Introduction to Analysis

Piotr Hajłasz

Contents

Chapter 1

Logic

1.1 Propositional calculus

All mathematical reasoning is based on logic. Below we provide a brief introduction to the main concepts of logic. We will try to keep a balance between rigorousness and intuition.

Mathematics consists of statements that are true or false, and we need to find out which statements are true. If a "statement" is neither true of false, it is not a statement in the mathematical sense.

$$2+2=4 \qquad \text{is a true statement,}$$

$$1+2=4 \qquad \text{is a false statement.}$$

Let's have a lunch. It is not a statement, because it has no logical value of being true or false. *This sentence is false.* It is also not a statement. Assuming it is true we arrive to a contradiction and assuming it is false we also arrive to a contradiction, so it is neither true or false and thus it is not a statement in the sense of logic. The problem with this sentence is that it refers to itself. Although this sentence is not mathematical in its nature, later we will see that some mathematical sentences[1] refer to themselves causing antinomies.

Some statements depend on variables and they are sometimes true and sometimes false, depending on the value of the variable. Consider the statement

$$x^2 = 1.$$

It is true if $x = 1$ or if $x = -1$, false otherwise.

We say that statements p and q are *logically equivalent* if p is true if and only if q is true. We denote logically equivalence by $p \equiv q$ or by $p \Leftrightarrow q$. For example the two statements

$$(x^2 = 1) \qquad \text{and} \qquad (x \in \{-1, 1\})$$

are logically equivalent, so we can write

$$(x^2 = 1) \quad \equiv \quad (x \in \{-1, 1\}).$$

[1]See Russel's paradox in Section 2.2.

Given statements p and q, we can build new statements.

Conjunction: "p and q" dentoed "$p \wedge q$".

Disjunction: "p or q" dentoed "$p \vee q$".

The statement $p \wedge q$ is true if and only if both statements p and q are true

$$(2+2=4) \quad \wedge \quad (2+2=5)$$

is false, while

$$(2+2=4) \quad \wedge \quad (2+3=5)$$

is true. The statement $p \vee q$ is true if and only if at least one of the statements p and q is true, so

$$(2+2=4) \quad \vee \quad (2+2=5)$$

is true, but

$$(2+2=5) \quad \vee \quad (2+2=6)$$

is false.

To a true statement we assign a logical value 1 and 0 to a false one. Thus we may summarize the above discussion using a *truth table*:

p	q	$p \wedge q$	$p \vee q$
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

Negation of a statement p is the statement

"not p" denoted $\neg p$

that is true if and only if p is false, so we have

p	$\neg p$
0	1
1	0

Another, less obvious construction is the *implication*

$$(p \text{ implies } q) \quad \text{or} \quad (\text{If } p, \text{ then } q) \quad \text{or} \quad (p \Rightarrow q).$$

Here p is called *hypothesis* and q is called *conclusion*. When is the implication true? Let's consider an example.

$$\underbrace{\text{If I win the lottery,}}_{p} \ \underbrace{\text{then}}_{\Rightarrow} \ \underbrace{\text{I will buy a house}}_{q}.$$

If I win the lottery, but decide not to buy a house, I break my promise making the above statement false,

p is true, q is false; $\quad p \Rightarrow q$ is false.

If I win the lottery and buy a house I fulfill my promise and the statement is true,

$$p \text{ is true, } q \text{ is true;} \qquad p \Rightarrow q \text{ is true.}$$

If I do not win in the lottery, then I have no obligation to buy a house. Thus no matter what I do, I will not break my promise. I can decide to buy a house or not to buy a house and in either case the statement is true.

$$p \text{ is false, } q \text{ is true or false;} \qquad p \Rightarrow q \text{ is true.}$$

Thus the only situation when the implication is false is when the hypothesis p is true, but the conclusion q is false. In all other cases the implication is true.

p	q	$p \Rightarrow q$
0	0	1
0	1	1
1	0	0
1	1	1

That seems obvious, but there are situations when it is not.

Example 1.1. The following implication is true since false assumption always implies a true conclusion

$$2 + 2 = 5 \quad \Longrightarrow \quad 2 + 2 = 4.$$

Although is does seem right, we can deduce the equality $2 + 2 = 4$ from the assumption that $2 + 2 = 5$ as follows. If $2 + 2 = 5$, then $0 \cdot (2 + 2) = 0 \cdot 5$, $0 = 0$. Adding 4 yields $4 = 4$. Replacing the left hand side with $2 + 2$ (we can do it because we know that $2 + 2 = 4$) gives $2 + 2 = 4$.

Example 1.2.

$$2 + 2 = 4 \qquad \Longrightarrow \qquad \text{Ronald Reagan was a president.}$$

The implication is certainly true, because both the hypothesis and the conclusion are true, but there is no way we can logically deduce from a simple mathematical equality $2 + 2 = 4$, the historical fact that Reagan was a president. That does not matter. In the mathematical logic, for the implication to be true we do not require the conclusion to be deducible from the hypothesis.

Example 1.3. Let us consider the following bizarre statement:

If I will eat my car, my right hand will turn into a leg.

That sounds quite insane, but the statement is true. Since I will not ear my car[2], the hypothesis is false, and thus the implication is true.

[2] Michel Lotito eat Cesna 150 so eating a car is possible, but I am not that crazy.

If p and q are statements, then the *equivalence* $p \equiv q$ is also a statement with obvious rules

p	q	$p \equiv q$
0	0	1
0	1	0
1	0	0
1	1	1

Indeed, $p \equiv q$ is true if and only if both p and q are true or false at the same time. Intuitively it is obvious that the following statements are equivalent

$$(p \equiv q) \qquad \text{and} \qquad \big((p \Rightarrow q) \wedge (q \Rightarrow p)\big).$$

The reader may verify the equivalence formally using a truth table in a similar way to the verification of the De Morgan Laws in (1.1) provided below.

Another, easy to check example of equivalent statements is that p is logically equivalent to $\neg(\neg p)$,

$$p \quad \equiv \quad \neg(\neg p).$$

As an application of the above discussion we will prove *De Morgan's Laws:*

$$\neg(p \wedge q) \quad \equiv \quad (\neg p) \vee (\neg q)$$

$$\neg(p \vee q) \quad \equiv \quad (\neg p) \wedge (\neg q).$$

We will discuss the first law only, the argument for the second one is similar. The law says that the conjunction $p \wedge q$ is false if and only if p is false or q is false, and that seems obvious, but we can (and we should to be mathematically rigorous) justify De Morgan's Laws formally using a truth table.

(1.1)

p	q	$p \wedge q$	$\neg(p \wedge q)$	$\neg p$	$\neg q$	$(\neg p) \vee (\neg q)$
0	0	0	**1**	1	1	**1**
0	1	0	**1**	1	0	**1**
1	0	0	**1**	0	1	**1**
1	1	1	**0**	0	0	**0**

By looking at the fourth and the seventh column in the above table we see that the statements $\neg(p \wedge q)$ and $(\neg p) \vee (\neg q)$ are both true or both false at the same time, so they are logically equivalent. That proves the first of the two De Morgan's laws.

Another example, less obvious than De Morgan's Laws, is that the following statements are equivalent

$$p \wedge (q \vee r) \quad \equiv \quad (p \wedge q) \vee (p \wedge r) \tag{1.2}$$

$$p \vee (q \wedge r) \quad \equiv \quad (p \vee q) \wedge (p \vee r)$$

We will prove the equivalence of statements in (1.2) leaving details of the second equivalence as an exercise.

p	q	r	$p \wedge (q \vee r)$	$(p \wedge q) \vee (p \wedge r)$
0	0	0	**0**	**0**
0	0	1	**0**	**0**
0	1	0	**0**	**0**
0	1	1	**0**	**0**
1	0	0	**0**	**0**
1	0	1	**1**	**1**
1	1	0	**1**	**1**
1	1	1	**1**	**1**

Example 1.4. *Prove that the statement* $p\wedge\neg(q\wedge\neg r)$ *is logically equivalent to* $(p\wedge\neg q)\vee(p\wedge r)$*.*

Proof. We could check it using a table as above, but instead we will prove it directly using the rules that we have already verified.

$$\begin{array}{cc} p \wedge \neg(q \wedge \neg r) & \stackrel{\text{De Morgan}}{\equiv} \\ p \wedge ((\neg q) \vee (\neg(\neg r))) & \equiv \\ p \wedge ((\neg q) \vee r) & \stackrel{(1.2)}{\equiv} \\ (p \wedge \neg q) \vee (p \wedge r) \ . & \end{array}$$

□

1.2 Proofs

In this section we will discuss the logical structure of proofs.

1.2.1 Direct proof

The only situation when an implication is false is when the hypothesis is true and the conclusion is false, so in order to prove that $p \Rightarrow q$ we can assume that p is true and we need to show that q is true as well. This is so called *direct proof.*

Example 1.5. *Prove that if* $a, b \geq 0$*, then* $\frac{a+b}{2} \geq \sqrt{ab}$*.*

Proof. Let $a, b \geq 0$. Then $(a-b)^2 \geq 0$, $a^2 - 2ab + b^2 \geq 0$, $a^2 + 2ab + b^2 \geq 4ab$, $(a+b)^2 \geq 4ab$ and taking the square root gives $a+b \geq 2\sqrt{ab}$. Dividing both sides by 2 yields $\frac{a+b}{2} \geq \sqrt{ab}$. □

1.2.2 Proof by contradiction

Let us start with an example of a simple proof by contradiction. Then we will carefully analyze the structure of the proof from the perspective of logic.

Example 1.6. *Let n be an integer. If n^3+5 is odd, then n is even.*

Proof. Suppose to the contrary that n^3+5 is odd and n is odd to. Then $n^3+5=2k+1$, $n=2\ell+1$ for some integers k and ℓ and a simple calculation yields

$$2k+1=(2\ell+1)^3+5, \qquad 2k=8\ell^3+12\ell^2+6\ell+5$$

which is a contradiction since in the last equality the left hand side is even while the right one is odd. The contradiction completes the proof. $\square$

The logical argument used in the proof is based on the fact that the following two statements are logically equivalent[3]

$$(p\Rightarrow q) \quad\equiv\quad \neg(p\wedge\neg q). \tag{1.3}$$

Therefore, in order to prove that the statement $p\Rightarrow q$ is true it suffices to prove that the statement $\neg(p\wedge\neg q)$ is true, i.e. that the statement $p\wedge\neg q$ is false. How do we prove that the statement $p\wedge\neg q$ is false? We assume that it is true and we arrive to a contradiction.

Let us look at the above proof again to see that it follows the logical argument described here. We want to prove

$$\underbrace{\text{If } n^3+5 \text{ is odd}}_{p},\ \underbrace{\text{then}}_{\Rightarrow}\ \underbrace{n \text{ is even}}_{q}.$$

To this end it suffices to prove that the following statement is false

$$\underbrace{n^3+5 \text{ is odd}}_{p}\ \underbrace{\text{and}}_{\wedge}\ \underbrace{n \text{ is odd}}_{\neg q}. \tag{1.4}$$

So we assume that (1.4) is true and, after some calculations, we arrive to a contradiction. This completes the proof.

The argument by contradiction is very common in mathematics and we often employ it using our intuition without realizing that the method of the proof by contradiction can be formally justified in the mathematical logic.

Sometimes a simpler argument is enough. In order to prove that the statement p is true it suffices to prove that the statement $\neg p$ is false. Here are two examples. The proof of Example 1.7 is due to Euclid (c. 300 BC).

Example 1.7. *There are infinitely many prime numbers.*

Proof. Suppose to the contrary that there are only finitely many primes. List them as $p_1,p_2,\ldots,p_n$. Consider the number $q=p_1p_2\cdot\ldots\cdot p_n+1>1$. Let p be a prime factor of q. Then $p\not\in\{p_1,p_2,\ldots,p_n\}$, because otherwise it would divide $q-p_1p_2\cdot\ldots\cdot p_n=1$. Contradiction. The proof is complete. $\square$

Example 1.8. *$\sqrt{2}$ is irrational.*

[3] We leave the proof of the equivalence (1.3) as an exercise.

Proof. Suppose to the contrary that $\sqrt{2}$ is rational. Then $\sqrt{2} = p/q$ for some positive integers and we may assume that p and q have no common factors. We have

$$2 = \frac{p^2}{q^2}, \qquad p^2 = 2q^2.$$

Therefore p^2 is even. Since the square of an odd number is odd, we have that p is even, $p = 2k$ so

$$(2k)^2 = 2q^2, \qquad 4k^2 = 2q^2, \qquad 2k^2 = q^2.$$

Thus q^2 is even and since the square of an odd number is odd, we have that q is even. We proved that both numbers p and q are even which contradicts the assumption that the numbers p and q have no common factors. □

1.2.3 Proof by contrapositive

This proof is based on the equivalence

$$(p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p). \tag{1.5}$$

Here is an example.

Example 1.9. *Let $x, y \in \mathbb{R}$. If $y^3 + yx^2 \leq x^3 + xy^2$, then $y \leq x$.*

In order to prove

$$\underbrace{y^3 + yx^2 \leq x^3 + xy^2}_{p} \Longrightarrow \underbrace{y \leq x}_{q}$$

it suffices to prove that

$$\underbrace{y > x}_{\neg q} \Longrightarrow \underbrace{y^3 + yx^2 > x^3 + xy^2}_{\neg p}.$$

Proof. Suppose that it is not true that $y \leq x$, i.e. $y > x$. Then $y - x > 0$ and we have

$$(y - x)(x^2 + y^2) > 0 \qquad (\text{because } x^2 + y^2 > 0)$$

$$yx^2 + y^3 - x^3 - xy^2 > 0$$

$$y^3 + yx^2 > x^3 + xy^2$$

so it is not true that $y^3 + yx^2 \leq x^3 + xy^2$. The proof is complete. □

The proof by contrapositive is less common than the proof by contradiction.

Here is a non mathematical application of the equivalence (1.5)

Example 1.10. The following two sentences were written on a jar of a barbecue sauce[4]:

Good goods are never cheap.

Cheap goods are never good.

Do the two sentences have the same meaning? A product is either good (G) or not good ($\neg G$), cheap (C) or not cheap ($\neg C$) so the sentences can be stated as follows

$$\begin{aligned} G &\Rightarrow \neg C \\ C &\Rightarrow \neg G. \end{aligned}$$

According to (1.5) we have

$$(G \Rightarrow \neg C) \quad \equiv \quad (\neg(\neg C) \Rightarrow \neg G) \quad \equiv \quad (C \Rightarrow \neg G).$$

Therefore the sentences have the same meaning so adding the second one was redundant since it did not provide any new information.

1.3 Context and quantifiers.

Mathematical statements often involve variables like e.g.

$$x^2 = 1.$$

It is important to fix the context in which we consider this statement. That means we need to fix a set from which we choose x. It can be the set of all complex numbers, the set of all real numbers, the set of all integers, etc.

Very often mathematical statements involve phrases *for all* or *there exists*. For example the statement

For all integeres x, $x^2 = 1$

is false, but the statement

There is an integer x such that $x^2 = 1$

is true.

If we fix the context first, we can express the two statements in a shorter way. Let x be an integer.

For all x, $x^2 = 1$.

There is x such that $x^2 = 1$.

The two statements are equivalent to the previous ones. We did not have to include the information that x is an integer into the statements, because it was assumed earlier.

The phrase *for all* or *for every* is called the *universal quantifier* and is dented by the symbol $\forall$. The phrase *there exists* is called the *existential quantifier* and is denoted by $\exists$.

[4] Daddy Sam's Bar-B-Que Sawce.

