION 8.3.3. A sequence $\{s_n\}$ is *eventually constant* if there exists $L \in \mathbb{R}$ and $N_0 \in \mathbb{R}$ such that $s_n = L$ for all $n \in \mathbb{N}$ such that $n > N_0$.

EXERCISE 8.3.4. Prove that the sequence $s_n = \left\lceil \dfrac{3n-2}{2n+3} \right\rceil$, $n \in \mathbb{N}$, is eventually constant. ◁

EXERCISE 8.3.5. Prove that the sequence $s_n = \left\lfloor \dfrac{5n-(-1)^n}{n/2+5} \right\rfloor$, $n \in \mathbb{N}$, is eventually constant. ◁

DEFINITION 8.3.6. A sequence $\{s_n\}$ *converges* if there exists $L \in \mathbb{R}$ such that for each $\epsilon > 0$ there exists a real number $N(\epsilon)$ such that

$$n \in \mathbb{N}, \quad n > N(\epsilon) \quad \Rightarrow \quad |s_n - L| < \epsilon.$$

The number L is called the *limit* of the sequence $\{s_n\}$. We also say that $\{s_n\}$ *converges to* L and write

$$\lim_{n\to\infty} s_n = L \quad \text{or} \quad s_n \to L \ \ (n \to \infty).$$

If a sequence does not converge we say that it *diverges*.

REMARK 8.3.7. The definition of convergence is a complicated statement. Formally it can be written as:

$$\exists\, L \in \mathbb{R} \ \text{ s.t. } \ \forall\, \epsilon > 0 \ \ \exists\, N(\epsilon) \in \mathbb{R} \ \text{ s.t. } \ \forall\, n \in \mathbb{N}, \quad n > N(\epsilon) \ \Rightarrow \ |s_n - L| < \epsilon.$$

◁

EXERCISE 8.3.8. State the negation of the statement in Remark 8.3.7. ◁

8.3.1. My informal discussion of convergence. It is easy to agree that the constant sequences are simplest possible sequences. For example the sequence

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
c_n	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

or formally, $c_n = 1$ for all $n \in \mathbb{N}$, is a very simple sequence. No action here! In this case, clearly, $\lim_{n\to\infty} c_n = 1$.

Now, I define $s_n = \dfrac{n - (-1)^n}{n}$, $n \in \mathbb{N}$, and I ask: Is $\{s_n\}$ a constant sequence? Just looking at the first few terms

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
s_n	2	$\frac{1}{2}$	$\frac{4}{3}$	$\frac{3}{4}$	$\frac{6}{5}$	$\frac{5}{6}$	$\frac{8}{7}$	$\frac{7}{8}$	$\frac{10}{9}$	$\frac{9}{10}$	$\frac{12}{11}$	$\frac{11}{12}$	$\frac{14}{13}$	$\frac{13}{14}$	$\frac{16}{15}$	$\frac{15}{16}$	$\frac{18}{17}$

indicates that this sequence is not constant. The table above also indicates that the sequence $\{s_n\}$ is not eventually constant. But imagine that you have a calculator which is capable of displaying only one decimal place. On this calculator the first terms of this sequence would look like:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
s_n	2.0	0.5	1.3	0.8	1.2	0.8	1.1	0.9	1.1	0.9	1.1	0.9	1.1	0.9	1.1

and the next 15 terms would look like:

n	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
s_n	0.9	1.1	0.9	1.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Basically, after the 20-th term this calculator does not distinguish s_n from 1. That is, this calculator leads us to think that $\{s_n\}$ is eventually constant. Why is this?

On this calculator all numbers between $0.95 = 1 - 1/20$ and $1.05 = 1 + 1/20$ are represented as 1, and for our sequence we can prove that

$$n \in \mathbb{N}, \quad n > 20 \qquad \Rightarrow \qquad 1 - \frac{1}{20} < s_n < 1 + \frac{1}{20},$$

or, equivalently,

$$n \in \mathbb{N}, \quad n > 20 \qquad \Rightarrow \qquad \big|s_n - 1\big| < \frac{1}{20}.$$

In the notation of Definition 8.3.6 this means $N(1/20) = 20$.

One can reasonably object that the above calculator is not very powerful and propose to use a calculator that can display three decimal places. Then the terms of $\{s_n\}$ starting with $n = 21$ are

n	21	22	23	24	25	26	27	28	29	30
s_n	1.048	0.955	1.043	0.958	1.040	0.962	1.037	0.964	1.034	0.967

Now the question is: Can we fool this powerful calculator to think that $\{s_n\}$ is eventually constant? Notice that on this calculator all numbers between $0.9995 = 1 - 1/2000$ and $1.0005 = 1 + 1/2000$ are represented as 1. Therefore, in the notation of Definition 8.3.6, we need $N(1/2000)$ such that

$$n \in \mathbb{N}, \quad n > N(1/2000) \qquad \Rightarrow \qquad 1 - \frac{1}{2000} < s_n < 1 + \frac{1}{2000}.$$

A calculation shows that $N(1/2000) = 2000$. That is

$$n \in \mathbb{N}, \quad n > 2000 \qquad \Rightarrow \qquad \big|s_n - 1\big| < \frac{1}{2000}.$$

This is illustrated by the following table of approximate values:

n	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
s_n	0.999	1.001	0.999	1.001	1.000	1.000	1.000	1.000	1.000	1.000

Hence, even this more powerful calculator is fooled into thinking that $\{s_n\}$ is eventually constant.

In computer science the precision of a computer is measured by the number called the *machine epsilon* (also called *macheps, machine precision* or *unit roundoff*). It is the smallest number that gives a number greater than 1 when added to 1.

Now, Definition 8.3.6 can be paraphrased as: A sequence converges if on each computer it appears to be eventually constant. This is the reason why I think that instead of the phrase "a sequence is convergent" we could use the phrase "a sequence is constantish."

8.4. Finding $N(\epsilon)$ for a convergent sequence

Example 8.4.1. Prove that $\displaystyle\lim_{n\to\infty} \frac{2n-1}{n+3} = 2.$ ◁

Solution. We prove the given equality using Definition 8.3.6. To do that for each $\epsilon > 0$ we have to find $N(\epsilon)$ such that

$$n \in \mathbb{N}, \quad n > N(\epsilon) \qquad \Rightarrow \qquad \left|\frac{2n-1}{n+3} - 2\right| < \epsilon. \tag{8.4.1}$$

Let $\epsilon > 0$ be given. We can think of n as an unknown in $\left|\dfrac{2n-1}{n+3} - 2\right| < \epsilon$ and solve this inequality for n. To this end first simplify the left-hand side:

$$\left|\frac{2n-1}{n+3} - 2\right| = \left|\frac{2n-1-2n-6}{n+3}\right| = \frac{|-7|}{|n+3|} = \frac{7}{n+3}. \tag{8.4.2}$$

Now, $\dfrac{7}{n+3} < \epsilon$ is much easier to solve for $n \in \mathbb{N}$:

$$\frac{7}{n+3} < \epsilon \quad \Leftrightarrow \quad \frac{n+3}{7} > \frac{1}{\epsilon} \quad \Leftrightarrow \quad n+3 > \frac{7}{\epsilon} \quad \Leftrightarrow \quad n > \frac{7}{\epsilon} - 3. \tag{8.4.3}$$

Now (8.4.3) indicates that we can choose $N(\epsilon) = \frac{7}{\epsilon} - 3$.