Let x be an integer, then we can reformulate the above two statements as follows

$$\forall x \ (x^2 = 1)$$

$$\exists x \ (x^2 = 1).$$

Sometimes we may explain the context directly in the formula as in the following true statement:

$$\forall x \in \mathbb{R} \ \left(x^2 = 1 \Rightarrow x \in \{-1, 1\}\right)$$

That fact that it is true seems obvious, but let's look at it more carefully. The statement says that for every $x \in \mathbb{R}$ the statement that follow the quantifier is true. Since it is true for every $x \in \mathbb{R}$, in particular, it is true for $x = 0$:

$$0^2 = 1 \quad \Rightarrow \quad 0 \in \{-1, 1\}.$$

While this implication seems ridiculous, it is certainly true, because an implication with the false hypothesis is always true!

Let x and y be real numbers. The statement

$$\forall x \ \exists y \ (x + y = 0)$$

means:

For every real x there is a real y such that $x + y = 0$.

This statement is obviously true.

The order of quantifiers is very important.

$$\forall x \ \exists y \ (x + y = 0) \qquad \text{is true}$$

while

$$\exists y \ \forall x \ (x + y = 0) \qquad \text{is false.}$$

More generally if $P(x, y)$ is a statement depending of two variables x and y (e.g. $x+y = 0$) we can consider statements like

$$\forall x \ \exists y \ P(x, y).$$

The statement $P(x, y)$ is often called a *property*. Indeed, this is a property that is satisfied (when true) by the given x and y or is not satisfied (when false). The example of the property $x + y = 0$ shows that in general $\exists y \ \forall x \ P(x, y)$ does not imply $\exists y \ \forall x \ P(x, y)$ since a true statement does not imply a false one. However we have.

Example 1.11. *For any property P, $\left(\exists y \ \forall x \ P(x, y)\right) \Rightarrow \left(\forall x \ \exists y \ P(x, y)\right)$.*

Proof. Since the only situation when an implication can be false is when the hypothesis is true and the conclusion is false, it suffices to assume that the hypothesis is true and then to show that the conclusion is true as well.

Assume that the statement

$$\exists y \ \forall x \ P(x, y) \tag{1.6}$$

is true. We need to show that the other statement

$$\forall x \; \exists y \; P(x,y) \tag{1.7}$$

is true as well. From the assumption that (1.6) is true we see that there is y, denote it by y_0, such that for all x, $P(x, y_0)$ is true so for all x we can find y (namely y_0) such that $P(x, y)$ is true which is precisely (1.7). □

It is of fundamental importance to understand how to negate statements that involve quantifiers. Let A be a set. Observe that the following two statements are equivalent

$$\neg(\forall x \in A \;\; P(x)) \quad \equiv \quad \exists x \in A \;\; (\neg P(x))$$

Indeed, the fact that for all $x \in A$, $P(x)$ is satisfied is not true means that there is $x \in A$ such that $P(x)$ is not satisfied. Similarly we justify the equivalence

$$\neg(\exists x \in A \; P(x)) \quad \equiv \quad \forall x \in A \; (\neg P(x)).$$

Note that we reverse the quantifier and negate the statement, but we do not negate the condition $x \in A$. For example the negation of the (false) statement

$$\forall \varepsilon > 0 \; (\varepsilon^2 \geq 1)$$

is the (true) statement

$$\exists \varepsilon > 0 \; (\varepsilon^2 < 1).$$

Now we can apply the two rules to the statements that involve more than one quantifier. For example

$$\neg(\forall x \; \exists y \; P(x,y)) \quad \equiv \quad \exists x \; (\neg(\exists y \; P(x,y))) \quad \equiv \quad \exists x \; \forall y \; (\neg P(x,y)).$$

We can apply the same argument to the case in which we have more than two quantifiers, e.g.

$$\neg(\forall x \; \exists y \; \exists z \; \forall w \; P) \quad \equiv \quad \exists x \; \forall y \; \forall z \; \exists w \; (\neg P).$$

The rule is quite clear, I hope. Each time we change the quantifier to the other one and then we negate the condition P. Let's see how it works. The statement

$$\exists x \; \forall y \; (x + y = 0)$$

is clearly false, so its negation should be true

$$\neg(\exists x \; \forall y \; (x + y = 0)) \quad \equiv \quad \forall x \; \exists y \; (x + y \neq 0)$$

and that is indeed true.

Example 1.12. *Prove the statement*

$$(\forall \varepsilon > 0) \; (\exists n \in \mathbb{N}) \; (n^{-1} < \varepsilon).$$

Proof. That statement means that for every positive ε there is a natural number n such that $n^{-1} < \varepsilon$. Suppose to the contrary that the statement is false. Then its negation is true, i.e.[5]

$$\begin{aligned} \neg\big((\forall \varepsilon > 0)\ (\exists n \in \mathbb{N})\ (n^{-1} < \varepsilon)\big) &\equiv (\exists \varepsilon > 0)\ (\forall n \in \mathbb{N})\ \neg(n^{-1} < \varepsilon) \\ &\equiv (\exists \varepsilon > 0)\ (\forall n \in \mathbb{N})\ (n^{-1} \geq \varepsilon)\,. \end{aligned}$$

That is

$$(\exists \varepsilon > 0)\ (\forall n \in \mathbb{N})\ (n^{-1} \geq \varepsilon)$$

is true. It says that we can find an $\varepsilon > 0$ such that for every n, $n^{-1} \geq \varepsilon$. Fix such $\varepsilon > 0$. Then for all positive integers n, $n^{-1} \geq \varepsilon$.

Let n_0 be any positive integer such that $n_0 > \varepsilon^{-1}$. Since the statement above is true for all positive integers, it is true, in particular, for $n = n_0$, so we have $n_0^{-1} \geq \varepsilon$. This however, implies that $n_0 \leq \varepsilon^{-1}$ which contradicts the fact that n_0 is bigger than ε^{-1}. The contradiction proves the claim. □

Remark 1.13. A good question is, how did we know that we should consider $n_0 > \varepsilon^{-1}$. We will learn it later, but for checking that the proof is correct we do not need to know how we got this argument. The next example will have a similar issue. We will take $\delta = \sqrt{\varepsilon}/2$ without any clue why.

The next example is similar in nature, but more complicated because of a more involved logical structure of the statement.

Example 1.14. *Prove that the following statement is true*[6]

$$(\forall \varepsilon > 0)\ (\exists \delta > 0)\ (\forall x \in \mathbb{R})\ \big(|x| < \delta \Rightarrow |x^2| < \varepsilon\big)\,.$$

Proof. The statement that we plan to prove should be read as follows. For every $\varepsilon > 0$ there is $\delta > 0$ such that for ever $x \in \mathbb{R}$, if $|x| < \delta$, then $|x^2| < \varepsilon$.

Suppose to the contrary that the statement is false, then its negation is true

$$\begin{array}{c} \neg\big((\forall \varepsilon > 0)\ (\exists \delta > 0)\ (\forall x \in \mathbb{R})\ (|x| < \delta \Rightarrow |x^2| < \varepsilon)\big) \quad \equiv \\ (\exists \varepsilon > 0)\ (\forall \delta > 0)\ (\exists x \in \mathbb{R})\ \neg(|x| < \delta \Rightarrow |x^2| < \varepsilon) \quad \equiv \\ (\exists \varepsilon > 0)\ (\forall \delta > 0)\ (\exists x \in \mathbb{R})\ (|x| < \delta\ \wedge\ |x^2| \geq \varepsilon). \end{array}$$

In the last step we used equivalence of the statements

$$\neg(p \Rightarrow q) \quad \equiv \quad p \wedge \neg q\,.$$

Thus the following statement is true and we need to arrive to a contradiction.

$$(\exists \varepsilon > 0)\ (\forall \delta > 0)\ (\exists x \in \mathbb{R})\ (|x| < \delta\ \wedge\ |x^2| \geq \varepsilon).$$

[5] As we have already explained earlier we reverse quantifiers and negate the statement $n^{-1} < \varepsilon$, but we keep conditions $\varepsilon > 0$ and $n \in \mathbb{N}$ unchanged.

[6] A reader familiar with a formal definition of continuity will see that the statement says that the function $f(x) = x^2$ is continuous at $x = 0$. However, at this point we are not interested in studying continuity and we only want to learn logic behind a proof of a statement.

The statement says that there is $\varepsilon > 0$ such that for every $\delta > 0$ there is x such that $|x| < \delta$ and $|x^2| \geq \varepsilon$.

Fix such $\varepsilon > 0$. Since for every $\delta > 0$ there is x such that $|x| < \delta$ and $|x^2| \geq \varepsilon$, it is true for $\delta = \sqrt{\varepsilon}/2$. That means for $\delta = \sqrt{\varepsilon}/2$ there is x such that

$$|x| < \sqrt{\varepsilon}/2 \quad \text{and} \quad |x^2| \geq \varepsilon.$$

However, the first inequality gives $|x^2| < \varepsilon/4$ which contradicts the second one. The proof is complete. □

This proof is difficult, but arguments of this type will appear in our reasoning very often and eventually we will learn how to use them.

1.4 Problems

Problem 1. Use the truth table to prove equivalence of the statements:
$p \equiv q$ and $(p \Rightarrow q) \wedge (q \Rightarrow p)$.

Problem 2. Use the truth table to prove equivalence of the statements:

$$(p \Rightarrow q) \quad \equiv \quad \neg(p \wedge \neg q) \quad \equiv \quad (\neg q \Rightarrow \neg p).$$

Problem 3. Use the equivalence

$$p \wedge (q \vee r) \quad \equiv \quad (p \wedge q) \vee (p \wedge r) \tag{1.8}$$

to prove

$$p \vee (q \wedge r) \quad \equiv \quad (p \vee q) \wedge (p \vee r).$$

To this end apply (1.8) to $\neg p$, $\neg q$, $\neg r$ in place of p, q, r, and negate the statement using De Morgan's Laws.

Problem 4. Negate the statement[7]

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in \mathbb{R} \;\; \forall y \in \mathbb{R} \;\; (|x - y| < \delta \;\Rightarrow\; |\sin x - \sin y| < \varepsilon).$$

Problem 5. Negate the statement: *For all real numbers x, y satisfying $x < y$, there is a rational number q such that $x < q < y$.* Formulate the negation as a sentence and not as a formula involving quantifiers.

Problem 6. Use an argument by contradiction prove that $\sqrt{3}$ is irrational.

Problem 7. In Example 1.5 we provided a direct proof. Prove the same statement using a proof by contradiction.

Problem 8. Prove the following statement[8]

$$\forall \varepsilon > 0 \;\; \exists n_0 \in \mathbb{N} \;\; \forall n \in \mathbb{N} \quad (n \geq n_0 \;\Rightarrow\; n^{-1} \leq \varepsilon).$$

[7] This is a true statement known as uniform continuity of the function $\sin x$. However, you are not asked to prove the statement only to negate it.

[8] Compare with Example 1.12.

Problem 9. Find all functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy the following property[9]

$$\forall \varepsilon > 0 \, \forall x, y \in \mathbb{R} \, \exists \delta > 0 \quad (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Problem 10. Find a mistake in the following 'solution' to Problem 9.

"Solution". We will prove that only constant functions satisfy the condition. Clearly constand functions satisfy it, because $|f(x) - f(y)| = 0 < \varepsilon$ for all $x, y \in \mathbb{R}$. Suppose now that a function f satisfies the condition. We have to prove that f is constant. The condition says that for any $\varepsilon > 0$ and any $x, y \in \mathbb{R}$ we can find $\delta > 0$ such that

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Since the inequality $|f(x) - f(y)| < \varepsilon$ is required to hold for all $\varepsilon > 0$ we have that $|f(x) - f(y)| \leq 0$ and hence $|f(x) - f(y)| = 0$, $f(x) = f(y)$. Since for all $x, y \in \mathbb{R}$ we have $f(x) = f(y)$ the function f is constant. "□"

[9] A reader familiar with uniform continuity should notice that this is not the condition for the uniform continuity, although it looks very similar: the order of quantifiers has been changed, see Section 11.11.

Chapter 2

Set theory

2.1 Elementary set theory

A set is a collection of arbitrary objects. A priori there are no restrictions of what objects we consider, but we will soon see that one has to be careful. There is no formal definition of a set and we only use our intuition and imagination to create sets.

The simplest set is the *empty set* denoted by $\emptyset$. This is a set with no elements at all.

By writing $x \in A$ we denote that x is an element of the set A. If x is *not* an element of A, we write $x \notin A$. That is

$$\neg(x \in A) \qquad \equiv \qquad x \notin A.$$

Two sets are equal if they have the same elements. Formally

$$(A = B) \qquad \equiv \qquad \forall x \;\; (x \in A \;\equiv x \in B).$$

We say that A is a subset of B, $A \subset B$ (or $B \supset A$) if every element of A is also an element of B. Formally

$$(A \subset B) \qquad \equiv \qquad \forall x \;\; (x \in A \;\Rightarrow x \in B)$$

Therefore

$$(A = B) \qquad \equiv \qquad (A \subset B \;\wedge\; B \subset A).$$

Is the empty set of real numbers the same as the empty set of chairs? Yes. They both have the same elements – no elements at all.

Sets can be defined by listing all of the elements, e.g.

$$A = \{1, 3, 5\} = \{5, 1, 3\}.$$

This is a set with three elements 1, 3 and 5 only. The order of elements in a set is not important, this is why $\{1, 3, 5\} = \{5, 1, 3\}$. Also

$$\{1, 1, 2, 5, 5\} = \{1, 2, 5\},$$

because both sets have the same elements: 1, 2, 5.

$$\mathbb{N} = \{1, 2, 3, 4, \ldots\}$$

is the set of all natural numbers, or, in other words, all positive integers.

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

is the set of all integers.

If A is a given set and $P(x)$ is a property of elements of A that may or may not be satisfied by a given $x \in A$, then we can define a new set

$$\{x \in A : \, P(x)\}.$$

This is the set of all elements in A such that the property $P(x)$ is satisfied. For example

$$\{x \in \mathbb{N} : \, x \text{ is even}\} = \{2, 4, 6, 8, \ldots\}$$

$$\{x \in \mathbb{R} : \, x^2 \leq 1\} = [-1, 1].$$

The *union* of sets A and B is defined by

$$A \cup B = \{x : \, x \in A \vee x \in B\}.$$

That means the union $A \cup B$ is a set that consists of all elements of A, all elements of B and no other element. The *intersection* of the sets $A \cap B$ consists of elements that belong to both sets A and B at the same time, that is

$$A \cap B = \{x : \, x \in A \wedge x \in B\}.$$

Example 2.1. If $a \in A$, then $\{a\} \subset A$, $\{a\} \cup A = A$ and $\{a\} \cap A = \{a\}$.