Now we have $N(\epsilon)$, but to complete the formal proof, we have to prove implication (8.4.1). The proof follows. Let $n \in \mathbb{N}$ and $n > \frac{7}{\epsilon} - 3$. Then the equivalences in (8.4.3) imply that $\frac{7}{n+3} < \epsilon$. Since by (8.4.3), $\left|\frac{2n-1}{n+3} - 2\right| = \frac{7}{n+3}$, it follows that $\left|\frac{2n-1}{n+3} - 2\right| < \epsilon$. This completes the proof of implication (8.4.1). □

Remark 8.4.2. This remark is essential for the understanding of the process described in the following examples. In the solution of Example 8.4.1 we found (in some sense) the smallest possible $N(\epsilon)$. It is important to notice that implication (8.4.1) holds with any larger value for "$N(\epsilon)$." For example, implication (8.4.1) holds if we set $N(\epsilon) = \frac{7}{\epsilon}$. With this new $N(\epsilon)$ we can prove implication (8.4.1) as follows. Let $n \in \mathbb{N}$ and $n > \frac{7}{\epsilon}$. Then $\frac{7}{n} < \epsilon$. Since clearly $\frac{7}{n+3} < \frac{7}{n}$, the last two inequalities imply that $\frac{7}{n+3} < \epsilon$ and we can continue with the same proof as in the solution of Example 8.4.1. ◁

Example 8.4.3. Prove that $\displaystyle\lim_{n\to\infty} \frac{1}{n^3 - n + 1} = 0$. ◁

Solution. We prove the given equality using Definition 8.3.6. To do that for each $\epsilon > 0$ we have to find $N(\epsilon)$ such that

$$n \in \mathbb{N}, \quad n > N(\epsilon) \qquad \Rightarrow \qquad \left|\frac{1}{n^3 - n + 1} - 0\right| < \epsilon. \tag{8.4.4}$$

Let $\epsilon > 0$ be given. We can think of n as an unknown in $\left|\dfrac{1}{n^3 - n + 1} - 0\right| < \epsilon$ and solve this inequality for n. To this end first simplify the left-hand side:

$$\left|\frac{1}{n^3 - n + 1} - 0\right| = \left|\frac{1}{n^3 - n + 1}\right| = \frac{|1|}{|n^3 - n + 1|} = \frac{1}{n^3 - n + 1}. \tag{8.4.5}$$

Unfortunately $\dfrac{1}{n^3 - n + 1} < \epsilon$ is not easy to solve for $n \in \mathbb{N}$. Therefore we use the idea from Remark 8.4.2 and replace the quantity $\dfrac{1}{n^3 - n + 1}$ with a larger quantity. To make a fraction larger we have to make the denominator smaller. Notice that $n^2 - n = n(n-1) \geq n - 1$ for all $n \in \mathbb{N}$. Therefore for all $n \in \mathbb{N}$ we have

$$n^3 - n + 1 = n^3 - (n-1) \geq n^3 - n(n-1) = n(n^2 - n + 1) \geq n.$$

Consequently,

$$\frac{1}{n^3 - n + 1} \leq \frac{1}{n}. \tag{8.4.6}$$

Now, $\dfrac{1}{n} < \epsilon$ is truly easy to solve for $n \in \mathbb{N}$:

$$\frac{1}{n} < \epsilon \qquad \Leftrightarrow \qquad n > \frac{1}{\epsilon}. \tag{8.4.7}$$

Hence we set $N(\epsilon) = \frac{1}{\epsilon}$.

Now we have $N(\epsilon)$, but to complete the formal proof, we have to prove implication (8.4.4). The proof follows. Let $n \in \mathbb{N}$ and $n > \frac{1}{\epsilon}$. Then the equivalence in (8.4.7) implies that $\frac{1}{n} < \epsilon$. By (8.4.6), $\frac{1}{n^3-n+1} \leq \frac{1}{n}$. The last two inequalities yield that $\frac{1}{n^3-n+1} < \epsilon$. By (8.4.5) it follows that $\left|\frac{1}{n^3-n+1} - 0\right| < \epsilon$. This completes the proof of implication (8.4.4). □

Example 8.4.4. Prove that $\displaystyle\lim_{n\to\infty} \frac{n^2-1}{n^2-2n+2} = 1.$ ◁

Solution. We prove the given equality using Definition 8.3.6. To do that for each $\epsilon > 0$ we have to find $N(\epsilon)$ such that

$$n \in \mathbb{N}, \quad n > N(\epsilon) \qquad \Rightarrow \qquad \left|\frac{n^2-1}{n^2-2n+2} - 1\right| < \epsilon. \tag{8.4.8}$$

Let $\epsilon > 0$ be given. We can think of n as an unknown in $\left|\dfrac{n^2-1}{n^2-2n+2} - 1\right| < \epsilon$ and solve this inequality for n. To this end first simplify the left-hand side:

$$\left|\frac{n^2-1}{n^2-2n+2} - 1\right| = \left|\frac{n^2-1-n^2+2n-2}{n^2-2n+2}\right| = \frac{|2n-3|}{n^2-2n+2}. \tag{8.4.9}$$

Unfortunately $\dfrac{|2n-3|}{n^2-2n+2} < \epsilon$ is not easy to solve for $n \in \mathbb{N}$. Therefore we use the idea from Remark 8.4.2 and replace the quantity $\dfrac{|2n-3|}{n^2-2n+2}$ with a larger quantity. Here is one way to discover a desired inequality. We first notice that for all $n \in \mathbb{N}$ the following two inequalities hold

$$|2n-3| \leq 2n \tag{8.4.10}$$

and

$$n^2 - 2n + 2 = \frac{n^2}{2} + \frac{1}{2}\bigl(n^2 - 4n + 4\bigr) = \frac{n^2}{2} + \frac{1}{2}(n-2)^2 \geq \frac{n^2}{2}. \tag{8.4.11}$$

Consequently

$$\frac{|2n-3|}{n^2-2n+2} \leq \frac{2n}{n^2/2} = \frac{4}{n}. \tag{8.4.12}$$

Now, $\dfrac{4}{n} < \epsilon$ is truly easy to solve for $n \in \mathbb{N}$:

$$\frac{4}{n} < \epsilon \qquad \Leftrightarrow \qquad n > \frac{4}{\epsilon}. \tag{8.4.13}$$

Hence we set $N(\epsilon) = \frac{4}{\epsilon}$.

Finally we have $N(\epsilon)$. But to complete the formal proof we have to prove implication (8.4.8). The proof follows. Let $n \in \mathbb{N}$ and $n > \frac{4}{\epsilon}$. Then the equivalence in (8.4.13) implies $\frac{4}{n} < \epsilon$. By (8.4.12), $\frac{|2n-3|}{n^2-2n+2} \leq \frac{4}{n}$. The last two inequalities yield $\frac{|2n-3|}{n^2-2n+2} < \epsilon$. By (8.4.9) it follows that $\left|\frac{n^2-1}{n^2-2n+2} - 1\right| < \epsilon$. This completes the proof of implication (8.4.8). $\square$

REMARK 8.4.5. For most sequences $\{s_n\}$ a proof of $\lim_{n\to\infty} s_n = L$ based on Definition 8.3.6 should consist from the following steps.

(1) Use algebra to simplify the expression $|s_n - L|$. It is desirable to eliminate the absolute value.
(2) Discover an inequality of the form

$$|s_n - L| \leq b(n) \quad \text{valid for all} \quad n \in \mathbb{N}. \tag{8.4.14}$$

Here $b(n)$ should be a simple function with the following properties:
   (a) $b(n) > 0$ for all $n \in \mathbb{N}$.
   (b) $\lim_{n\to\infty} b(n) = 0$. (Just check this property "mentally.")
   (c) $b(n) < \epsilon$ is easily solvable for n for <u>every</u> $\epsilon > 0$. The solution should be of the form "$n >$ some expression involving ϵ, call it $N(\epsilon)$."
(3) Use inequality (8.4.14) to prove the implication $n \in \mathbb{N}$, $n > N(\epsilon) \Rightarrow |s_n - L| < \epsilon$. $\lhd$

EXERCISE 8.4.6. Determine the limits (if they exist) of the sequences (e), (f), (g), (h), (i), and (n) in Example 8.1.4. Prove your claims. $\lhd$

EXERCISE 8.4.7. Determine whether the sequence $\left\{\dfrac{3n+1}{7n-4}\right\}_{n=1}^{\infty}$ converges and, if it converges, give its limit. Provide a formal proof. $\lhd$

EXERCISE 8.4.8. Determine the limits (if they exist) of the sequences in Example 8.1.5. Prove your claims. $\lhd$

8.5. Two standard sequences

EXERCISE 8.5.1. Let $a \in \mathbb{R}$ be such that $-1 < a < 1$.

(a) Prove that for all $n \in \mathbb{N}$ we have

$$|a|^n \leq \frac{1}{n\,(1-|a|)}.$$

(b) Prove that

$$\lim_{n\to\infty} a^n = 0. \qquad \lhd$$

EXERCISE 8.5.2. Let a be a positive real number. Prove that

$$\lim_{n\to\infty} a^{1/n} = 1. \qquad \lhd$$

SOLUTION. Let $a > 0$. If $a = 1$, then $a^{1/n} = 1$ for all $n \in \mathbb{N}$. Therefore $\lim_{n\to\infty} a^{1/n} = 1$.

Assume $a > 1$. Then $a^{1/n} > 1$. We shall prove that

$$a^{1/n} - 1 \leq a\,\frac{1}{n} \quad \big(\forall\, n \in \mathbb{N}\big). \tag{8.5.1}$$

Put $x = a^{1/n} - 1 > 0$. Then, by Bernoulli's inequality we get

$$a = (1+x)^n \geq 1 + nx.$$

Consequently, solving for x we get that $x = a^{1/n} - 1 \leq (a-1)/n$. Since $a - 1 < a$, (8.5.1) follows.

Assume $0 < a < 1$. Then $1/a > 1$. Therefore, by already proved (8.5.1), we have

$$\left(\frac{1}{a}\right)^{1/n} - 1 \leq \frac{1}{a}\frac{1}{n} \quad \bigl(\forall\, n \in \mathbb{N}\bigr).$$

Since $(1/a)^{1/n} = 1/\bigl(a^{1/n}\bigr)$, simplifying the last inequality, together with the inequality $a^{1/n} < 1$, yields

$$1 - a^{1/n} \leq \frac{a^{1/n}}{a}\frac{1}{n} \leq \frac{1}{a}\frac{1}{n} \quad \bigl(\forall\, n \in \mathbb{N}\bigr). \tag{8.5.2}$$

As $a < a + 1/a$ and $1/a < a + 1/a$, the inequalities (8.5.1) and (8.5.2) imply

$$\bigl|a^{1/n} - 1\bigr| \leq \left(a + \frac{1}{a}\right)\frac{1}{n} \quad \bigl(\forall\, n \in \mathbb{N}\bigr). \tag{8.5.3}$$

Let $\epsilon > 0$ be given. Solving $\bigl(a + 1/a\bigr)\frac{1}{n} < \epsilon$ for n, reveals $N(\epsilon)$:

$$N(\epsilon) = \left(a + \frac{1}{a}\right)\frac{1}{\epsilon}$$

Now it is easy to prove the implication (Do it as an exercise!)

$$n \in \mathbb{N}, \quad n > \left(a + \frac{1}{a}\right)\frac{1}{\epsilon} \quad \Rightarrow \quad |a^{1/n} - 1| < \epsilon. \qquad \square$$

8.6. Non-convergent sequences

Exercise 8.6.1. Prove that the sequence (d) in Example 8.1.4 does not converge. Use Remark 8.3.7 and Exercise 8.3.8 ◁

Exercise 8.6.2. (Prove or Disprove) If $\{s_n\}$ does not converge to L, then there exist $\epsilon > 0$ and $N(\epsilon)$ such that $|s_n - L| \geq \epsilon$ for all $n \geq N(\epsilon)$. ◁

8.7. Convergence and boundedness

Exercise 8.7.1. Consider the following two statements:

(A) The sequence $\{s_n\}$ is bounded.
(B) The sequence $\{s_n\}$ converges.

Is (A)⇒(B) true or false? Is (B)⇒(A) true or false? Justify your answers. ◁

8.8. Algebra of limits of convergent sequences

Exercise 8.8.1. Let $\{s_n\}$ be a sequence in $\mathbb{R}$ and let $L \in \mathbb{R}$. Set $t_n = s_n - L$ for all $n \in \mathbb{N}$.

Prove that $\{s_n\}$ converges to L if and only if $\{t_n\}$ converges to 0. ◁

Exercise 8.8.2. Let $c \in \mathbb{R}$. If $\lim\limits_{n\to\infty} x_n = X$ and $z_n = c\,x_n$ for all $n \in \mathbb{N}$, then $\lim\limits_{n\to\infty} z_n = c\,X$. ◁

Exercise 8.8.3. Let $\{x_n\}$ and $\{y_n\}$ be sequences in $\mathbb{R}$. Assume

(a) $\{x_n\}$ converges to 0,
(b) $\{y_n\}$ is bounded,
(c) $z_n = x_n y_n$ for all $n \in \mathbb{N}$.

Prove that $\{z_n\}$ converges to 0. ◁

EXERCISE 8.8.4. Let $\{x_n\}$ and $\{y_n\}$ be sequences in $\mathbb{R}$. Assume
(a) $\lim\limits_{n\to\infty} x_n = X$,
(b) $\lim\limits_{n\to\infty} y_n = Y$,
(c) $z_n = x_n + y_n$ for all $n \in \mathbb{N}$.

Prove that $\lim\limits_{n\to\infty} z_n = X + Y$. ◁

EXERCISE 8.8.5. Let $\{x_n\}$ and $\{y_n\}$ be sequences in $\mathbb{R}$. Assume
(a) $\lim\limits_{n\to\infty} x_n = X$,
(b) $\lim\limits_{n\to\infty} y_n = Y$,
(c) $z_n = x_n\, y_n$ for all $n \in \mathbb{N}$.

Prove that $\lim\limits_{n\to\infty} z_n = X\,Y$. ◁

EXERCISE 8.8.6. If $\lim\limits_{n\to\infty} x_n = X$ and $X > 0$, then there exists a real number N such that $n \geq N$ implies $x_n \geq X/2$. ◁

EXERCISE 8.8.7. Let $\{x_n\}$ be a sequence in $\mathbb{R}$. Assume
(a) $x_n \neq 0$ for all $n \in \mathbb{N}$,
(b) $\lim\limits_{n\to\infty} x_n = X$,
(c) $X > 0$,
(d) $w_n = \dfrac{1}{x_n}$ for all $n \in \mathbb{N}$.

Prove that $\lim\limits_{n\to\infty} w_n = \dfrac{1}{X}$. ◁

EXERCISE 8.8.8. Let $\{x_n\}$ and $\{y_n\}$ be sequences in $\mathbb{R}$. Assume
(a) $x_n \neq 0$ for all $n \in \mathbb{N}$,
(b) $\lim\limits_{n\to\infty} x_n = X$,
(c) $\lim\limits_{n\to\infty} y_n = Y$,
(d) $X \neq 0$,
(e) $z_n = \dfrac{y_n}{x_n}$ for all $n \in \mathbb{N}$.