We can also define the union and the intersection of an infinite family of sets. If $A_1, A_2, A_3, \ldots$ are sets, then we define

$$\bigcup_{n=1}^{\infty} A_n = A_1 \cup A_2 \cup A_3 \cup \ldots = \{x : \, \exists n \; x \in A_n\}$$

$$\bigcap_{n=1}^{\infty} A_n = A_1 \cap A_2 \cap A_3 \cap \ldots = \{x : \, \forall n \; x \in A_n\}.$$

In the above unions and intersections of infinite families of sets the sets were arranged into a sequence, but we can also think about other infinite families of sets that even cannot be arranged into a sequence. For example we may think of a situation that we have a set associated with each real number

$$A_t, \quad t \in \mathbb{R}.$$

Then we define

$$\bigcup_{t \in \mathbb{R}} A_t = \{x : \, \exists t \in \mathbb{R} \; (x \in A_t)\}, \qquad \bigcap_{t \in \mathbb{R}} A_t = \{x : \, \forall t \in \mathbb{R} \; (x \in A_t)\}.$$

Example 2.2.

$$\bigcup_{t\in\mathbb{R}}\{t\} = \mathbb{R}, \qquad \bigcap_{t\in\mathbb{R}}\{t\} = \emptyset.$$

More generally, if I is a set and for each $i \in I$, A_i is a set we can talk about the set $\{A_i\}_{i\in I}$ of all lets A_i. Then we can define

$$\bigcup_{i\in I} A_i = \{x :\ \exists i \in I\ (x \in A_i)\} \quad \text{and} \quad \bigcap_{i\in I} A_i = \{x :\ \forall i \in I\ (x \in A_i)\}.$$

A set $\{A_i\}_{i\in I}$ whose elements are sets is often called a *family of sets*. In the definition of the sum and intersection there is no restriction how large family of sets we consider.

The *complement of A relative to B* or just the *difference* of sets A and B is defined by

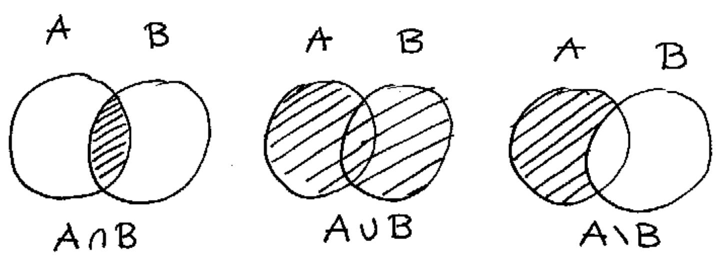
$$B \setminus A = \{x \in B :\ x \notin A\}$$

so

$$x \in B \setminus A \quad \equiv \quad (x \in B) \wedge \neg(x \in A).$$

Example 2.3. $\mathbb{N} \setminus \{2, 4, 6, 8, \ldots\} = \{1, 3, 5, 7, \ldots\}$

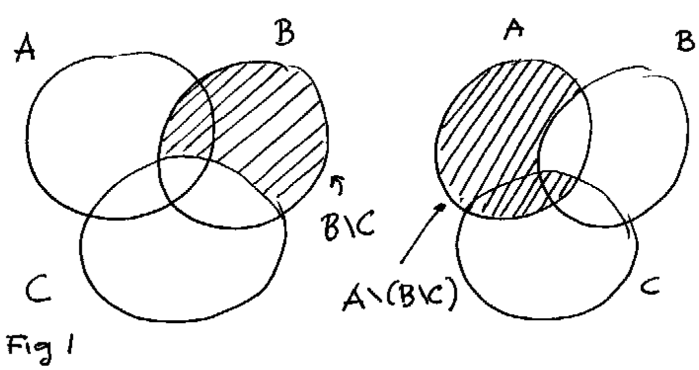
The operations on sets can be represented graphically using *Venn diagrams*.



Example 2.4. *Prove that $A \setminus (B \setminus C) = (A \setminus B) \cup (A \cap C)$.*

Proof. This is an important example, because it shows different methods that can be used to study operations on sets and also it shows a direct link to logic.

First we will present a geometric and not very rigorous proof based on Venn diagrams.



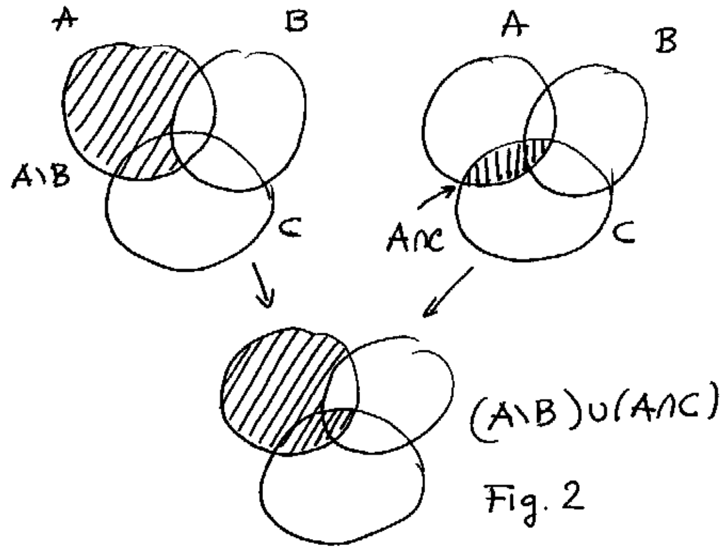


Fig. 2

The result follows after comparing Fig. 1 with Fig. 2.

Now we will show a rigorous proof that will employ methods of logic developed earlier. We will prove that the two sets are equal by showing that they have the same elements. Recall that

$$p \wedge (q \vee r) \quad \equiv \quad (p \wedge q) \vee (p \wedge r).$$

We have

$$\begin{array}{cc} x \in A \setminus (B \setminus C) & \equiv \\ x \in A \ \wedge \ \neg(x \in B \setminus C) & \equiv \\ x \in A \ \wedge \ \neg(x \in B \ \wedge \ x \notin C) & \equiv \\ x \in A \ \wedge \ (x \notin B \ \vee \ x \in C) & \equiv \\ (x \in A \ \wedge \ x \notin B) \ \vee \ (x \in A \ \wedge \ x \in C) & \equiv \\ (x \in A \setminus B) \ \vee \ (x \in A \cap C) & \equiv \\ x \in (A \setminus B) \cup (A \cap C) \ . & \end{array}$$

We proved that x is an element of $A \setminus (B \setminus C)$ if and only if x is an element of $(A \setminus B) \cup (A \cap C)$. Therefore the two sets are equal. □

Remark 2.5. Observe that the equivalence between the third and the fifth line follows from the equivalence

$$p \wedge \neg(q \wedge \neg r) \quad \equiv \quad (p \wedge \neg q) \vee (p \wedge r)$$

that we proved in Example 1.4. Indeed,

$$\underbrace{x \in A}_{p} \wedge \neg(\underbrace{x \in B}_{q} \ \wedge \ \underbrace{x \notin C}_{\neg r}) \quad \equiv \quad (\underbrace{x \in A}_{p} \ \wedge \ \underbrace{x \notin B}_{\neg q}) \vee (\underbrace{x \in A}_{p} \ \wedge \underbrace{x \in C}_{r})$$

The next result is a version of De Morgan Laws for sets.

Proposition 2.6.

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

Proof. We will prove only the first equality, the proof of the second one is similar. We have

$$\begin{array}{cc}
x \in A \setminus (B \cup C) & \equiv \\
x \in A \ \wedge \ \neg(x \in B \cup C) & \equiv \\
x \in A \ \wedge \ \neg(x \in B \ \vee \ x \in C) & \equiv \\
x \in A \ \wedge \ (x \notin B \ \wedge \ x \notin C) & \equiv \\
(x \in A \ \wedge \ x \notin B) \ \wedge \ (x \in A \ \wedge \ x \notin C) & \equiv \\
(x \in A \setminus B) \ \wedge \ (x \in A \setminus C) & \equiv \\
x \in (A \setminus B) \cap (A \setminus C) \ \ . &
\end{array}$$

We proved that x is an element of $A \setminus (B \cup C)$ if and only if x is an element of $(A \setminus B) \cap (A \setminus C)$. Therefore the two sets are equal. □

Similarly one can prove a more general form of the De Morgan Laws. We leave a proof as an exercise.

Proposition 2.7 (De Morgan Laws)**.** *For any set A and any family of sets $\{A_i\}_{i \in I}$*

$$A \setminus \bigcup_{i \in I} A_i = \bigcap_{i \in I} (A \setminus A_i)\,,$$

$$A \setminus \bigcap_{i \in I} A_i = \bigcup_{i \in I} (A \setminus A_i)\,.$$

2.2 Russell's paradox

Our intuition is that a set can be *any* collection of *any* elements. In particular elements of a set can also be sets. Let X be a set. Since its elements can be sets the statement $X \notin X$ makes sense and we may consider the set

$$A = \{X : \ X \text{ is a set } \ \wedge \ X \notin X\}.$$

Now we can ask an important question

Is A is an element of A?

Suppose that $A \in A$. Then it satisfies the condition $X \notin X$ from the definition of A so $A \notin A$ and we have a contradiction. Suppose now that $A \notin A$, then A is a set and it satisfies the condition $X \notin X$ from the definition of A so $A \in A$ and we arrive at a contradiction again. That simply means the set A cannot exist.

But then we may ask: what is a set? How can we be sure that a set we define will not lead to a contradiction?

The only way to handle this unpleasant and annoying situation is to create a list of "safety rules" which tell us what operations on sets are allowed and how we can create new sets from the ones that already exist. Such safety rules are called *axioms* and we are not allowed to do anything which cannot be concluded from axioms.

The axioms will be discusses in the next section, but let us mention one of the axioms now. The *axiom of specification* says that if X is a set and P is a property of elements of X, then the set

$$\{Y \in X : P(Y)\}$$

exists. In the Russel Paradox we tried to create the set:

$$A = \{X \in \text{All sets} : X \notin X\}.$$

Suppose that the set of all sets exists. Then according to the axiom of specification the set A exists too and we arrive to a contradiction which is explained above. Assuming that the set of all sets exists leads to a contradiction and that proves that the set of all sets cannot exist. Since the set of all sets does not exists, A is not a set, because the axiom of specification does not apply. We proved

Proposition 2.8. *The set of all sets does not exist.*

2.3 Axioms of the set theory*

The material of this section is difficult, because of a high level of abstraction. It is okay if after going through the material of this section you will be confused and lost. What is written here will not be really used in the subsequent sections so if you decide to skip this section it is okay, but I would advise to read it anyway. If you like it that is great. If not, then no much harm is done.

Formal axiomatic theory of sets, the so called *Zermelo-Fraenkel set theory with the Axiom of Choice* or just ZFC for short is difficult and it is not of our intention to present it with details. Instead, we will present a rather intuitive approach which however, will give a quite good understanding of what is it all about.

Mathematics deals with abstract entities. Integers are also abstract objects. While we can use them to count apples, integers by itself have nothing to do with objects in a real life.

The same applies to sets. We said that sets can be collections of any objects, but in mathematics we would need to know how to define these objects. Thus for simplicity we assume that all objects existing in the set theory are sets and so elements of sets can be sets only[1]. That is bizarre, I know, but that makes the theory a lot simpler. In fact we will only assume that the empty set exist and starting from that we will create the entire mathematical realm, including integers and real numbers. We will create everything from nothing (i.e. form the empty set). Isn't that great?

[1]In the following sections however, we will not be that restrictive and by a set we will mean any collection of objects that are not necessarily sets.

Axiom of Existence. *The empty set $\emptyset$ exists.*

Formally the empty set is defined by the condition

$$\exists Y \, \forall X \, (X \notin Y).$$

By know we know that at least one set exists: empty set. That is not much, but other axioms will allow us to construct new sets from it. Actually we will be able to construct all the sets that are needed in mathematics. The next axiom explains what it means that two sets are equal.

Axiom of Equality[2]**.** *Two sets are equal $X = Y$ if they have the same elements, i.e.,*

$$(X = Y) \;\equiv\; \forall Z (Z \in X \equiv Z \in Y)$$

It implies that there is exactly one empty set: the empty set of integers is equal to the empty set of complex numbers, because both sets have exactly the same elements – no elements at all.

For sets we define what it means that X is a subset of Y exactly as in Section 2 and then two sets X and Y are equal if and only if $X \subset Y$ and $Y \subset X$.

The next axiom tells us how to build new sets from the sets that exist.

Axiom of Pairing. *If the sets X and Y exist, then also the set $\{X, Y\}$ exists.*

Let us emphasize that the set $\{X, Y\}$ is a set with elements X and Y and it is *not* the union of sets X and Y.

As for now we only know that $\emptyset$ exists, but the Axiom of Pairing implies that also

$$\{\emptyset, \emptyset\} = \{\emptyset\}$$

exists. Note that both of the sets $\{\emptyset, \emptyset\}$ and $\{\emptyset\}$ are equal, because both have the same elements:

$$(X \in \{\emptyset, \emptyset\}) \quad \equiv \quad (X = \emptyset \;\vee\; X = \emptyset) \quad \equiv \quad (X = \emptyset) \quad \equiv \quad (X \in \{\emptyset\}).$$

Then also the following sets also exist:

$$\emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\},\ \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$$

It is easy to prove that these sets are different. For example $\emptyset \neq \{\emptyset\}$. Indeed, $\emptyset \in \{\emptyset\}$ while $\emptyset \notin \emptyset$, because $\emptyset$ has no elements. Therefore the sets $\emptyset$, and $\{\emptyset\}$ do not have the same elements so they are different.

We can now define natural numbers $\mathbb{N}_*$ including zero[3] by calling

$$0 := \emptyset,\ 1 := \{\emptyset\},\ 2 := \{\emptyset, \{\emptyset\}\},\ 3 := \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots \tag{2.1}$$

[2]Known also as the*Axiom of Extensionality.*

[3]While here we include zero among natural numbers, in the remaining sections of the book we will always assume that natural numbers are $\mathbb{N} = \{1, 2, 3, \ldots\}$.

Axiom of Union. *If I is a set (of sets), then there is a set whose elements belong to at least one of the elements of I.*

In other words the following set exists:

$$\bigcup_{X\in I} X = \bigcup\{X : X \in I\} := \{Y : (\exists\, X \in I)\,(Y \in X)\}.$$

Now we can define the operation of adding 1 to an integer in the following way:

If $n = X$ is a natural number, that is a member of the list (2.1), then we define $n + 1 = X \cup \{X\}$.

For example

$$2 = \underbrace{\{\emptyset, \{\emptyset\}\}}_{X} \quad \text{so} \quad 2+1 = X \cup \{X\} = \{\emptyset, \{\emptyset\}\} \cup \underbrace{\{\{\emptyset, \{\emptyset\}\}\}}_{\text{this set has one element}} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = 3.$$

Then we can define addition and multiplication of integers in terms of operations on sets (we will not provide details).

Once we know how to add natural numbers we can define the inequality $n \leq m$ as follows:

$$(n \leq m) \quad \equiv \quad (\exists a \in \mathbb{N}_* \; (n = a + m)).$$

In particular we have the relation $n \leq m$ between natural numbers and clearly 0 is the smallest natural number with respect to this relation.

Axiom of Specification[4]**.** *If X is a set and P is a property of elements of X, then the set*

$$\{Y \in X : P(Y)\}$$

exists.

The Axiom of Specification implies the existence intersection of sets. If X and Y are sets, then we define

$$X \cap Y = \{Z \in X : Z \in Y\}.$$

More generally if I is any nonempty set of sets, we define the intersection of all sets in I as follows. Fix any $X_0 \in I$ and set

$$\bigcap_{X\in I} X = \bigcap\{X : X \in I\} := \{Y \in X_0 : (\forall X \in I)\ (Y \in X)\}.$$

Therefore we do not need to add as an axiom the existence of an intersection of sets since it it a consequence of the axiom of specification. Although the Axiom of Specification can be used to define intersection of sets it cannot be used to define the union of sets so we had to state the existence of union of sets as an axiom.

[4]Known also as the *Axiom Schema of Specification*.

In particular, if $I = \{X_1, X_2, \ldots\}$, we write

$$\bigcap\{X : X \in I\} = \bigcap_{i=1}^{\infty} X_i = \{Y \in X_1 : \forall i \ (Y \in X_i)\}.$$

Here however, we went a bit ahead, because the set I is infinite and we do not know yet that infinite sets exist. We will get to that.