Prove that $\lim\limits_{n\to\infty} z_n = \dfrac{Y}{X}$. (Hint: Use previous exercises.) ◁

EXERCISE 8.8.9. Prove that $\lim\limits_{n\to\infty} \dfrac{2n^2 + n - 5}{n^2 + 2n + 2} =$ (insert correct value) by using the results we have proved (Exercises 8.8.2, 8.8.4, 8.8.5, 8.8.7, 8.8.8) and a small trick. You may use Definition 8.3.6 of convergence directly in this problem only to evaluate limit of the special form $\lim\limits_{n\to\infty} \dfrac{1}{n}$. ◁

REMARK 8.8.10. The point of Exercise 8.8.9 is to see that the general properties of limits (Exercises 8.8.2, 8.8.4, 8.8.5, 8.8.7, 8.8.8) can be used to reduce complicated situations to a few simple ones, so that when the few simple ones have been done it is no longer necessary to go back to Definition 8.3.6 of convergence every time. ◁

8.9. Convergent sequences and the order in $\mathbb{R}$

EXERCISE 8.9.1. Let $\{s_n\}$ be a sequence in $\mathbb{R}$. Assume

(a) $\lim_{n\to\infty} s_n = L$.
(b) There exists a real number N_0 such that $s_n \geq 0$ for all $n \in \mathbb{N}$ such that $n > N_0$.

Prove that $L \geq 0$. $\triangleleft$

EXERCISE 8.9.2. Let $\{a_n\}$ and $\{b_n\}$ be sequences in $\mathbb{R}$. Assume

(a) $\lim_{n\to\infty} a_n = K$.
(b) $\lim_{n\to\infty} b_n = L$.
(c) There exists a real number N_0 such that $a_n \leq b_n$ for all $n \in \mathbb{N}$ such that $n > N_0$.

Prove that $K \leq L$. $\triangleleft$

EXERCISE 8.9.3. Is the following refinement of Exercise 8.9.1 true? If $\{s_n\}$ converges to L and if $s_n > 0$ for all $n \in \mathbb{N}$, then $L > 0$. $\triangleleft$

EXERCISE 8.9.4. Let $\{x_n\}$ be a sequence in $\mathbb{R}$. Assume

(a) $x_n \geq 0$ for all $n \in \mathbb{N}$,
(b) $\lim_{n\to\infty} x_n = X$,
(c) $w_n = \sqrt{x_n}$ for all $n \in \mathbb{N}$.

Prove that $\lim_{n\to\infty} w_n = \sqrt{X}$. $\triangleleft$

EXERCISE 8.9.5. There are three sequences in this exercise: $\{a_n\}$, $\{b_n\}$ and $\{s_n\}$. Assume the following

(a) The sequence $\{a_n\}$ converges to L.
(b) The sequence $\{b_n\}$ converges to L.
(c) There exists a real number n_0 such that

$$a_n \leq s_n \leq b_n \quad \text{for all} \quad n \in \mathbb{N}, \ n > n_0.$$

Prove that $\{s_n\}$ converges to L. $\triangleleft$

EXERCISE 8.9.6. (1) Let $x \geq 0$ and $n \in \mathbb{N}$. Prove the inequality

$$(1+x)^n \geq 1 + n\,x + \frac{n(n-1)}{2}\,x^2.$$

(2) Prove that for all $n \in \mathbb{N}$ we have $1 \leq n^{1/n} \leq 1 + \dfrac{2}{\sqrt{n}}$.
HINT: Apply the inequality proved in (1) to $\bigl(1 + 2/\sqrt{n}\bigr)^n$.
(3) Prove that the sequence $\{n^{1/n}\}$ converges and determine its limit. $\triangleleft$

EXERCISE 8.9.7. (1) Prove that $(n!)^2 \geq n^n$ for all $n \in \mathbb{N}$. HINT: Write

$$\bigl(n!\bigr)^2 = \bigl(1\cdot n\bigr)\bigl(2\cdot(n-1)\bigr)\cdots\bigl((n-1)\cdot 2\bigr)\bigl(n\cdot 1\bigr) = \prod_{k=1}^{n} k\bigl(n-k+1\bigr).$$

Then prove $k\bigl(n-k+1\bigr) \geq n$ for all $k = 1, \ldots, n$.
(2) Prove that

$$\lim_{n\to\infty} \frac{1}{\bigl(n!\bigr)^{1/n}} = 0.$$ $\triangleleft$

8.10. The monotonic convergence theorem

DEFINITION 8.10.1. A sequence $\{s_n\}$ of real numbers is said to be *non-decreasing* if $s_n \leq s_{n+1}$ for all $n \in \mathbb{N}$, *strictly increasing* if $s_n < s_{n+1}$ for all $n \in \mathbb{N}$, *non-increasing* if $s_n \geq s_{n+1}$ for all $n \in \mathbb{N}$, *strictly decreasing* if $s_n > s_{n+1}$ for all $n \in \mathbb{N}$. A sequence with any of these properties is said to be *monotonic*.

EXERCISE 8.10.2. Again a huge task here. Which of the sequences in Examples 8.1.4, 8.1.5, and 8.1.6 are monotonic? Find few monotonic ones in each example. Provide rigorous proofs. ◁

EXERCISE 8.10.3. (Prove or Disprove) If $\{x_n\}$ is non-increasing, then $\{x_n\}$ converges. ◁

The following two exercises give powerful tools for establishing convergence of a sequence.

EXERCISE 8.10.4. If $\{s_n\}$ is non-increasing and bounded below, then $\{s_n\}$ converges. ◁

EXERCISE 8.10.5. If $\{s_n\}$ is non-decreasing and bounded above, then $\{s_n\}$ converges. ◁

PROOF. Assume that the sequence $\{s_n\}$ is non-decreasing and bounded above. Consider the range of the sequence $\{s_n\}$. That is consider the set

$$A = \{s_n \ : \ n \in \mathbb{N}\}.$$

The set A is nonempty and bounded above. Therefore $\sup A$ exists. Put $L = \sup A$.

We will prove that $s_n \to L$ $(n \to \infty)$. Let $\epsilon > 0$ be arbitrary. Since $L = \sup A$ we have

(1) $L \geq s_n$ for all $n \in \mathbb{N}$.
(2) There exists $a_\epsilon \in A$ such that $L - \epsilon < a_\epsilon$.

Since $a_\epsilon \in A$, there exists $N_\epsilon \in \mathbb{N}$ such that $a_\epsilon = s_{N_\epsilon}$. It remains to prove that

$$n \in \mathbb{N}, \quad n > N_\epsilon \quad \Rightarrow \quad |s_n - L| < \epsilon. \tag{8.10.1}$$

Let $n \in \mathbb{N}$, $n > N_\epsilon$ be arbitrary. Since we assume that $\{s_n\}$ is non-decreasing, it follows that $s_n \geq s_{N_\epsilon}$. Since $L - \epsilon < a_\epsilon = s_{N_\epsilon} \leq s_n$, we conclude that $L - s_n < \epsilon$. Since $L \geq s_n$, we have $|s_n - L| = L - s_n < \epsilon$. The implication (8.10.1) is proved. □

EXERCISE 8.10.6. There is a huge task here. Consider the sequences given in Example 8.1.6. Prove that each of these sequences converges and determine its limit. ◁

8.11. Two important sequences with the same limit

In this section we study the sequences defined in (8.1.3) and (8.1.5).

$$E_n = \left(1 + \frac{1}{n}\right)^n, \quad n \in \mathbb{N},$$

$$S_1 = 2 \quad \text{and} \quad \bigl(\forall\, n \in \mathbb{N}\bigr) \quad S_{n+1} = S_n + \frac{1}{(n+1)!}.$$

EXERCISE 8.11.1. Prove by mathematical induction that $S_n \leq 3 - 1/n$ for all $n \in \mathbb{N}$. ◁

EXERCISE 8.11.2. Prove that the sequence $\bigl\{S_n\bigr\}$ converges. ◁

EXERCISE 8.11.3. Let $n, k \in \mathbb{N}$ and $n \geq k$. Use Bernoulli's inequality to prove that

$$\frac{n!}{(n-k)!\, n^k} \geq 1 - \frac{(k-1)k}{n}$$

HINT: Notice that

$$\frac{n!}{n^k (n-k)!} = 1 \cdot \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \geq \left(1 - \frac{k-1}{n}\right)^k.$$ ◁

EXERCISE 8.11.4. The following inequalities hold: $E_1 = S_1$ and for all integers n greater than 1,

$$S_n - \frac{3}{n} < E_n < S_n. \tag{8.11.1}$$

HINT: Let n be an integer greater than 2. Notice that by the Binomial Theorem

$$E_n = \left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{1}{n^k} = 1 + 1 + \sum_{k=2}^{n} \frac{n!}{(n-k)!\, n^k} \frac{1}{k!}.$$

Then use Exercise 8.11.3 to prove $E_n > S_n - S_{n-2}/n$. Then use Exercise 8.11.1. ◁

EXERCISE 8.11.5. The sequences $\{E_n\}$ and $\{S_n\}$ converge to the same limit. ◁

Exercise 8.11.5 justifies the following definition.

DEFINITION 8.11.6. The number e is the common limit of the sequences $\{E_n\}$ and $\{S_n\}$.

REMARK 8.11.7. The sequence $\{E_n\}$ is increasing. To prove this claim let $n \in \mathbb{N}$ be arbitrary. Consider the fraction

$$\begin{aligned} \frac{E_{n+1}}{E_n} &= \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^n} = \frac{n+1}{n} \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^{n+1}} = \frac{n+1}{n} \left(\frac{\frac{n+2}{n+1}}{\frac{n+1}{n}}\right)^{n+1} \\ &= \frac{n+1}{n} \left(\frac{n(n+2)}{(n+1)^2}\right)^{n+1} = \frac{n+1}{n} \left(1 - \frac{1}{(n+1)^2}\right)^{n+1} \end{aligned} \tag{8.11.2}$$

Since $-\frac{1}{(n+1)^2} > -1$ for all $n \in \mathbb{N}$, applying Bernoulli's Inequality with $x = -\frac{1}{(n+1)^2}$ we get

$$\left(1 - \frac{1}{(n+1)^2}\right)^{n+1} > 1 - (n+1)\frac{1}{(n+1)^2} = 1 - \frac{1}{n+1}. \tag{8.11.3}$$

The relations (8.11.2) and (8.11.3) imply

$$\frac{E_{n+1}}{E_n} = \frac{n+1}{n} \left(1 - \frac{1}{(n+1)^2}\right)^{n+1} > \frac{n+1}{n} \left(1 - \frac{1}{n+1}\right) = 1.$$

Thus

$$\frac{E_{n+1}}{E_n} > 1 \quad \text{for all} \quad n \in \mathbb{N},$$

that is the sequence $\{E_n\}$ is increasing. ◁

8.12. Subsequences

Composing functions is a common way how functions interact with each other. Can we compose two sequences? Let $x : \mathbb{N} \to \mathbb{R}$ and $y : \mathbb{N} \to \mathbb{R}$ be two sequences. Does the composition $x \circ y$ make sense? This composition makes sense only if the range of y is contained in $\mathbb{N}$. In this case $y : \mathbb{N} \to \mathbb{N}$. That is the composition $x \circ y$ makes sense only if y is a sequence in $\mathbb{N}$. It turns out that the most important composition of sequences involve increasing sequences in $\mathbb{N}$. In this section, the Greek letters μ and ν will always denote increasing sequences of natural numbers.

DEFINITION 8.12.1. A *subsequence* of a sequence $\{x_n\}$ is a composition of the sequence $\{x_n\}$ and an increasing sequence $\{\mu_k\}$ of natural numbers. This composition will be denoted by $\{x_{\mu_k}\}$ or $\{x(\mu_k)\}$.

REMARK 8.12.2. The concept of subsequence consists of two ingredients:

- the sequence $\{x_n\}$ (remember it's really a function: $x : \mathbb{N} \to \mathbb{R}$)
- the increasing sequence $\{\mu_k\}$ of natural numbers (remember this is an increasing function: $\mu : \mathbb{N} \to \mathbb{N}$).

The composition $x \circ \mu$ of these two sequences is a new sequence $y : \mathbb{N} \to \mathbb{R}$. The k-th term y_k of this sequence is $y_k = x_{\mu_k}$. Note the analogy with the usual notation for functions: $y(k) = x(\mu(k))$. Usually we will not introduce the new name for a subsequence: we will write $\{x_{\mu_k}\}_{k=1}^{\infty}$ to denote a subsequence of the sequence $\{x_n\}$. Here $\{\mu_k\}_{k=1}^{\infty}$ is and increasing sequence of natural numbers which selects particular elements of the sequence $\{x_n\}$ to be included in the subsequence. ◁

REMARK 8.12.3. Roughly speaking, a subsequence of $\{x_n\}$ is a sequence formed by selecting some of the terms in $\{x_n\}$, keeping them in the same order as in the original sequence. It is the sequence $\{\mu_k\}$ of positive integers that does the selecting. ◁

EXAMPLE 8.12.4. Few examples of increasing sequences in $\mathbb{N}$ are:

(a) $\mu_k = 2k, \quad k \in \mathbb{N}$. (The sequence of even positive integers.)
(b) $\nu_k = 2k - 1, \quad k \in \mathbb{N}$. (The sequence of odd positive integers.)
(c) $\mu_k = k^2, \quad k \in \mathbb{N}$. (The sequence of perfect squares.)
(d) Let j be a fixed positive integer. Set $\nu_k = j + k$ for all $k \in \mathbb{N}$.
(e) The sequence $2, 3, 5, 7, 11, 13, 17, \ldots$ of prime numbers. For this sequence no formula for $\{\mu_k\}$ is known.

◁

EXERCISE 8.12.5. Let $\{\mu_n\}$ be an increasing sequence in $\mathbb{N}$. Prove that $\mu_n \geq n$ for all $n \in \mathbb{N}$. ◁

EXERCISE 8.12.6. Each subsequence of a convergent sequence is convergent with the same limit. ◁

Remark 8.12.7. The "contrapositive" of Exercise 8.12.6 is a powerful tool for proving that a given sequence does not converge. As an illustration prove that the sequence $\{(-1)^n\}$ does not converge in two different ways: using the definition of convergence and using the "contrapositive" of Exercise 8.12.6. $\lhd$

Exercise 8.12.8. [The Zipper Theorem] Let $\{x_n\}$ be a sequence in $\mathbb{R}$ and let $\{\mu_k\}$ and $\{\nu_k\}$ be increasing sequences in $\mathbb{N}$. Assume

(a) $\{\mu_k : k \in \mathbb{N}\} \bigcup \{\nu_k : k \in \mathbb{N}\} = \mathbb{N}$.
(b) $\{x_{\mu_k}\}$ converges to L.
(c) $\{x_{\nu_k}\}$ converges to L.