Axiom of Powers. *For a given set X there is a set $P(X)$ whose elements are all subsets of X.*

The set $P(X)$ is called the *power set.* That explains the name of the axiom. One can prove that if a set X has n elements, then $P(X)$ has 2^n elements and for that reason the power set is sometimes denoted by 2^X.

The axioms listed above allow us to construct new sets starting from the empty set. However, in all the constructions we end up with finite sets only. We could define natural numbers, and clearly there are infinitely many natural numbers, but we do not know yet whether the set of all natural numbers exist. We cannot conclude yet that any infinite set exists. The problem is solved by introducing the Axiom of Infinity.

Axiom of Infinity. *There is a set X such that $\emptyset \in X$ and whenever Y is an element of X, then also $Y \cup \{Y\}$ is an element of X.*

Formally,

$$(\exists X)\Big(\emptyset \in X \ \wedge \ \big((\forall Y \in X)\,(Y \cup \{Y\} \in X)\big)\Big).$$

In particular the set contains

$$\underbrace{\emptyset}_{0},\ \emptyset \cup \{\emptyset\} = \underbrace{\{\emptyset\}}_{1},\ \{\emptyset\} \cup \{\{\emptyset\}\} = \underbrace{\{\emptyset, \{\emptyset\}\}}_{2}, \ldots \tag{2.2}$$

so it contains all natural numbers. However, the set X may contain more elements than those listed in (2.2). More precisely, there are many infinite sets with the above property, some of them are larger than just the set of natural numbers and we define the set of natural numbers $\mathbb{N}_*$ as the intersection of all such sets. This is the smallest set with that property.

Thus the set $\mathbb{N}_*$ has the three properties:

1. $\emptyset \in \mathbb{N}_*$ (i.e. $0 \in \mathbb{N}_*$) ,
2. if $X \in \mathbb{N}_*$, then $X \cup \{X\} \in \mathbb{N}_*$ (i.e. if $n \in \mathbb{N}_*$, then $n + 1 \in \mathbb{N}_*$).

Being the smallest set with the two properties means that

3. If a set S has the above two properties, then $\mathbb{N}_* \subset S$.

It can also be formulated it as follows as the Principle of Mathematical Induction.

Theorem 2.9 (Principle of Mathematical Induction)**.** *Let S be a subset of $\mathbb{N}_*$ that has two properties*

1. $0 \in S$,

2. *For every natural number n, if $n \in S$, then $n+1 \in S$.*

Then $S = \mathbb{N}_*$.

Proof. The set S has the properties 1. and 2., so by property 3. of the set of natural numbers $\mathbb{N}_* \subset S$. On the other hand we know that $S \subset \mathbb{N}_*$ so $S = \mathbb{N}_*$. □

The Principle of Mathematical Induction has many applications and we will discuss them in Chapter 4.

Once we have the definition of the set of natural numbers it is also not difficult to define the set of integers $\mathbb{Z}$ and the set of rational numbers $\mathbb{Q}$. Rational numbers p/q can be identified with pairs of integers (p,q) such that $q \in \mathbb{N}$ and p and q are relatively prime. Then we can define real numbers. We will return to these constructions in Section 6.

The next axiom is the Axiom of Replacement. The formal statement of the Axiom of Replacement is very difficult to comprehend and we will provide only a not very rigorous version of it. It refers to functions. A formal definition of a function will be provided in Section 3.1.

Axiom of Replacement[5]**.** *Let $f : A \to B$ be a function. Then the image of the function exists. More precisely the image $f(A)$ is a subset of B defined by*

$$f(A) = \{Y \in B : \exists\, X \in A\, (Y = f(X))\}.$$

The last and the least obvious axiom is the Axiom of Choice. Although the statement seems quite natural it leads to very unexpected examples that often contradict our intuition.

Axiom of Choice. *For every set X whose elements are nonempty sets, there is a function f with domain X such that for all $Y \in X$, $f(Y) \in Y$.*

The function f is called a *choice function.* To each set $Y \in X$ the function f selects one element $f(Y) \in Y$ from Y. Then the image of the function f is a set that contains at least one element from each of the sets in X. If the sets in X are pairwise disjoint, then the image of f contains exactly one element from each of the sets in X.

There are many statements that are equivalent to the Axiom of Choice. Among the statements equivalent to the Axiom of Choice the most important are Zorn's lemma and the existence of well-ordering of sets. Another equivalent statement is that every linear space has a basis.

[5]Known also as the *Axiom Schema of Replacement*

2.4 Problems

Problem 1. Let $A = \{1, 2, 5\}$. Find the power set $P(A)$.

Problem 2. Let $A = \{2n - 1 : \; n \in \mathbb{N}\}$ and $B = \{4n : \; n \in \mathbb{N}\}$. Find $A \cap B$.

Problem 3. Prove that if $A \subset B$, then $A = B \setminus (B \setminus A)$.

Problem 4. Prove that $(A \setminus B) \setminus C = A \setminus (B \cup C)$

Problem 5. Prove that $\{\emptyset\} \neq \{\emptyset, \{\emptyset\}\}$.

Problem 6. Prove that $\{\{a\}, \{a, \{b\}\}\} = \{\{c\}, \{c, \{d\}\}\}$ if and only if $a = b$ and $c = d$.

MORE PROBLEMS WILL BE ADDED!

Chapter 3

Cartesian products and functions

The sets do not specify the order of its elements, e.g.

$$\{1, 2\} = \{2, 1\}.$$

Thus in order to take the order into consideration we need the notion of *ordered pair* (a, b). Namely we define

$$(a, b) = (c, d) \qquad \text{if and only if } a = c \text{ and } b = d.$$

In particular $(1, 2) \neq (2, 1)$.

Formally we need to be able to define the ordered pair using the set theory (that is all what we have) and a formal definition due to Kuratowski is as follows:

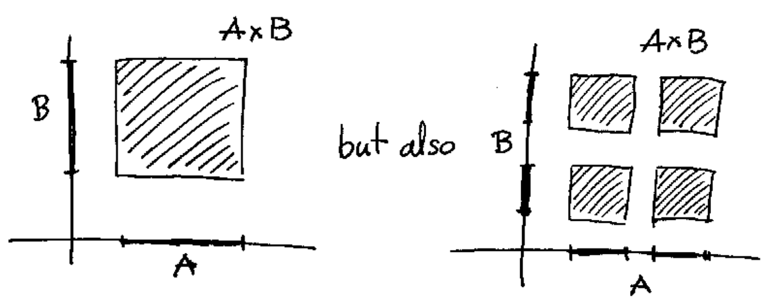
$$(a, b) := \{\{a\}, \{a, \{b\}\}\}.$$

It is an easy exercise[1] that with this definition $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.

If A and B are sets, then we define the *Cartesian product* by

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

The Cartesian product has a transparent geometric interpretation.



[1]See Problem 6 in Chapter 2.

3.1 Functions

Let X and Y be given sets. A function $f : X \to Y$ is a rule that assigns to each $x \in X$ an element of Y denoted by $f(x)$. X is called the *domain* of f and Y the *target* of f.

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

is the *range* of f or the *image* of f.

Example 3.1. If $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, then the range of f is $f(\mathbb{R}) = [0, \infty)$.

The *graph of* f is

$$G(f) = \{(x, f(x)) \in X \times Y : \, x \in X\}.$$

Observe that a graph of a function is not an arbitrary subset of $X \times Y$. The graph $G(f) \subset X \times Y$ has the property that for every $x \in X$ there is exactly one $y \in Y$ such that $(x, y) \in G(f)$. We denote *exists exactly one* by $\exists!$. Thus the graph $G(f)$ has the property[2]

$$(\forall x \in X \ \exists! y \in Y) \ \ ((x, y) \in G(f)).$$

Suppose now that we have any subset $R \subset X \times Y$ with the above property that is:

$$(\forall x \in X \ \exists! y \in Y) \ \ ((x, y) \in R). \tag{3.1}$$

Then the set R is a graph of some function $f : X \to Y$. Namely, the function f that is defined as follows. For $x \in X$ we define $f(x)$ to be the unique y in Y such that $(x, y) \in R$.

In other words, if we know the graph, we also know the function. All information about the function is contained in its graph.[3]

The definition of a function as a certain 'rule' is vague and not very rigorous so in mathematics we define a function to be any subset $R \subset X \times Y$ with the property (3.1). That is we identify a function with its graph and that is a good definition, because we can reconstruct the function if we know its graph. However, in what follows we will use a less formal definition that a function is a 'rule'.

A function $f : X \to Y$ is called *one-to-one* if[4]

$$\forall x_1, x_2 \in X \ (x_1 \neq x_2 \ \Rightarrow \ f(x_1) \neq f(x_2)).$$

If a function $f : X \to Y$ is one-to-one, then there is an inverse function

$$f^{-1} : f(X) \to X$$

defined by $f^{-1}(f(x)) = x$ for all $x \in X$.

---

[2]This is an abstract version of the vertical line test.

[3]This definition of a function leads to a funny example. If $X = Y = \emptyset$, then $X \times Y = \emptyset$ and $R = \emptyset \subset \emptyset = X \times Y$ satisfies the definition of the function as given above. Therefore $\emptyset : \emptyset \to \emptyset$ is a function! That is correct. We have to accept such examples if we want to have theory to be simple and consistent. Although we will never deal with such a function, there is no need to exclude it from the set of functions.

[4]This is an abstract version of the horizontal line test.

Example 3.2. If $f:\mathbb{R}\to\mathbb{R}$, $f(x)=e^x$, then f is one-to-one, the range if $f(\mathbb{R})=(0,\infty)$ and the inverse function is $f^{-1}:(0,\infty)\to\mathbb{R}$, $f^{-1}(x)=\ln x$.

A function is said to be *onto* or *a surjection* or *surjective* if $f(X)=Y$, i.e.,

$$\forall y\in Y\ \exists x\in X\ (y=f(x)).$$

A function $f:X\to Y$ that is one-to-one and surjective is called a *bijection*. In this situation the inverse function is defined on $f(X)=Y$ and the inverse function $f^{-1}:Y\to X$ is also a bijection. Thus a bijection is a one-to-one correspondence between all points of X and all points of Y.

Example 3.3. $f:\mathbb{R}\to\mathbb{R}$, $f(x)=x^3$ is one-to-one and onto so it is a bijection. The inverse function is $f^{-1}(x)=\sqrt[3]{x}$. It is also a bijection.

If $f:X\to Y$ and $A\subset X$, then we define

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

and we call $f(A)$ the *image of A under f.*

If $f:X\to Y$ and $A\subset Y$, then we define

$$f^{-1}(A)=\{x\in X:\ f(x)\in A\}$$

and we call this set the *preimage of A under f* or the *inverse image of A under f.* Note that

$$x\in f^{-1}(A)\quad\equiv\quad f(x)\in A.$$

The function $f:X\to X$, $f(x)=x$ for all $x\in X$ is called *identity* and is often denoted by $I_X:X\to X$.

If the image of f consists of one point, then f is called a *constant function.* That is $f:X\to Y$ is a constant function if there is $c\in Y$ such that $f(x)=c$ for all $x\in X$.

If $f:X\to Y$ and $g:Y\to Z$ are two functions, then the *composition of f and g* is the function

$$g\circ f:X\to Z$$

defined by the formula

$$(g\circ f)(x)=g(f(x)).$$

Example 3.4. If $f:\mathbb{R}\to\mathbb{R}$, $f(x)=2x+1$ and $g:\mathbb{R}\to\mathbb{R}$, $g(x)=x^2$, then $g\circ f:\mathbb{R}\to\mathbb{R}$ is given by $(g\circ f)(x)=g(f(x))=g(2x+1)=(2x+1)^2$.

If $f:X\to Y$ and $A\subset X$, then the *restriction of f to A* is

$$f|_A:A\to Y,\qquad (f|_A)(x)=f(x)\text{ for all }x\in A.$$

In this situation we say that f is an *extension* of $f|_A$.

Proposition 3.5. *Let $f : X \to Y$ be a function and let $A, B \subset Y$. Then*

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B), \tag{3.2}$$
$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$$

$$f^{-1}(Y \setminus A) = X \setminus f^{-1}(A). \tag{3.3}$$

Remark 3.6. Given a set $E \subset X$ we often write $E^c = X \setminus E$ to denote the *complement* of the set E in X. With this notation (3.3) can be stated as

$$f^{-1}(A^c) = (f^{-1}(A))^c.$$

Proof. We will only prove (3.2) leaving the proofs of the other two statements as an exercise. We need to prove that

$$x \in f^{-1}(A \cup B) \quad \text{if and only if} \quad x \in f^{-1}(A) \cup f^{-1}(B). \tag{3.4}$$

Let $x \in f^{-1}(A \cup B)$. This means that $f(x) \in A \cup B$, i.e. $f(x) \in A$ or $f(x) \in B$. If $f(x) \in A$, then $x \in f^{-1}(A) \subset f^{-1}(A) \cup f^{-1}(B)$. If $f(x) \in B$, then $x \in f^{-1}(B) \subset f^{-1}(A) \cup f^{-1}(B)$. In either case $x \in f^{-1}(A) \cup f^{-1}(B)$.

Suppose now that $x \in f^{-1}(A) \cup f^{-1}(B)$. Then $x \in f^{-1}(A)$ of $x \in f^{-1}(B)$. If $x \in f^{-1}(A)$, then $f(x) \in A \subset A \cup B$. If $x \in f^{-1}(B)$, then $f(x) \in B \subset A \cup B$. In either case $f(x) \in A \cup B$, i.e. $x \in f^{-1}(A \cup B)$. This completes the proof of the equivalence (3.4).

We could also prove the equivalence (3.4) more formally as follows

$$\begin{array}{cc} x \in f^{-1}(A \cup B) & \equiv \\ f(x) \in A \cup B & \equiv \\ (f(x) \in A) \;\vee\; (f(x) \in B) & \equiv \\ (x \in f^{-1}(A)) \;\vee\; (x \in f^{-1}(B)) & \equiv \\ x \in f^{-1}(A) \cup f^{-1}(B). & \end{array}$$

□

Proposition 3.7. *Let $f : X \to Y$ be a function and let $A, B \subset X$. Then*

$$f(A \cup B) = f(A) \cup f(B), \qquad f(A \cap B) \subset f(A) \cap f(B). \tag{3.5}$$

If in addition f is one-to-one, then

$$f(A \cap B) = f(A) \cap f(B). \tag{3.6}$$

We will only prove (3.5) leaving the proof of (3.6) as an exercise.