Prove that $\{x_n\}$ converges to L. $\lhd$

Example 8.12.9. The sequence (o) in Example 8.1.4 does not converge, but it does have convergent subsequences, for instance the subsequence $\left\{\dfrac{2k}{2k+1}\right\}_{k=1}^{\infty}$ (Here $\mu_k = 2k,\ k \in \mathbb{N}$) and the subsequence $\left\{\dfrac{1}{(2k-1)2k}\right\}_{k=1}^{\infty}$ (Here $\nu_k = 2k - 1,\ k \in \mathbb{N}$). $\lhd$

Remark 8.12.10. The notation for subsequences is a little tricky at first. Note that in x_{μ_k} it is k that is the variable. Thus the successive elements of the subsequence are $x_{\mu_1}, x_{\mu_2}, x_{\mu_3}$, etc. To indicate a different subsequence of the same sequence $\{x_n\}_{n=1}^{\infty}$ it would be necessary to change not the variable name, but the selection sequence. For example $\{x_{\mu_k}\}_{k=1}^{\infty}$ and $\{x_{\nu_k}\}_{k=1}^{\infty}$ in Example 8.12.9 are distinct subsequences of $\{x_n\}$. (Thus $\{x_{\mu_k}\}_{k=1}^{\infty}$ and $\{x_{\mu_j}\}_{j=1}^{\infty}$ are the same subsequence of $\{x_n\}_{n=1}^{\infty}$ for exactly the same reason that $x \mapsto x^2\,(x \in \mathbb{R})$ and $t \mapsto t^2\,(t \in \mathbb{R})$ are the same function. To make a different function it's the rule you must change, not the variable name.) $\lhd$

Example 8.12.11. Let $\{x_n\}$ be the sequence defined by

$$x_n = \frac{(-1)^n (n+1)^{(-1)^n}}{n}, \qquad n \in \mathbb{N}.$$

The values of $\{x_n\}$ are

$$-\frac{1}{1\cdot 2}, \frac{3}{2}, -\frac{1}{3\cdot 4}, \frac{5}{4}, -\frac{1}{5\cdot 6}, \frac{7}{6}, -\frac{1}{7\cdot 8}, \frac{9}{8}, -\frac{1}{9\cdot 10}, \frac{11}{10}, \ldots.$$

$\lhd$

Exercise 8.12.12. Every sequence has a monotonic subsequence. $\lhd$

Hint: Let $\{x_n\}$ be an arbitrary sequence. Consider the set

$$\mathbb{M} = \{n \in \mathbb{N} : \forall\, k > n \text{ we have } x_k \geq x_n\}.$$

The set $\mathbb{M}$ is either finite or infinite. Construct a monotonic subsequence in each case.

Exercise 8.12.13. Every bounded sequence of real numbers has a convergent subsequence. $\lhd$

Exercise 8.12.14. [Sequential characterization of compactness] Let F be a subset of $\mathbb{R}$. Prove that F is compact if and only if every sequence in K has a subsequence which converges in K. $\lhd$

8.13. The Cauchy criterion

DEFINITION 8.13.1. A sequence $\{s_n\}$ of real numbers is called a *Cauchy sequence* if for every $\epsilon > 0$ there exists a real number N_ϵ such that

$$\forall\, n, m \in \mathbb{N}, \quad n, m > N_\epsilon \quad \Rightarrow \quad |s_n - s_m| < \epsilon.$$

EXERCISE 8.13.2. Prove that every convergent sequence is a Cauchy sequence. ◁

EXERCISE 8.13.3. Prove that every Cauchy sequence is bounded. ◁

EXERCISE 8.13.4. If a Cauchy sequence has a convergent subsequence, then it converges. ◁

EXERCISE 8.13.5. Prove that each Cauchy sequence has a convergent subsequence. ◁

EXERCISE 8.13.6. Prove that a sequence converges if and only if it is a Cauchy sequence. ◁

8.14. Sequences and supremum and infimum

EXERCISE 8.14.1. Let $A \subset \mathbb{R}$, $A \neq \emptyset$ and assume that A is bounded above. Prove that $a = \sup A$ if and only if

(a) a is an upper bound of A, that is, $a \geq x$, for all $x \in A$;
(b) there exists a sequence $\{x_n\}$ such that

$$x_n \in A \quad \text{for each} \quad n \in \mathbb{N} \quad \text{and} \quad \lim_{n\to\infty} x_n = a. \qquad ◁$$

EXERCISE 8.14.2. Let $A \subset \mathbb{R}$, $A \neq \emptyset$ and assume that A is bounded above. Let $a = \sup A$ and assume that $a \notin A$. Prove that there exists a strictly increasing sequence $\{x_n\}$ such that

$$x_n \in A \quad \text{for each} \quad n \in \mathbb{N} \quad \text{and} \quad \lim_{n\to\infty} x_n = a. \qquad ◁$$

EXERCISE 8.14.3. State and prove the characterization of infimum which is analogous to the characterization of $\sup A$ given in Exercise 8.14.1. ◁

EXERCISE 8.14.4. State and prove an exercise involving infimum of a set which is analogous to Exercise 8.14.2. ◁

8.15. Limit inferior and limit superior

Consider the sequence $\{(-1)^n n/(n+1)\}$. This sequence does not have a limit. However, one can prove that the subsequence of even terms converges to 1 and the subsequence of odd terms converges to -1. The concepts of the limit inferior and the limit superior capture this phenomenon for every bounded sequence.

DEFINITION 8.15.1. Let $\{x_n\}$ be a bounded sequence. For $m \in \mathbb{N}$ let

$$y_m = \inf\{x_n \,:\, n \geq m\} \qquad \text{and} \qquad z_m = \sup\{x_n \,:\, n \geq m\}.$$

The *limit inferior* of $\{x_n\}$, denoted by $\varliminf_{n\to\infty} x_n$ or $\liminf_{n\to\infty} x_n$, is given by

$$\varliminf_{n\to\infty} x_n = \sup\bigl\{y_m \,:\, m \in \mathbb{N}\bigr\}.$$

The *limit superior* of $\{x_n\}$, denoted by $\varlimsup_{n\to\infty} x_n$ or $\limsup_{n\to\infty} x_n$, is given by

$$\varlimsup_{n\to\infty} x_n = \inf\bigl\{z_m \,:\, m \in \mathbb{N}\bigr\}.$$

Exercise 8.15.2. Compute

$$\varliminf_{n\to\infty} \frac{(-1)^n n}{n+1} \quad \text{and} \quad \varlimsup_{n\to\infty} \frac{(-1)^n n}{n+1}.$$

Rigorously justify your reasoning. ◁

Exercise 8.15.3. Let $\{x_n\}$ be a bounded sequence. Prove that the sequences $\{y_m\}$ and $\{z_m\}$ from Definition 8.15.1 converge and state their limits. ◁

Exercise 8.15.4. Let $\{x_n\}$ be a bounded sequence. Explore the relationship between these four numbers

$$\varliminf_{n\to\infty} x_n, \quad \varlimsup_{n\to\infty} x_n, \quad \varliminf_{n\to\infty} (-x_n), \quad \varlimsup_{n\to\infty} (-x_n).$$ ◁

Exercise 8.15.5. Let $\{a_n\}$ and $\{b_n\}$ be bounded sequences. Prove that

$$\varlimsup_{n\to\infty} (a_n + b_n) \le \varlimsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$$

and

$$\varliminf_{n\to\infty} (a_n + b_n) \ge \varliminf_{n\to\infty} a_n + \varliminf_{n\to\infty} b_n.$$

Find an example for which the inequalities above are strict. ◁

Exercise 8.15.6. Let $\{x_n\}$ be a bounded sequence. Let

$$L_1 = \varliminf_{n\to\infty} x_n \quad \text{and} \quad L_2 = \varlimsup_{n\to\infty} x_n.$$

Prove that for every $\epsilon > 0$ there exists $N(\epsilon)$ such that

$$n \in \mathbb{N}, \quad n > N(\epsilon) \quad \Rightarrow \quad x_n \in (L_1 - \epsilon, L_2 + \epsilon).$$ ◁

Exercise 8.15.7. Let $\{x_n\}$ be a bounded sequence. Prove that $\{x_n\}$ converges if and only if

$$\varliminf_{n\to\infty} x_n = \varlimsup_{n\to\infty} x_n.$$ ◁

CHAPTER 9

Continuous functions

In this chapter, the letter I will always denote a nonempty subset of $\mathbb{R}$. This includes more general sets, but the most common examples of I are intervals.