Proof that $f(A \cup B) = f(A) \cup f(B)$. If $y \in f(A \cup B)$, then $y = f(x)$ for some $x \in A \cup B$ and we have two cases. If $x \in A$, then $y = f(x) \in f(A) \subset f(A) \cup f(B)$. If $x \in B$, then $y = f(x) \in f(B) \subset f(A) \cup f(B)$. In either case, $y \in f(A) \cup f(B)$. We proved that

$$y \in f(A \cup B) \quad \Rightarrow \quad y \in f(A) \cup f(B),$$

i.e.,

$$f(A \cup B) \subset f(A) \cup f(B). \tag{3.7}$$

If $y \in f(A) \cup f(B)$, then $y \in f(A)$ or $y \in f(B)$ and we have two cases. If $y \in f(A)$, then $y = f(x)$ for some $x \in A \subset A \cup B$ so $y = f(x) \in f(A \cup B)$. If $y \in f(B)$, then $y = f(x)$ for some $x \in B \subset A \cup B$ so $y = f(x) \in f(A \cup B)$. We proved

$$y \in f(A) \cup f(B) \quad \Rightarrow \quad y \in f(A \cup B),$$

i.e.,

$$f(A) \cup f(B) \subset f(A \cup B). \tag{3.8}$$

Now (3.7) and (3.8) yield equality $f(A \cup B) = f(A) \cup f(B)$. □

Proof that $f(A \cap B) \subset f(A) \cap f(B)$. If $y \in f(A \cap B)$, then $y = f(x)$ for some $x \in A \cap B$. Hence $x \in A$ and $x \in B$ at the same time so $y = f(x) \in f(A)$ and $y = f(x) \in f(B)$ at the same time and hence $f \in f(A) \cap f(B)$. We proved that

$$y \in f(A \cap B) \quad \Rightarrow \quad y \in f(A) \cap f(B),$$

i.e. $f(A \cap B) \subset f(A) \cap f(B)$. □

Remark 3.8. In general we do not have equality in $f(A \cap B) \subset f(A) \cap f(B)$. That is, in general it is *not true* that $f(A) \cap f(B) \subset f(A \cap B)$. This is somewhat surprising since we proved that $f(A \cup B) = f(A) \cup f(B)$ and one could expect a similar equality for the intersection of sets.

Let us try to prove that $f(A) \cap f(B) \subset f(A \cap B)$ to see where the proof falls apart.

Let $y \in f(A) \cap f(B)$. Then $y \in f(A)$ and $y \in f(B)$ at the same time so

$$y = f(x) \quad \text{for some } x \in A \quad \text{and} \quad y = f(x) \quad \text{for some } x \in B.$$

Therefore $y = f(x)$ for some $x \in A \cap B$ so $y \in f(A \cap B)$ and we proved that

$$y \in f(A) \cap f(B) \quad \Rightarrow \quad y \in f(A \cap B),$$

i.e.,

$$f(A) \cap f(B) \subset f(A \cap B).$$

The proof is not correct! But where did we make a mistake?

We proved (and that is correct) that

$$y = f(x) \quad \text{for some } x \in A, \tag{3.9}$$

$$y = f(x) \quad \text{for some } x \in B, \tag{3.10}$$

and we concluded that $y = f(x)$ for some $x \in A \cap B$. This conclusion is however, not correct since $x \in A$ for which (3.9) is true need not be the same as $x \in B$ for which (3.10) is true and so we cannot claim that there is a common $x \in A \cap B$ for which $y = f(x)$!

3.2 More logic*

We will prove now (3.5) using a more formal logical notation. In practice we do not do this and we write proofs as those written above, but the point is to learn what logical structure is hidden behind the above proofs. First we have to learn some more logic.

Assume that $P(x)$ and $Q(x)$ are some statements. Then we have

$$\exists x\ (P(x) \vee Q(x)) \quad \equiv \quad (\exists x\ P(x)) \vee (\exists x\ Q(x)).$$

This is correct. $\exists x\ (P(x) \vee Q(x))$ is true, when there is x such that $P(x)$ is true or $Q(x)$ which happens exactly when there is x such that $P(x)$ is true or there is x such that $Q(x)$ is true i.e., $(\exists x\ P(x)) \vee (\exists x\ Q(x))$. Thus the two statements are equivalent.

$$\forall x\ (P(x) \wedge Q(x)) \quad \equiv \quad (\forall x\ P(x)) \wedge (\forall x\ Q(x)).$$

This is also true. For every x both statements $P(x)$ and $Q(x)$ are true exactly when for all x, $P(x)$ is true and for all x, $Q(x)$ is true.

However in the following statements we only have implication in one direction

$$\exists x\ (P(x) \wedge Q(x)) \quad \Rightarrow \not\Leftarrow \quad (\exists x\ P(x)) \wedge (\exists x\ Q(x)).$$

Indeed, if $\exists x\ (P(x) \wedge Q(x))$ is true, then we can find x such that $P(x)$ and $Q(x)$ are true at the same time. However, if the statement $(\exists x\ P(x)) \wedge (\exists x\ Q(x))$ is true, then $P(x)$ and $Q(x)$ can be true for different values of x. Here is an example. We assume that $x \in \mathbb{R}$.

$$\underbrace{\exists x\ (x^2 \geq 1 \wedge x^2 \leq 1/2)}_{\text{false}} \quad \Rightarrow \not\Leftarrow \quad \underbrace{\underbrace{(\exists x\ x^2 \geq 1)}_{\text{true}} \wedge \underbrace{(\exists x\ x^2 \leq 1/2)}_{\text{true}}}_{\text{true}}.$$

A false statement implies a true statement, but a true statement does not imply a false one so we only have an implication from left to right. Similarly we have

$$\forall x(P(x) \vee Q(x)) \quad \not\Rightarrow \Leftarrow \quad (\forall x\ P(x)) \vee (\forall x\ Q(x))$$

Here is an example

$$\underbrace{\forall x\ (x^2 \geq 1 \vee x^2 \leq 1)}_{\text{true}} \quad \not\Rightarrow \Leftarrow \quad \underbrace{(\forall x\ x^2 \geq 1) \vee (\forall x\ x^2 \leq 1)}_{\text{false}}.$$

Now we are ready to provide a formal proof of Proposition 3.7. We will only prove (3.5).

Proof that $f(A \cup B) = f(A) \cup f(B)$.

$$\begin{array}{cc}
y \in f(A \cup B) & \equiv \\
\exists x\ ((x \in A \cup B) \wedge (y = f(x)) & \equiv \\
\exists x\ ((x \in A \vee x \in B) \wedge (y = f(x)) & \equiv \\
\exists x\ (x \in A \wedge y = f(x)) \vee (x \in B \wedge y = f(x)) & \equiv \\
(\exists x\ (x \in A \wedge y = f(x))) \vee (\exists x\ (x \in B \wedge y = f(x))) & \equiv \\
y \in f(A) \vee\ y \in f(B) & \equiv \\
y \in f(A) \cup f(B)\ \ . &
\end{array}$$

□

Proof that $f(A \cap B) \subset f(A) \cap f(B)$.

$$\begin{array}{cc}
y \in f(A \cap B) & \equiv \\
\exists x \; ((x \in A \cap B) \wedge (y = f(x)) & \equiv \\
\exists x \; (x \in A \wedge x \in B \wedge y = f(x)) & \equiv \\
\exists x \; ((x \in A \wedge y = f(x)) \wedge (x \in B \wedge y = f(x)) & \Rightarrow \not\Leftarrow \\
(\exists x \; (x \in A \wedge y = f(x))) \wedge (\exists x \; (x \in B \wedge y = f(x))) & \equiv \\
y \in f(A) \wedge \; y \in f(B) & \equiv \\
y \in f(A) \cap f(B) \; . &
\end{array}$$

All statements but one are equivalent and in one case we only have the implication $\Rightarrow$ so we proved that $y \in f(A \cap B) \;\Rightarrow\; y \in f(A) \cap f(B)$ and hence $f(A \cap B) \subset f(A) \cap f(B)$. □

3.3 Problems

Problem 1. Prove Proposition 3.7.

Problem 2. Prove that if $f : X \to Y$ is a function and $A_1, A_2, A_3, \ldots$ are subsets of X, then

$$f\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} f(A_i)\,,$$

and

$$f\left(\bigcap_{i=1}^{\infty} A_i\right) \subset \bigcap_{i=1}^{\infty} f(A_i). \tag{3.11}$$

Provide an example to show that we do not necessarily have equality in (3.11)

Problem 3. Prove that if $f : X \to Y$ is one-to-one and $A_1, A_2, A_3, \ldots$ are subsets of X, then

$$f\left(\bigcap_{i=1}^{\infty} A_i\right) = \bigcap_{i=1}^{\infty} f(A_i)\,.$$

MORE PROBLEMS WILL BE ADDED!

Chapter 4

Mathematical induction

In this section we will assume familiarity with real numbers although a formal set of axioms and a formal construction of real numbers will be presented later. While it might look like a lack of consistency, mathematical induction is so important and straightforward that postponing it for later would be a mistake.

4.1 Patterns

Consider a well known formula

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}\,.$$

How can we check that it is true? We can check it for $n = 1, 2, 3, \ldots, 100$ on a calculator and we see that both sides are equal. We observe a pattern. We often guess general formulas from a pattern. Since the formula is true for all $n = 1, 2, 3, \ldots, 100$ it seems obvious that it must be true for all n. Is it a convincing argument?

Here is another example. Consider the sequence 2^n, $n = 0, 1, 2, 3, \ldots$ and look at the first digits in the decimal representation of 2^n. The sequence 2^n starts with $\mathbf{1}, \mathbf{2}, \mathbf{4}, \mathbf{8}, \mathbf{1}6, \mathbf{3}2, \mathbf{6}4, \mathbf{1}28, \mathbf{2}56, \mathbf{5}12 \ldots$ If we keep playing with the calculator we see that the first digits in the sequence are

$$\begin{aligned}
&1,\ 2,\ 4,\ 8,\ 1,\ 3,\ 6,\ 1,\ 2,\ 5,\\
&1,\ 2,\ 4,\ 8,\ 1,\ 3,\ 6,\ 1,\ 2,\ 5,\\
&1,\ 2,\ 4,\ 8,\ 1,\ 3,\ 6,\ 1,\ 2,\ 5,\\
&1,\ 2,\ 4,\ 8,\ 1,\ 3,\ 6,\ 1,\ 2,\ 5,\\
&1,\ 2,\ 4,\ 8,\ 1,\ 3, \ldots
\end{aligned}$$

The above sequence shows the first digits of 2^n for $n = 0, 1, 2, 3, \ldots, 45$. We can easily see a pattern that keeps repeating. The sequence of digits $1, 2, 4, 8, 1, 3, 6, 1, 2, 5$ appears with a periodic regularity. In particular digits 7 and 9 do not appear in the sequence. That seems like a general rule. Can we state it as a theorem? We seem to have enough evidence. However, if we would add one more term to the sequence we would see that $2^{46} = 7.036874418\ldots \times 10^{13}$. Moreover $2^{53} = 9.007199255\ldots \times 10^{15}$. One can prove even a more surprising result that eventually the digit 7 will appear more often[1] than the digit 8!

Let us consider now another, more complicated, example.[2]

Example 4.1. Let $a_1 = 14, a_2 = 128, a_3 = 1170, a_4 = 10695, a_5 = 97763$, and let the following elements in the sequence be defined by the recursive function

$$a_{n+5} = 10a_{n+4} - 8a_{n+3} + a_{n+2} + 3a_{n+1} + 2a_n \quad \text{for } n \geq 1.$$

This sequence is defined in a similar way as the famous Fibonacci sequence. Consider also another sequence $b_1 = 14, b_2 = 128$ and

$$b_{n+2} = \left[\frac{b_{n+1}^2}{b_n} + \frac{1}{2}\right], \quad \text{for } n \geq 1.$$

Here $[x]$ denotes the integer part of x, i.e., the largest integer less than or equal to x. Are the two sequences a_n and b_n equal for all n? The formulas are very different, but if we experiment with a calculator or a computer, we will see that they agree for $n = 1, 2, \ldots, 100$. We see a pattern, so two sequences seem equal. We are not sure, so we keep checking for larger $n = 101, 102, \ldots, 5000$. Still works. Now we can be sure that they must be equal for all n. Really? If we would be more patient and have checked the formula for still larger n we would see that $a_n \neq b_n$ for $n = 5016$. That is really surprising! Isn't it?

Example 4.2. Let $a_0 = a_1 = 1$ and

$$a_{n+1} = \frac{a_0^{97} + a_1^{97} + a_2^{97} + \ldots + a_n^{97}}{n} \quad \text{for } n = 1, 2, 3, \ldots$$

One can prove that the numbers a_n are integers for $n \leq 2039$, but a_{2040} is not an integer.

4.2 Induction

If we observe a pattern that suggest a formula for a given sequence, it can never be regarded as a proof and we need a rigorous mathematical argument that will allow us to verify the formula for all n. A method of proving such statements is based on the Principle of Mathematical Induction (Theorem 2.9). Let's recall it again.[3]

Theorem 4.3 (Principle of Mathematical Induction)**.** *Let S be a subset of $\mathbb{N}$ that has two properties*

[1] The reason why the sequence of 10 digits seems to appear periodically has a rational explanation. $2^{10} = 1024$ is very close to 1000. Thus multiplying a given number 10 times by 2 it is almost like multiplying it by 1000, and multiplication by 1000 just adds three zeros at the end — it does not change the digits at the beginning. However, we do not multiply by 1000, but by 1024 and 24 creates an error which, like a tumor, grows so much that at the time of taking 2^{46} the error changes the anticipated first digit 6 to 7. Tumor kills the patient.

[2] This example and the next one are due to J. Wróblewski.

[3] While in Theorem 2.9 we started induction form $n = 0$, here we start the induction from $n = 1$.

1. $1 \in S$,

2. *For every natural number n, if $n \in S$, then $n+1 \in S$.*

Then $S = \mathbb{N}$.

The above theorem leads to a beautiful and powerful method of proving that a statement $P(n)$ about a natural number n is true for all $n \in \mathbb{N}$. The method is called *Mathematical Induction*. We formulate the method as a theorem.

Theorem 4.4 (Principle of Mathematical Induction)**.** *Let for each $n \in \mathbb{N}$, $P(n)$ be a statement about a natural number n. Suppose also that*

1. *$P(1)$ is true,*

2. *For every $n \in \mathbb{N}$, if $P(n)$ is true, then $P(n+1)$ is true.*

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof. If $S = \{n \in \mathbb{N} : P(n) \text{ is true}\}$, then S satisfies the assumptions of from Theorem 4.3 and hence $S = \mathbb{N}$. □

The assumption "if $P(n)$ is true" in **2.** is called the *induction hypothesis*.

Let us check how the theorem works in practice. As a first application of Mathematical Induction we will prove

Example 4.5. *Prove that for every $n \in \mathbb{N}$, $2^{6n+1} + 3^{2n+2}$ is divisible by 11.*

Proof.

1. For $n = 1$ we have
$$2^{6n+1} + 3^{2n+2} = 2^7 + 3^4 = 209 = 11 \cdot 19.$$

2. Suppose now $11 | 2^{6n+1} + 3^{2n+2}$. We need to prove that $11 | 2^{6(n+1)+1} + 3^{2(n+1)+2}$. By the assumption $2^{6n+1} + 3^{2n+2} = 11k$ for some $k \in \mathbb{N}$. We have

$$\begin{aligned} & 2^{6(n+1)+1} + 3^{2(n+1)+2} = 2^{6n+1} \cdot 2^6 + 3^{2n+2} \cdot 3^2 = 64 \cdot 2^{6n+1} + 9 \cdot 3^{2n+2} \\ = \quad & 64(2^{6n+1} + 3^{2n+2}) - 55 \cdot 3^{2n+2} = 64 \cdot 11k - 5 \cdot 11 \cdot 3^{2n+2} = 11(64k - 5 \cdot 3^{2n+2})\,. \end{aligned}$$

The proof is complete. □

Proposition 4.6 (Bernoulli's inequality)**.** *If n is a positive integer and $a \geq -1$, a real number, then*
$$(a+1)^n \geq 1 + na.$$

Proof.

1. For $n = 1$ the inequality is obvious.

2. Suppose that $(1+a)^n \geq 1+na$. We need to prove that $(1+a)^{n+1} \geq 1+(n+1)a$. We have

$$(1+a)^{n+1} = (1+a)^n(1+a) \geq (1+na)(1+a) = 1+(n+1)a+na^2 \geq 1+(n+1)a\,.$$

In the proof of the first inequality we used the fact that $(1+a)^n \geq 1+na$ and $1+a \geq 0$. The proof is complete. □

Example 4.7. *Prove that for $n \geq 1$*

$$\frac{1}{n+1}+\frac{1}{n+2}+\frac{1}{n+3}+\ldots+\frac{1}{3n}+\frac{1}{3n+1} > 1\,.$$

Proof.

1. For $n=1$ we have

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{4} = \frac{6+4+3}{12} = \frac{13}{12} > 1\,.$$

2. Suppose the inequality is true for n, we have to prove it for $n+1$.

$$\begin{aligned}
&\frac{1}{n+2}+\frac{1}{n+3}+\ldots+\frac{1}{3(n+1)+1} \\
&= \frac{1}{n+2}+\frac{1}{n+3}+\ldots+\frac{1}{3n+1}+\frac{1}{3n+2}+\frac{1}{3n+3}+\frac{1}{3n+4} \\
&= \left(\frac{1}{n+1}+\frac{1}{n+2}+\ldots+\frac{1}{3n+1}\right)-\frac{1}{n+1}+\frac{1}{3n+2}+\frac{1}{3n+3}+\frac{1}{3n+4}\,.
\end{aligned}$$

Now it suffices to prove that

$$-\frac{1}{n+1}+\frac{1}{3n+2}+\frac{1}{3n+3}+\frac{1}{3n+4} > 0\,.$$

We have

$$\frac{1}{3n+2}+\frac{1}{3n+3}+\frac{1}{3n+4} > \frac{1}{n+1} = \frac{3}{3n+3} \qquad \equiv$$

$$\frac{1}{3n+2}+\frac{1}{3n+4} > \frac{2}{3n+3} \qquad \equiv$$

$$\frac{3n+4+3n+2}{(3n+2)(3n+4)} > \frac{2}{3n+3} \qquad \equiv$$

$$(6n+6)(3n+3) > 2(3n+2)(3n+4) \qquad \equiv$$

$$(3n+3)^2 > (3n+2)(3n+4) \qquad \equiv$$

$$9n^2+18n+9 > 9n^2+18n+8 \qquad \equiv$$

$$9 > 8.$$

Since the last statement is true and all statements above are equivalent, the inequality we wanted to prove is true. The proof is complete. □

Recall that n *factorial* is defined by $n! = 1 \cdot 2 \cdot \ldots \cdot n$ and $0! = 1$. We also define *binomial coefficients*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdot \ldots \cdot (n-k+1)}{k!}\,. \tag{4.1}$$

Theorem 4.8 (Binomial theorem)**.** *For $a, b \in \mathbb{R}$ and $n \in \mathbb{N}$ we have*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = a^n + \frac{n}{1!}\, a^{n-1} b + \frac{n(n-1)}{2!}\, a^{n-2} b^2 + \ldots + \frac{n(n-1)\cdots 2}{(n-1)!}\, a b^{n-1} + b^n.$$

Remark 4.9. The first three and the last three terms of the above formula are

$$a^n + n a^{n-1} b + \frac{n(n-1)}{2}\, a^{n-2} b^2 + \ldots + \frac{n(n-1)}{2}\, a^2 b^{n-2} + n a b^{n-1} + b^n$$

Also the coefficients in the binomial formula are symmetric in the sense that the coefficient at $a^{n-k}b^k$ is equal to the coefficient at $a^k b^{n-k}$, because one can easily prove that

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

Proof. For $n = 1$ the equality is obvious. Suppose it is true for n and we will prove it for $n+1$. One can easily prove (see problem 1) that

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

By the induction hypothesis we assume that binomial formula is true for n. Multiplying it by $a+b$ yields

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

Observe that

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

Hence

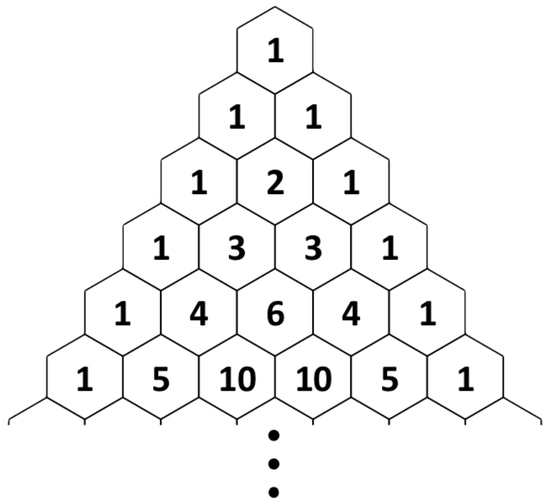
$$\begin{aligned}
(a+b)^{n+1} &= a^{n+1} + \sum_{k=1}^{n} \binom{n}{k} a^{n+1-k} b^k \\
&+ b^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} a^{n+1-k} b^k \\
&= a^{n+1} + \left(\sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1} \right] a^{n+1-k} b^k \right) + b^{n+1} \\
&= a^{n+1} + \left(\sum_{k=1}^{n} \binom{n+1}{k} a^{n+1-k} b^k \right) + b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k\,.
\end{aligned}$$

The proof is complete. □

You need to memorize two cases:

$$(a \pm b)^2 = a^2 \pm 2ab + b^2 \quad \text{and} \quad (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

For the cases with higher exponents the easiest way to obtain formula is to use the *Pascal triangle*:



Then we have

$(a+b)^n$	Expansion	Pascal's triangle
$(a+b)^0$	1	1
$(a+b)^1$	$1a + 1b$	1,1
$(a+b)^2$	$1a^2 + 2ab + 1b^2$	1,2,1
$(a+b)^3$	$1a^3 + 3a^2b + 3ab^2 + 1b^3$	1,3,3,1
$(a+b)^4$	$1a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + 1b^4$	1,4,6,4,1
$(a+b)^5$	$1a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + 1b^5$	1,5,10,10,5,1

Theorem 4.10. *If n is a natural number and $a, b \in \mathbb{R}$, then*

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \ldots + ab^{n-2} + b^{n-1}).$$

Proof. While one can prove the equality using mathematical induction, here is a shorter argument

$$\begin{aligned}
&(a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \ldots + ab^{n-2} + b^{n-1})\\
&= a(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \ldots + ab^{n-2} + b^{n-1})\\
&\quad - b(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \ldots + ab^{n-2} + b^{n-1})\\
&= (a^n + a^{n-1}b + a^{n-2}b^2 + \ldots + a^2b^{n-2} + ab^{n-1})\\
&\quad - (a^{n-1}b + a^{n-2}b^2 + a^{n-3}b^3 + \ldots + ab^{n-1} + b^n)\\
&= a^n - b^n.
\end{aligned}$$

The last equality follows from the fact that most of the terms will cancel out. □

Theorem 4.11 (Cauchy-Schwarz inequality)**.** *For $n \in \mathbb{N}$ and real numbers a_i, b_i, $i = 1, 2, \ldots, n$ we have*

$$\left|\sum_{i=1}^{n} a_i\, b_i\right| \le \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \left(\sum_{i=1}^{n} b_i^2\right)^{1/2}.$$

Proof. While it is possible to prove the inequality using mathematical induction (Problem 6) we will present a much more clever and a shorter proof.

For any real number t we have

$$0 \le \sum_{i=1}^{n} (a_i + t b_i)^2 = \sum_{i=1}^{n} a_i^2 + 2t \sum_{i=1}^{n} a_i b_i + t^2 \sum_{i=1}^{n} b_i^2.$$

That means the quadratic function

$$t \mapsto \left(\sum_{i=1}^{n} b_i^2\right) t^2 + \left(2 \sum_{i=1}^{n} a_i b_i\right) t + \left(\sum_{i=1}^{n} a_i^2\right)$$

is always non-negative so it can have at more one root and therefore its discriminant is less than or equal to zero, that is

$$\left(2 \sum_{i=1}^{n} a_i b_i\right)^2 - 4 \left(\sum_{i=1}^{n} b_i^2\right)\left(\sum_{i=1}^{n} a_i^2\right) \le 0,$$

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \le \left(\sum_{i=1}^{n} b_i^2\right)\left(\sum_{i=1}^{n} a_i^2\right)$$

and the Cauchy-Schwarz inequality follows upon taking the square root.[4] □

Example 4.12. *"Prove" that for all $n \in \mathbb{N}$ the inequality $30n < 2^n + 110$ is true.*

"Proof". For $n = 1$ we have $30 < 2 + 110$ which is true. Suppose the inequality is true for given n and we need to prove it for $n + 1$. We have

$$30(n+1) = 30n + 30 < 2^n + 110 + 30 = 2^{n+1} + 110 + 30 - 2^n < 2^{n+1} + 110,$$

and clearly the last inequality is true for $n \ge 5$. Therefore it remains to prove the inequality $30n < 2^n + 110$ for $n = 2, 3, 4$ and we do it by a direct computation. For $n = 2$ we have $60 < 4 + 110$, For $n = 3$ we have $90 < 8 + 110$ and for $n = 4$ we have $120 < 16 + 110$. "□"

In particular for $n = 6$ we obtain an amazing inequality $180 < 174$. The problem is that we proved the inequality for $n = 1, 2, 3, 4$ and then we proved that if it is true for $n \ge 5$, then it is also true for $n + 1$. We did not however, checked the inequality for $n = 5$ and the inequality is false for $n = 5$. Be careful!

[4] Remember that $\sqrt{x^2} = |x|$.

There are many modifications of the method of induction, for example in many situations we need to start with $n = n_0$ rather than $n = 1$. For example if we want to prove $P(n)$ for all $n \geq 3$ it suffices to prove $P(3)$ and that $P(n)$ implies $P(n+1)$ for $n \geq 3$.

As an application of the Principle of Mathematics Induction we will prove the following important result.

Theorem 4.13 (Well-Ordering Principle)**.** *Every nonempty subset $S \subset \mathbb{N}$ has the smallest element, i.e., there is $n_0 \in S$ such that $n_0 \leq n$ for all $n \in S$.*

Proof. Suppose to the contrary that $\emptyset \neq S \subset \mathbb{N}$ does not have the smallest element. Let $X = \mathbb{N} \setminus S$ and let

$$Y = \{n \in X : \forall m \leq n \ (m \in X)\} = \{n \in X : 1, 2, 3, \ldots, n \in X\}.$$

That is, the set Y consists of all natural numbers $n \in X$ such that none of the elements $1, 2, \ldots, n$ belong to S. Clearly $Y \subset X = \mathbb{N} \setminus S$. We will use the Principle of Mathematical Induction to prove that $Y = \mathbb{N}$ and hence $S = \emptyset$, so we will have a contradiction. Clearly $1 \in X$ as otherwise $1 \in S$ would be the smallest element. Hence also $1 \in Y$. Suppose now that $n \in Y$. That means all natural numbers $m \leq n$ are not in S. Thus $n+1$ cannot belong to S as otherwise it would be the smallest element in S, so $n+1 \in X$ and since $n+1$ and all smaller numbers than $n+1$ are also in X we have that $n+1 \in Y$. It follows now from the Principle of Mathematical Induction (Theorem 4.3) that $Y = \mathbb{N}$ which is a contradiction as pointed above. □

Sometimes in order to prove $P(n+1)$ it is not enough to use $P(n)$ but we also need to use $P(k)$ for all $k \leq n$. In this situation the following modification of the method of induction applies.

Theorem 4.14 (Principle of Strong Induction)**.** *Let for each $n \in \mathbb{N}$, $P(n)$ be a statement about a natural number n. Suppose also that*

1. *$P(1)$ is true,*

2. *For every $n \in \mathbb{N}$, if $P(k)$ is true, for $k = 1, 2, \ldots n$, then $P(n+1)$ is true.*

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Again this theorem seems obvious and we will not prove it.

Theorem 4.15 (Binary representation)**.** *Every non-negative integer can be represented as*

$$n = c_k 2^k + c_{k-1} 2^{k-1} + \ldots + c_0 2^0,$$

where $c_i \in \{0, 1\}$. It is called a binary representation*.*

Proof. For $n = 0$ we have $0 = 0 \cdot 2^0$. Let $n \geq 0$ and assume that all integers $0, 1, 2, \ldots, n$ have binary representations (Principle of Strong Induction). We need to prove that $n+1$ also has it.

If $n+1$ is even, $n \geq 1$ and hence $(n+1)/2 \leq n$ has a binary representation

$$\frac{n+1}{2} = c_k 2^k + c_{k-1} 2^{k-1} + \ldots + c_0 2^0.$$

Accordingly

$$n+1 = c_k 2^{k+1} + c_{k-1} 2^k + \ldots + c_0 2^1 + 0 \cdot 2^0.$$

is a binary representation of $n+1$. If $n+1$ is odd, then $n/2 \geq 0$ has a binary representation

$$\frac{n}{2} = c_k 2^k + c_{k-1} 2^{k-1} + \ldots + c_0 2^0$$

and hence

$$n+1 = c_k 2^{k+1} + c_{k-1} 2^k + \ldots + c_0 2^1 + 1 \cdot 2^0$$

is a binary representation of $n+1$. □

A quite unusual modification of the method of induction will be used in the proof of the following result (no cheating this time!).

Theorem 4.16 (Arithmetic-Geometric Mean Inequality)**.** *If $a_1, a_2, a_3, \ldots, a_n \geq 0$, then*

$$\sqrt[n]{a_1 \cdot a_2 \cdot \ldots \cdot a_n} \leq \frac{a_1 + a_2 + \ldots + a_n}{n}$$

and the equality holds if and only if $a_1 = a_2 = \ldots = a_n$.

Proof. We will prove the inequality only, but the reader may conclude from the proof that the equality holds if and only if $a_1 = a_2 = \ldots = a_n$. We leave this last conclusion as an exercise.

First we will prove the inequality for $n = 2^k$, $k = 1, 2, 3, \ldots$, i.e., we will prove it for $n = 2, 4, 8, 16, \ldots$

1. For $n = 2^1$ we have

$$\frac{a_1 + a_2}{2} \geq \sqrt{a_1 a_2} \quad \equiv \quad a_1 - 2\sqrt{a_1 a_2} + a_2 \quad \equiv \quad (\sqrt{a_1} - \sqrt{a_2})^2 \geq 0\,.$$

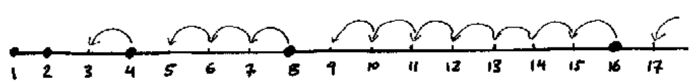
Since the last condition is obviously true, the inequality, as equivalent to the last statement, is also true.

2. Suppose that the inequality is true for $n = 2^k$. We need to prove it for $n = 2^{k+1}$. We have

$$
\begin{aligned}
& \sqrt[2^{k+1}]{a_1 \cdot \ldots \cdot a_{2^k} \cdot a_{2^k+1} \cdot \ldots \cdot a_{2^{k+1}}} \\
= & \sqrt{\sqrt[2^k]{a_1 \cdot \ldots \cdot a_{2^k}}\ \sqrt[2^k]{a_{2^k+1} \cdot \ldots \cdot a_{2^{k+1}}}} \\
\leq & \frac{\sqrt[2^k]{a_1 \cdot \ldots \cdot a_{2^k}} + \sqrt[2^k]{a_{2^k+1} \cdot \ldots \cdot a_{2^{k+1}}}}{2} \\
\leq & \frac{\frac{a_1+\ldots+a_{2^k}}{2^k} + \frac{a_{2^k+1}+\ldots a_{2^{k+1}}}{2^k}}{2} \\
= & \frac{a_1 + a_2 + \ldots + a_{2^{k+1}}}{2^{k+1}}.
\end{aligned}
$$

The above estimates require some explanations. The first equality is obvious. The second inequality is just a consequence of the arithmetic-geometric inequality for $n = 1$ which was proved in **1.** The third inequality follows from the inductive assumption that the inequality is true for $n = 2^k$ and the last equality is obvious again.