9.1. The ϵ-δ definition of a continuous function

DEFINITION 9.1.1. A function $f : I \to \mathbb{R}$ is *continuous at a point* $x_0 \in I$ if for every $\epsilon > 0$ there exists $\delta = \delta(\epsilon, x_0) > 0$ such that for all $x \in I$ we have

$$|x - x_0| < \delta(\epsilon, x_0) \quad \Rightarrow \quad \bigl|f(x) - f(x_0)\bigr| < \epsilon. \tag{9.1.1}$$

The function f is *continuous on* I if it is continuous at each point of I.

Since most of the time we will be dealing with functions which are continuous on their domains we will state the definition of a function continuous on its domain as a separate definition.

DEFINITION 9.1.2. A function $f : I \to \mathbb{R}$ with domain I and codomain $\mathbb{R}$ is *continuous on* I if for every $x_0 \in I$ and every $\epsilon > 0$ there exists $\delta = \delta(\epsilon, x_0) > 0$ such that for all $x \in I$ we have

$$|x - x_0| < \delta(\epsilon, x_0) \quad \Rightarrow \quad \bigl|f(x) - f(x_0)\bigr| < \epsilon.$$

Note that the implication in (9.1.1) can be restated as

$$x \in I \cap (x_0 - \delta, x_0 + \delta) \quad \Rightarrow \quad f(x) \in \bigl(f(x_0) - \epsilon, f(x_0) + \epsilon\bigr),$$

where we abbreviated $\delta = \delta(\epsilon, x_0) > 0$. Further, the preceding implication can be restated using the notation of neighborhoods introduced in Section 7.1, as follows:

$$x \in I \cap N(x_0, \delta) \quad \Rightarrow \quad f(x) \in N\bigl(f(x_0), \epsilon\bigr).$$

?????????????

To further simplify the

Next we restate Definition 9.1.1 using the terminology introduced in Section 7.1. For a function $f : I \to \mathbb{R}$ and a subset $A \subseteq I$ we will use the notation $f(A)$ to denote the set

$$\bigl\{y \in \mathbb{R} : \exists\, x \in A \text{ such that } f(x) = y\bigr\} = \bigl\{f(x) : x \in A\bigr\}$$

A function $f : I \to \mathbb{R}$ is continuous at a point $x_0 \in I$ if for each neighborhood V of $f(x_0)$ there exists a neighborhood U of x_0 such that

$$f\bigl(I \cap U\bigr) \subseteq V.$$

9.2. Finding $\delta(\epsilon)$ for a given function at a given point

In this and the next section we will prove that some familiar functions are continuous. This should be a review of what was done in Math 226.

A general strategy for proving that a given function f is continuous at a given point x_0 is as follows:

Step 1: Simplify the expression $|f(x)-f(x_0)|$ and try to establish a simple connection with the expression $|x-x_0|$. The simplest connection is to discover positive constants δ_0 and K such that

$$x \in I \ \text{ and } \ x_0-\delta_0 < x < x_0+\delta_0 \quad \Rightarrow \quad |f(x)-f(x_0)| \leq K\,|x-x_0|. \qquad (9.2.1)$$

Constants δ_0 and K might depend on x_0. Formulate your discovery as a lemma.

Step 2: Let $\epsilon > 0$ be given. Use the result in Step 1 to define your $\delta(\epsilon, x_0)$. For example, if (9.2.1) holds, then $\delta(\epsilon, x_0) = \min\bigl\{\epsilon/K, \delta_0\bigr\}$.

Step 3: Use the definition of $\delta(\epsilon, x_0)$ from Step 2 and the lemma from Step 1 to prove the implication (9.1.1).

Example 9.2.1. We will show that the function $f(x) = x^2$ is continuous at $x_0 = 3$. Here $I = \mathbb{R}$ and we do not need to worry about the domain of f.
Step 1. First simplify

$$\bigl|f(x)-f(x_0)\bigr| = \bigl|x^2-3^2\bigr| = \bigl|(x+3)(x-3)\bigr| = \bigl|x+3\bigr|\,\bigl|x-3\bigr|. \qquad (9.2.2)$$

Now we notice that if $2 < x < 4$ we have $\bigl|x+3\bigr| = x+3 \leq 7$. Thus (9.2.1) holds with $\delta_0 = 1$ and $K = 7$. We formulate this result as a lemma.

Lemma. *Let $f(x) = x^2$ and $x_0 = 3$. Then*

$$|x-3| < 1 \qquad \Rightarrow \qquad \bigl|x^2-3^2\bigr| < 7|x-3|.$$

Proof. Let $|x-3| < 1$. Then $2 < x < 4$. Therefore $x+3 > 0$ and $|x+3| = x+3 < 7$. By (9.2.2) we now have $\bigl|x^2-3^2\bigr| < 7|x-3|$. ◁

Step 2. Now we define $\delta(\epsilon) = \min\bigl\{\epsilon/7, 1\bigr\}$.
Step 3. It remains to prove (9.1.1). To this end, assume $|x-3| < \min\bigl\{\epsilon/7, 1\bigr\}$. Then $|x-3| < 1$. Therefore, by Lemma we have $\bigl|x^2-3^2\bigr| < 7|x-3|$. Since by the assumption $|x-3| < \epsilon/7$, we have $7|x-3| < 7\,\epsilon/7\epsilon$. Now the inequalities

$$\bigl|x^2-3^2\bigr| < 7|x-3| \quad \text{and} \quad 7|x-3| < \epsilon$$

imply that $\bigl|x^2-3^2\bigr| < \epsilon$. This proves (9.1.1) and completes the proof that the function $f(x) = x^2$ is continuous at $x_0 = 3$. ◁

Exercise 9.2.2. Prove that the reciprocal function $x \mapsto \dfrac{1}{x}$, $x \neq 0$, is continuous at $x_0 = 1/2$. ◁

Exercise 9.2.3. State carefully what it means for a function f *not* to be continuous at a point x_0 in its domain. (Express this as a formal mathematical statement.) ◁

Exercise 9.2.4. Consider the function $f(x) = \operatorname{sgn} x$. Find a point x_0 at which the function f is not continuous. Provide a formal proof. ◁

Exercise 9.2.5. Show that the function $f(x) = x^2$ is continuous on $\mathbb{R}$. ◁

Exercise 9.2.6. Prove that $q(x) = 3x^2 + 5$ is continuous on $\mathbb{R}$. ◁

9.3. Familiar continuous functions

EXERCISE 9.3.1. Let $m, k \in \mathbb{R}$ and $m \neq 0$. Prove that the linear function $\ell(x) = m\,x + k$ is continuous on $\mathbb{R}$. ◁

EXERCISE 9.3.2. Let $a, b, c \in \mathbb{R}$ and $a \neq 0$. Prove that the quadratic function $q(x) = a\,x^2 + b\,x + c$ is continuous on $\mathbb{R}$. ◁

EXERCISE 9.3.3. Let $n \in \mathbb{N}$ and let $x, x_0 \in \mathbb{R}$ be such that $x_0 - 1 \leq x \leq x_0 + 1$. Prove the following inequality

$$\bigl|x^n - x_0^n\bigr| \leq n\bigl(|x_0| + 1\bigr)^{n-1}\bigl|x - x_0\bigr|.$$

HINT: First notice that the assumption $x_0 - 1 \leq x \leq x_0 + 1$ implies that $|x| < |x_0| + 1$. Then use the Mathematical Induction and the identity

$$\bigl|x^{n+1} - x_0^{n+1}\bigr| = \bigl|x^{n+1} - x\,x_0^n + x\,x_0^n - x_0^{n+1}\bigr|.$$