We proved the inequality for $n = 2, 4, 8, 16, \ldots$. In order to prove that the inequality is true for all integers it suffices to prove that if it is true for n, then it is also true for $n - 1$ (reverse induction).



Thus suppose that the inequality is true for n. We will prove it is true for $n-1$. We have

$$
\begin{aligned}
& \sqrt[n]{a_1 \cdot \ldots \cdot a_{n-1} \cdot \left(\frac{a_1 + \ldots + a_{n-1}}{n-1}\right)} \\
\leq & \frac{a_1 + \ldots a_{n-1} + \left(\frac{a_1+\ldots+a_{n-1}}{n-1}\right)}{n} \\
= & \frac{a_1 + \ldots + a_{n-1}}{n-1}.
\end{aligned}
$$

The first inequality above follows from the assumption that the arithmetic-geometric inequality is true for n. Hence

$$\sqrt[n]{a_1 \cdot \ldots \cdot a_{n-1}}\ \sqrt[n]{\frac{a_1 + \ldots + a_{n-1}}{n-1}} \leq \frac{a_1 + \ldots + a_{n-1}}{n-1},$$

so

$$\left(a_1 \cdot \ldots \cdot a_{n-1}\right)^{1/n} \leq \left(\frac{a_1 + \ldots a_{n-1}}{n-1}\right)^{1-1/n}$$

and finally

$$\left(a_1 \cdot \ldots \cdot a_{n-1}\right)^{\frac{1}{n-1}} \leq \frac{a_1 + \ldots + a_{n-1}}{n-1}$$

which is what we wanted to prove. □

4.3 Problems

Problem 1. Prove that if $0 \leq k \leq n$ are integers, then

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

Problem 2. Prove that

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}.$$

Problem 3. Prove that for $q \neq 1$ we have

$$1 + q + q^2 + \ldots + q^n = \frac{1 - q^{n+1}}{1 - q}.$$

Problem 4. Prove using induction that if n is a natural number and $a, b \in \mathbb{R}$, then

$$a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \ldots + ab^{n-2} + b^{n-1}).$$

Problem 5. Prove that the number of diagonals in a convex polygon with n sides equals $n(n-3)/2$.

Problem 6. Use mathematical induction to prove the Cauchy-Schwarz inequality

$$\left|\sum_{i=1}^{n} a_i\, b_i\right| \leq \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \left(\sum_{i=1}^{n} b_i^2\right)^{1/2}.$$

Problem 7. Examine the sequence of first digits in the decimal expansion of 1.778^n, $n = 1, 2, 3, \ldots$ We see a periodic sequence $1, 3, 5, 9$. The sequence will eventually change. Using calculator find the smallest value of n for which the sequence changes. Why does the sequence of digits $1, 3, 5, 9$ appear periodically for such a long time? Can you modify the sequence so that the digits $1, 3, 5, 9$ will appear periodically even more times?

Chapter 5

Relations, integers and rational numbers

In Section 2.3 we learned how to define natural numbers $\mathbb{N}_* = \{0, 1, 2, \ldots\}$ directly from the axioms of the set theory. Those who did not follow this construction carefully, may assume that somehow we know what the natural numbers $\mathbb{N}_*$ are and that they are equipped with the operations of addition, multiplication and the order $n \leq m$. We also assume familiarity with basic properties of arithmetic like

$$n + m = m + n, \quad n(m + k) = nm + nk, \quad \text{etc.}$$

While all these properties can be proved using the definitions, we take them for granted. The main point of this chapter is to learn how to define all integers $\mathbb{Z}$ and all rational numbers $\mathbb{Q}$. In the next chapter we will learn how to define all real numbers $\mathbb{R}$.

We are born with a good understanding of what natural numbers are, however, existence of negative integers, rational numbers or real numbers is not so natural and it took humanity thousands of years to understand what they are and how they should be constructed. Leopold Kronecker (1823–1891) who worked on number theory, algebra and logic once said: *God created the natural numbers; all the rest is the work of man.*

However, before we can proceed to the construction of $\mathbb{Z}$ or $\mathbb{Q}$ we need to learn about *binary relations*.

5.1 Binary relations

Before we provide a general definition of a *binary relation*, called *relation* for short, let us look at some examples of relations:

1. $A \subset B$, where A and B are sets,
2. $A = B$, where A and B are sets,

3. $x \leq y$, where x, y are real numbers,

4. $n < m$, where n, m are integers,

5. *Congruence relation* on the set of integers. Given $n \in \mathbb{N}$, $n \geq 2$, two integers $a, b \in \mathbb{Z}$ are *congruent modulo* n, denoted

$$a \equiv b \pmod{n},$$

   if $a - b$ is divisible by n, $n | a - b$,

6. $\Delta ABC \cong \Delta A'B'C'$, relation of congruence of triangles.

More generally, if X is a set, and $\sim$ is a relation between elements of X, then we write $x \sim y$ if the elements x, y are in this relation. However, we still need a general and rigorous definition of a relation.

A (*binary*) *relation* on X is *any* subset $R \subset X \times X$ of the Cartesian product. Then we write

$$a \sim b \quad \text{if and only if} \quad (a, b) \in R.$$

Let us look at some examples.

Example 5.1. The real numbers satisfy the relation $x \leq y$ if and only if (x, y) belongs to the set:

Example 5.2. The natural numbers satisfy the relation $n = m$ is and only if (n, m) belongs to the set:

We say that a relation $\sim$ on a set X is

1. *Reflexive*, if $a \sim a$ for all $a \in X$,

2. *Symmetric*, if for all $a, b \in X$, $a \sim b$ implies $b \sim a$,

3. *Transitive*, if for all $a, b, c \in X$, if $a \sim b$ and $b \sim c$, then also $a \sim c$,

4. *Equivalence relation*, if it is reflexive, symmetric and transitive.

1. The relation $A \subset B$, where A and B are sets is reflexive and transitive, but not symmetric.

2. The relation $A = B$, where A and B are sets is an equivalence relation.

3. The relation $x \leq y$, where x, y are real numbers is reflexive and transitive, but not symmetric.

4. The relation $n < m$, where n, m are integers is transitive, but neither reflexive nor symmetric.

5. The congruence relation between integers is an equivalence relation. Indeed, it is reflexive, $a \equiv a \pmod{n}$, because $a - a = 0$ is divisible by n. It is symmetric because if $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$ (if $a - b$ is divisible by n, then obviously $b - a$ is divisible by n). Finally, the relation is transitive. If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a - b = kn$ and $b - c = \ell n$ so $a - c = (a - b) + (b - c) = (k + \ell)n$ and hence $a \equiv c \pmod{n}$.

6. The relation of congruence of triangles is an equivalence relation.

Formally, there is a 'small' problem with the example of sets. We defined a relation to be a relation between elements of some set X. However, the relation $A \subset B$ between any two sets should be defined on the set of all sets, but there is no set of all sets (Proposition 2.8)! Still we can think of the collection of all sets. Why not? While the collection of all sets is not a set, we call it a *category of sets* so the relation $A \subset B$ is defined in the category of all sets. We use the world 'category' to emphasize that this is not a set.

Let $\sim$ be an equivalence relation in a set X. For any $a \in X$ we define the *equivalence class of a* as

$$[a] = \{b \in X : b \sim a\}.$$

Clearly $a \in [a]$ (because the relation $\sim$ is reflexive) so $[a]$ is non-empty.

Example 5.3. The relation $x = y$ between real numbers is an equivalence relation and $[x] = \{x\}$ for very $x \in \mathbb{R}$.

Example 5.4. Consider the congruence relation $a \equiv b \pmod{5}$ in $\mathbb{Z}$. We have that $b \equiv a \pmod{5}$ if $b - a = 5k$ or $b = a + 5k$ for some $k \in \mathbb{Z}$. Therefore $[a] = \{a + 5k : k \in \mathbb{Z}\}$. We have

$$\begin{aligned}
[0] &= \{\ldots, -15, -10, -5, 0, 5, 10, 15, \ldots\} \\
[1] &= \{\ldots, -14, -9, -4, 1, 6, 11, 16, \ldots\} \\
[2] &= \{\ldots, -13, -8, -3, 2, 7, 12, 17, \ldots\} \\
[3] &= \{\ldots, -12, -7, -2, 3, 8, 13, 18, \ldots\} \\
[4] &= \{\ldots, -11, -6, -1, 4, 9, 14, 19, \ldots\} \\
[5] &= \{\ldots, -10, -5, 0, 5, 10, 15, 20, \ldots\}
\end{aligned}$$

Clearly $[0] = [5]$. Each equivalence class $[n]$ consists of a set of integers that is infinite in both directions with gaps of length 5 between consecutive integers. The set $[n + 1]$ is obtained from the set $[n]$ by shifting all integers to the right by 1. Since the gaps have length 5, the set $[n + 5]$ is obtained from $[n]$ by shifting all integers to the right by 5 and we obtain the same set as $[n]$, i.e., $[n + 5] = [n]$. This is to say, that we have only 5 different equivalence classes: $[0]$, $[1]$, $[2]$, $[3]$, $[4]$. Every integer belongs to one and only one equivalence class and $[a] = [b]$ if and only if $a \equiv b \pmod{5}$, that is if and only if a and b differ by a multiple of 5. The collection of these equivalence classes is denoted by

$$\mathbb{Z}_5 = \{[0], [1], [2], [3], [4]\},$$

and with some abuse of notation we can simply identify $\mathbb{Z}_5$ with the set $\{0, 1, 2, 3, 4\}$.

Example 5.5. The above example can be easily generalized as follows. Let $n \in \mathbb{N}$. Then the equivalence classes of the congruence relation $a \equiv b \pmod{n}$ are

$$[a] = \{a + kn : k \in \mathbb{Z}\} = \{\dots, a - 2n, a - n, a, a + n, a + 2n, \dots\},$$

$[a] = [a + n]$ and we have only n distinct equivalence classes $[0], [1], \dots, [n - 1]$. Every integer belongs to one and only one equivalence class and $[a] = [b]$ if and only if $a \equiv b \pmod{n}$. The equivalence classes form the set

$$\mathbb{Z}_n = \{[0], [1], [2], \dots, [n - 1]\},$$

and we can identify it with $\mathbb{Z}_n = \{0, 1, 2, \dots, n - 1\}$.

In the case of an arbitrary equivalence relation on a set X we have

Theorem 5.6. *Let $\sim$ be an equivalence relation on a set X. Then every element of X belongs to one and only one equivalence class. More precisely, for every $a \in X$, $a \in [a]$ and if $a \in [b]$, then $[a] = [b]$. Moreover $a \sim b$ if and only if $[a] = [b]$.*

That means the set X is partitioned into a collection of pairwise disjoint equivalence classes.

Proof. If $a \in A$, then $a \sim a$ (reflexive relation) so $a \in [a]$. Suppose now that $a \in [b]$. We need to show that $[a] = [b]$.

Since $a \in [b]$, it follows from the definition of $[b]$ that $a \sim b$ and hence $b \sim a$ (symmetric relation).

If $c \in [a]$ then $c \sim a$, $a \sim b$ so $c \sim b$ (transitive relation) and hence $c \in [b]$. This proves that $[a] \subset [b]$. If $c \in [b]$, then $c \sim b$, $b \sim a$ and hence $c \sim a$ so $c \in [a]$. This proves that $[b] \subset [a]$. Since $[a] \subset [b]$ and $[b] \subset [a]$ we have that $[a] = [b]$.

It remains to prove that $a \sim b$ if and only if $[a] = [b]$. If $a \sim b$, then $a \in [b]$, but (as we proved above) it implies that $[b] = [a]$. If $[a] = [b]$, then $a \in [a] = [b]$, $a \in [b]$ and hence $a \sim b$. This completes the proof that $a \sim b$ is equivalent to $[a] = [b]$. $\square$

5.2 Integers

Every child knows what the natural numbers are. I have 5 fingers in my palm, I can count them! But what is -5? I cannot think of -5 fingers! In the elementary school they teach us that the temperature can be negative. Just look at your thermometer! Or that you can have negative savings if you have a debt. All these examples are just to convince us that we should take negative integers for granted and just to learn how to use them. Imagine that you want to teach a child (who knows natural numbers including 0) about all integers, but you do not have in mind examples with the temperature or debt. Then the natural (but not very didactic) way to introduce integers would be to say:

You know, integers are just numbers $n - m$, where $n, m \in \mathbb{N}_*$. If $n \geq m$, then you know that it is a number in $\mathbb{N}_*$, but you can also write $n - m$, even if $n < m$. Don't worry what it is. Treat it formally and just learn how to compute. First, remember that $n_1 - m_1 = n_2 - m_2$ if $n_1 + m_2 = n_2 + m_1$. You know how to add natural numbers, don't you? You add integers as follows: $(n_1 - m_1) + (n_2 - m_2) = (n_1 + n_2) - (m_1 + m_2)$ and multiply them by $(n_1 - m_1)(n_1 - m_2) = (n_1 n_2 + m_1 m_2) - (n_1 m_2 + m_1 n_2)$. Aha, there is one more thing. You can regard natural numbers $n \in \mathbb{N}_*$ as integers of the form $n - 0$ and we also write $-n = 0 - n$. Now you know everything you need to know. Go and do your homework!

While, I am sure, the child would appreciate your lesson (a poor child left behind), this is exactly what the integers are and this is how we define them in mathematics. We identify integers with ordered pairs of natural numbers (n, m), $n, m \in \mathbb{N}_*$. Such a pair represents $n - m$ and we identify pairs (n_1, m_1) and (n_2, m_2) if $n_1 + m_2 = n_2 + m_1$. That means we introduce the relation: $(n_1, m_1) \sim (n_2, m_2)$ if and only if $n_1 + m_2 = n_2 + m_1$. This relation turns out to be an equivalence relation and we identify integers with equivalence classes of this relation. Okay, let's do it formally.

In the set $\mathbb{N}_* \times \mathbb{N}_*$ we define the relation $\sim$ as follows

$$(n_1, m_1) \sim (n_2, m_2) \quad \text{if and only if} \quad n_1 + m_2 = n_2 + m_1.$$

This is an equivalence relation. Indeed $(n, m) \sim (n, m)$ because $n + m = n + m$ so it is reflexive. If $(n_1, m_1) \sim (n_2, m_2)$, then $n_1 + m_2 = n_2 + m_1$ and hence $n_2 + m_1 = n_1 + m_2$ and the last condition means that $(n_2, m_2) \sim (n_1, m_1)$. Therefore the relation is symmetric. If $(n_1, m_1) \sim (n_2, m_2)$ and $(n_2, m_2) \sim (n_3, m_3)$, then $n_1 + m_2 = n_2 + m_1$ and $n_2 + m_3 = n_3 + m_2$ so

$$(n_1 + m_2) + (n_2 + m_3) = (n_2 + m_1) + (n_3 + m_2),$$

$$(n_1 + m_3) + (n_2 + m_2) = (n_3 + m_1) + (n_2 + m_2),$$

$$n_1 + m_3 = n_3 + m_1$$

and hence $(n_1, m_1) \sim (n_3, m_3)$, which proves that the relation is transitive.