◁

EXERCISE 9.3.4. Let $n \in \mathbb{N}$. Prove that the power function $x \mapsto x^n$, $x \in \mathbb{R}$, is continuous on $\mathbb{R}$. ◁

EXERCISE 9.3.5. Let $n \in \mathbb{N}$ and let $a_0, a_1, \ldots, a_n \in \mathbb{R}$ with $a_n \neq 0$. Prove that the n-th order polynomial

$$p(x) = a_0 + a_1\,x + \cdots + a_{n-1}\,x^{n-1} + a_n\,x^n$$

is a continuous function on $\mathbb{R}$. ◁

EXERCISE 9.3.6. Prove that the reciprocal function $x \mapsto \dfrac{1}{x}$, $x \neq 0$, is continuous on its domain. ◁

EXERCISE 9.3.7. Prove that the square root function $x \mapsto \sqrt{x}$, $x \geq 0$, is continuous on its domain. ◁

EXERCISE 9.3.8. Let $n \in \mathbb{N}$ and let x and a be positive real numbers. Prove that

$$\bigl|\sqrt[n]{x} - \sqrt[n]{a}\bigr| \leq \frac{\sqrt[n]{a}}{a}\,\bigl|x - a\bigr|.$$

HINT: Notice that the given inequality is equivalent to

$$b^{n-1}\,\bigl|y - b\bigr| \leq \bigl|y^n - b^n\bigr|, \quad y, b > 0.$$

This inequality can be proved using Exercise 5.2.7 (with $a = 1$ and $x = y/b$). ◁

EXERCISE 9.3.9. Let $n \in \mathbb{N}$. Prove that the n-th root function $x \mapsto \sqrt[n]{x}$, $x \geq 0$, is continuous on its domain. ◁

9.4. Various properties of continuous functions

EXERCISE 9.4.1. Let $f : I \to \mathbb{R}$ be continuous at $x_0 \in I$ and let y be a real number such that $f(x_0) < y$. Then there exists $\alpha > 0$ such that

$$x \in I \cap (x_0 - \alpha, x_0 + \alpha) \quad \Rightarrow \quad f(x) < y.$$

Illustrate with a diagram. ◁

EXERCISE 9.4.2. Let $f : I \to \mathbb{R}$ be a continuous function on I. Let S be a non-empty bounded above subset of I such that $u = \sup S$ belongs to I. Let $y \in \mathbb{R}$. Prove: If $f(x) \leq y$ for each $x \in S$, then $f(u) \leq y$. ◁

9.5. Algebra of continuous functions

All exercises in this section have the same structure. With the exception of Exercise 9.5.3, there are three functions in each exercise: f, g and h. The function h is always related in a simple (green) way to the functions f and g. Based on the given (green) information about f and g you are asked to prove a claim (red) about the function h.

Exercise 9.5.1. Let $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ be given functions with a common domain. Define the function $h : I \to \mathbb{R}$ by

$$h(x) = f(x) + g(x), \qquad x \in I.$$

(a) If f and g are continuous at $x_0 \in I$, then h is continuous at x_0.
(b) If f and g are continuous on I, then h is continuous on I. ◁

Exercise 9.5.2. Let $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ be given functions with a common domain. Define the function $h : I \to \mathbb{R}$ by

$$h(x) = f(x)g(x), \qquad x \in I.$$

(a) If f and g are continuous at $x_0 \in I$, then h is continuous at x_0.
(b) If f and g are continuous on I, then h is continuous on I. ◁

Exercise 9.5.3. Let $g : I \to \mathbb{R}$ be a given functions such that $g(x) \neq 0$ for all $x \in I$. Define the function $h : I \to \mathbb{R}$ by

$$h(x) = \frac{1}{g(x)}, \qquad x \in I.$$

(a) If g is continuous at $x_0 \in I$, then h is continuous at x_0.
(b) If g is continuous on I, then h is continuous on I. ◁

Exercise 9.5.4. Let $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ be given functions with a common domain. Assume that $g(x) \neq 0$ for all $x \in I$. Define the function $h : I \to \mathbb{R}$ by

$$h(x) = \frac{f(x)}{g(x)}, \qquad x \in I.$$

(a) If f and g are continuous at $x_0 \in I$, then h is continuous at x_0.
(b) If f and g are continuous on I, then h is continuous on I. ◁

Exercise 9.5.5. Let I and J be non-empty subsets of $\mathbb{R}$. Let $f : I \to \mathbb{R}$ and $g : J \to \mathbb{R}$ be given functions. Assume that the range of f is contained in J. Define the function $h : I \to \mathbb{R}$ by

$$h(x) = g(f(x)), \qquad x \in I.$$

(a) If f is continuous at $x_0 \in I$ and g is continuous at $f(x_0) \in J$, then h is continuous at x_0.
(b) If f is continuous on I and g is continuous on J, then h is continuous on I. ◁

9.6. Continuous functions on a closed bounded interval $[a, b]$

In this section we assume that $a, b \in \mathbb{R}$ and $a < b$.

EXERCISE 9.6.1. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. If $f(a) < 0$ and $f(b) > 0$, then there exists $c \in [a, b]$ such that $f(c) = 0$.
HINT: Consider the set

$$W = \bigl\{ w \in [a, b) : \forall x \in [a, w] \ f(x) < 0 \ \bigr\}.$$

Prove the following properties of W:

(i) W does not have a maximum.
(ii) W has a supremum. Set $w = \sup W$.
(iii) Review Exercise 9.4.2.
(iv) Connect the dots.

◁

EXERCISE 9.6.2. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then there exists $c \in [a, b]$ such that $f(x) \leq f(c)$ for all $x \in [a, b]$.
HINT: Consider the set

$$W = \Bigl\{ v \in [a, b) : \exists \ z \in [a, b] \ \text{ such that } \ \forall \ x \in [a, v] \ \ f(x) < f(z) \ \Bigr\}.$$

Here $[a, a]$ denotes the set $\{a\}$. Prove the following properties of the set W:

(i) If $a < u$ and $[a, u) \subseteq W$ and there exists $t \in [a, b]$ such that $f(t) > f(u)$, then $u \in W$.
(ii) W does not have a maximum.
(iii) W has a supremum. Set $w = \sup W$ and prove $[a, w) \subseteq W$.
(iv) The items (ii) and (iii) yield information about w.

◁

EXERCISE 9.6.3. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then there exists $d \in [a, b]$ such that $f(d) \leq f(x)$ for all $x \in [a, b]$.
HINT: Use Exercise 9.6.2. ◁

EXERCISE 9.6.4. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then the range of f is a closed bounded interval.
HINT: Use Exercises 9.6.2, 9.6.3, and 9.6.1. ◁

EXERCISE 9.6.5. Consider the function $f(x) = x^2$, $x \in \mathbb{R}$.

(a) Prove that 2 is in the range of f.
(b) Prove that the range of f equals $[0, +\infty)$.

◁

DEFINITION 9.6.6. A function f is *increasing* on an interval I if $x, y \in I$ and $x < y$ imply $f(x) < f(y)$. A function f is *decreasing* if $x, y \in I$ and $x < y$ imply $f(x) > f(y)$. A function which is increasing or decreasing is said to be *strictly monotonic*.

EXERCISE 9.6.7. If f is continuous and increasing on $[a, b]$ or continuous and decreasing on $[a, b]$, then for each y between $f(a)$ and $f(b)$ there is exactly one $x \in [a, b]$ such that $f(x) = y$. ◁

EXERCISE 9.6.8. Let $f(x) = x^3 + x$, $x \in \mathbb{R}$. Prove that f has an inverse. That is, prove that for each $y \in \mathbb{R}$ there exists unique $x \in \mathbb{R}$ such that $f(x) = y$. $\triangleleft$