We proved that $\sim$ is an equivalence relation and hence we can talk about equivalence classes of this relation.

The set of integers, denoted by $\mathbb{Z}$ is defined to be the set of all equivalence relations

$$\mathbb{Z} = \{[(n, m)] : (n, m) \in \mathbb{N}_* \times \mathbb{N}_*\}.$$

If $a, b \in \mathbb{Z}$, then there are natural numbers $n_1, m_1, n_2, m_2 \in \mathbb{N}_*$ such that $a = [(n_1, m_1)]$ and $b = [(n_2, m_2)]$ and we define

$$a + b = [(n_1 + n_2, m_1 + m_2)],$$

$$a \cdot b = [(n_1 n_2 + m_1 m_2, n_1 m_2 + m_1 n_2)]$$

Note that the natural numbers n_1, m_1, n_2, m_2 are not uniquely determined so in order to show that addition and multiplication of integers is well defined we need to show that if we represent a and b differently, that is $a = [(\tilde{n}_1, \tilde{m}_1)]$ and $b = [(\tilde{n}_2, \tilde{m}_2)]$, then the result of addition and multiplication computed with this new representation will be the same as with the old one, that is we need to show that

$$[(\tilde{n}_1+\tilde{n}_2,\tilde{m}_1+\tilde{m}_2)] = [(n_1+n_2,m_1+m_2)], \tag{5.1}$$

$$[(\tilde{n}_1\tilde{n}_2+\tilde{m}_1\tilde{m}_2,\tilde{n}_1\tilde{m}_2+\tilde{m}_1\tilde{n}_2)] = [(n_1n_2+m_1m_2,n_1m_2+m_1n_2)]. \tag{5.2}$$

We will verify (5.1) only, leaving verification of (5.2) as an exercise.

We have

$$a=[(n_1,m_1)]=[(\tilde{n}_1,\tilde{m}_1)], \quad b=[(n_2,m_2)]=[(\tilde{n}_2,\tilde{m}_2)]$$

so

$$n_1+\tilde{m}_1=\tilde{n}_1+m_1, \quad n_2+\tilde{m}_2=\tilde{n}_2+m_2.$$

Adding both equalities yields

$$n_1+\tilde{m}_1+n_2+\tilde{m}_2=\tilde{n}_1+m_1+\tilde{n}_2+m_2,$$

$$(\tilde{n}_1+\tilde{n}_2)+(m_1+m_2)=(n_1+n_2)+(\tilde{m}_1+\tilde{m}_2),$$

which proves (5.1).

Let $\underline{0}=[(0,0)]$ and $\underline{1}=[(1,0)]$. If $a=[(n,m)]\in\mathbb{Z}$, then we write $-a=[(m,n)]$.

Theorem 5.7. *The following properties are satisfied by all $a,b,c\in\mathbb{Z}$:*

(Z1) $(a+b)+c=a+(b+c)$,

(Z2) $a+b=b+a$,

(Z3) $a+\underline{0}=a$,

(Z4) $a+(-a)=\underline{0}$,

(Z5) $(a\cdot b)\cdot c=a\cdot(b\cdot c)$,

(Z6) $a\cdot b=b\cdot a$,

(Z7) $a\cdot\underline{1}=a$,

(Z8) $a\cdot(b+c)=(a\cdot b)+(a\cdot c)$.

(Z9) $\underline{1}\neq\underline{0}$.

All these properties are easy to verify directly from the definitions and the proof is left to the reader. In what follows we take them for granted.

The subtraction $k-\ell$ is well defined among elements of $\mathbb{N}_*$ as long as $k\geq\ell$. Namely, we define $k-\ell$ to be a unique $n\in\mathbb{N}_*$ such that $k=n+\ell$, but $k-\ell$ makes no sense in $\mathbb{N}_*$ if $k<\ell$.

Each integer can be uniquely represented as

$$[(n,0)] \quad \text{or} \quad [(0,n)]. \tag{5.3}$$

Indeed, let $[(k,\ell)] \in \mathbb{Z}$ be any integer. If $k \geq \ell$, then $[(k,\ell) = [(n,0)]$, where $n = k-\ell \in \mathbb{N}_*$. If $\ell \geq k$, then $[(k,\ell)] = [(0,n)]$, where $n = \ell - k \in \mathbb{N}_*$. The integer n representing $[(k,\ell)]$ is unique. If $[(n,0)] = [(\tilde{n},0)]$, then $n+0 = \tilde{n}+0$, $n = \tilde{n}$. Similarly, if $[(0,n)] = [(0,\tilde{n})]$, then $n = \tilde{n}$.

Note also that $[(0,n)] = -[(n,0)]$. Because of the (unique) representation (5.3) of integers, we can define $\underline{n} = [(n,0)]$ so $-\underline{n} = [(0,n)]$ and we can list all integers in $\mathbb{Z}$ as

$$\ldots,\ \underbrace{[(0,4)]}_{-\underline{4}},\ \underbrace{[(0,3)]}_{-\underline{3}},\ \underbrace{[(0,2)]}_{-\underline{2}},\ \underbrace{[(0,1)]}_{-\underline{1}},\ \underbrace{[(0,0)]}_{\underline{0}},\ \underbrace{[(1,0)]}_{\underline{1}},\ \underbrace{[(2,0)]}_{\underline{2}},\ \underbrace{[(3,0)]}_{\underline{3}},\ \underbrace{[(4,0)]}_{\underline{4}},\ \ldots \tag{5.4}$$

We now can identify natural numbers with integers, by identifying $n \in \mathbb{N}_*$ with $\underline{n} = [(n,0)]$, but we still need to check that the arithmetic operations of addition and multiplication in $\mathbb{N}_*$ are consistent with the corresponding operations in $\mathbb{Z}$. That is we need to check that for all $n,m \in \mathbb{N}_*$ we have

$$\underline{n+m} = \underline{n} + \underline{m} \quad \text{and} \quad \underline{n \cdot m} = \underline{n} \cdot \underline{m}.$$

The verification is easy and left to the reader.

Finally, we can equip integers with the inequality $\leq$ which is a relation defined as follows. For any $a,b \in \mathbb{Z}$

$$a \leq b \qquad \text{if and only if} \qquad b = a + \underline{n} \text{ for some } n \in \mathbb{N}_*.$$

With this inequality, integers in (5.4) are listed in the increasing order.

Theorem 5.8. *For all integers $a,b,c \in \mathbb{Z}$, the relation of inequality $\leq$ has the following properties:*

(Z10) $a \leq a$.

(Z11) $a \leq b \wedge b \leq a \Rightarrow a = c$.

(Z12) $a \leq b \wedge b \leq c \Rightarrow a \leq c$.

(Z13) *Either $a \leq b$ or $b \leq a$.*

(Z14) $a \leq b \Rightarrow a + c \leq b + c$

(Z15) $\underline{0} \leq a \wedge \underline{0} \leq b \Rightarrow 0 \leq a \cdot b$.

We also write $a \geq b$ if $b \leq a$.

From now one we will write n and $-n$ instead of $\underline{n}$ and $-\underline{n}$ and we will regard $\mathbb{N}_*$ (and hence $\mathbb{N}$) as a subset of $\mathbb{Z}$.

5.3 Rational numbers

In the ancient Greece they did not talk about rational numbers, but about proportions. Two segments (areas etc.) A and B are in the proportion like $n : m$ if A increased m times equals B increased n times. Here, of course m and n are positive integers. Clearly proportions $n_1 : m_1$ and $n_2 : m_2$ are equivalent if $n_1m_2 = n_2m_1$. Rational numbers are the abstraction of a notion of a proportion and formally they have to be defined as equivalence classes of proportions $n : m$. Here we however, allow the integers to be negative.

Consider

$$F = \{(n, m) \in \mathbb{Z} \times \mathbb{Z} :\ m \neq 0\}$$

and define the relation

$$(n_1, m_1) \sim (n_2, m_2) \quad \text{if and only if} \quad n_1m_2 = n_2m_1.$$

A reader will not have difficulty to check that this is an equivalence relation. The set of *rational numbers* is

$$\mathbb{Q} = \{[(n, m)] :\ (n, m) \in \mathbb{Z}\}$$

with the operations of addition and multiplication defined below.

The equivalence classes are denote by

$$\frac{n}{m} := [(n, m)].$$

Let $r, s \in \mathbb{Q}$. Then

$$r = \frac{n_1}{m_1},\ s = \frac{n_2}{m_2} \quad \text{for some } n_1, m_1, n_2, m_2 \in \mathbb{Z},\ nm_1, m_2 \neq 0$$

and we define

$$r + s = [(n_1m_2 + n_2m_1, m_1m_2)] = \frac{n_1m_2 + n_2m_1}{m_1m_2}, \quad rs = [(n_1n_2, m_1m_2)] = \frac{n_1n_2}{m_1m_2}.$$

As in the case of the construction of integers we need to show that the definition of $r + s$ and rs does not depend on how we represent r and s. That is, we need to show that if also

$$r = \frac{\tilde{n}_1}{\tilde{m}_1},\ s = \frac{\tilde{n}_2}{\tilde{m}_2}, \quad \text{that is} \quad \tilde{n}_1m_1 = n_1\tilde{m}_1,\ \tilde{n}_2m_2 = n_2\tilde{m}_2,$$

then

$$\frac{\tilde{n}_1\tilde{m}_2 + \tilde{n}_2\tilde{m}_1}{\tilde{m}_1\tilde{m}_2} = \frac{n_1m_2 + n_2m_1}{m_1m_2} \quad \text{and} \quad \frac{\tilde{n}_1\tilde{n}_2}{\tilde{m}_1\tilde{m}_2} = \frac{n_1n_2}{m_1m_2}.$$

We will prove the second equality, leaving the proof of the first one as an exercise. First note that

$$\frac{n}{m} = \frac{kn}{km}, \quad \text{for all } n, m, k \in \mathbb{Z},\ m, k \neq 0, \text{ because } n(km) = (kn)m.$$

Using this equality we can argue as follows:

$$\frac{\tilde{n}_1\tilde{n}_2}{\tilde{m}_1\tilde{m}_2} = \frac{m_1m_2\tilde{n}_1\tilde{n}_2}{m_1m_2\tilde{m}_1\tilde{m}_2} = \frac{(\tilde{n}_1m_1)(\tilde{n}_2m_2)}{m_1m_2\tilde{m}_1\tilde{m}_2} = \frac{(n_1\tilde{m}_1)(n_2\tilde{m}_2)}{m_1m_2\tilde{m}_1\tilde{m}_2} = \frac{n_1n_2}{m_1m_2}.$$

If $r \neq \frac{0}{1}$, i.e., $r = n/m$, where $n, m \neq 0$, then we write $r^{-1} = m/n$. It is not difficult to prove that addition and multiplication in $\mathbb{Q}$ have the following properties.

Theorem 5.9. *For all $r, s, t \in \mathbb{Q}$ we have*

(Q1) $(r + s) + t = r + (s + t)$,

(Q2) $r + s = s + t$,

(Q3) $r + \frac{0}{1} = r$,

(Q4) $r + (-r) = \frac{0}{1}$,

(Q5) $(r \cdot s) \cdot t = r \cdot (s \cdot t)$,

(Q6) $r \cdot s = s \cdot r$,

(Q7) $r \cdot \frac{1}{1} = r$,

(Q8) *If* $r \neq \frac{0}{1}$, *then* $r \cdot (r^{-1}) = \frac{1}{1}$,

(Q9) $r \cdot (s + t) = (r \cdot s) + (r \cdot t)$.

(Q10) $\frac{1}{1} \neq \frac{0}{1}$.

We leave the proof as an exercise.

Note that the property (Q8) has no counterpart in $\mathbb{Z}$ (see Theorem 5.7), all other properties do.

Now, each integer $n \in \mathbb{Z}$ can be identified with a rational number $\frac{n}{1}$. Namely the function

$$\Phi : \mathbb{Z} \to \mathbb{Q} \quad \text{defined by} \quad \Phi(n) = \frac{n}{1}$$

is one-to-one and it has the following properties $\Phi(0) = \frac{0}{1}$, $\Phi(1) = \frac{1}{1}$ and

$$\Phi(n + m) = \Phi(n) + \Phi(m),\ \Phi(n \cdot m) = \Phi(n) \cdot \Phi(m) \quad \text{for all } n, m \in \mathbb{Z}.$$

That means the operations of addition and multiplication of integers n and m are consistent with the operations of addition and multiplication of these integers regarded as rational numbers $\frac{n}{1}$ and $\frac{m}{1}$.

In what follows we will regard $\mathbb{Z}$ as a subset of $\mathbb{Q}$ and simply write n instead of $\frac{n}{1}$.

Each rational number can be represented as

$$r = \frac{n}{m}, \quad \text{where } n \in \mathbb{Z} \text{ and } m \in \mathbb{N}. \tag{5.5}$$

Indeed, if $r = k/\ell$, and $\ell > 0$, then we take $n = k$, $m = \ell$. If $r = k/\ell$, and $\ell < 0$, then $r = (-k)/(-\ell)$ and we take $n = -k$, $m = -\ell$.

Therefore for any $q \in \mathbb{Q}$, there is $m \in \mathbb{N}$ (denominator in (5.5)) such that $mq \in \mathbb{Z}$. If $r, s \in \mathbb{Q}$, then we can find $m \in \mathbb{N}$ such that $mr, ms \in \mathbb{Z}$ (common denominator).

Integers are equipped with the relation of the inequality $\leq$. It can be extended to the set of all rational numbers $\mathbb{Q}$ as follows: $r \leq s$ if and only if there is $m \in \mathbb{N}$ such that $mr, ms \in \mathbb{Z}$ and $mr \leq ms$.

Theorem 5.10. *For all $r, s, t \in \mathbb{Q}$ we have*

(Q11) $r \leq s$.

(Q12) $r \leq s \wedge s \leq r \Rightarrow r = s$.

(Q13) $r \leq s \wedge s \leq t \Rightarrow r \leq t$.

(Q14) *Either* $r \leq s$ *or* $s \leq r$.

(Q15) $r \leq s \Rightarrow r + t \leq s + t$

(Q16) $\frac{0}{1} \leq r \wedge \frac{0}{1} \leq s \Rightarrow \frac{0}{1} \leq r \cdot s$.

5.4 Problems

Problem 1. Prove the property (Z14): for all $a, b, c \in \mathbb{Z}$ we have $a \leq b \Rightarrow a + c \leq b + c$.

Problem 2. Let

$$F = \{(n, m) \in \mathbb{Z} \times \mathbb{Z} : m \neq 0\}.$$

Prove that the relation

$$(n_1, m_1) \sim (n_2, m_2) \quad \text{if and only if} \quad n_1 m_2 = n_2 m_1.$$

is an equivalence relation.

Problem 3. Let[1] $F_* = \mathbb{Z} \times \mathbb{Z}$ and define the relation

$$(n_1, m_1) \sim (n_2, m_2) \quad \text{if and only if} \quad n_1 m_2 = n_2 m_1.$$

We can try to mimic the construction of rational numbers with the set $\mathbb{F}_*$ in place of $\mathbb{F}$. Where does the theory fall apart?

Problem 4. $\mathbb{Z}_{11} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ is a field. Find $7 + 8$, $7 \cdot 8$, $7/8$ in $\mathbb{Z}_{11}$.

Problem 5. Let X be a non-empty set and let $P(X)$ be the power set. In the power set we have the operation of addition $A + B := A \cup B$. Check which of the axioms A1, A2, A3, A4 are satisfied.

Problem 6. Give an example of a non-empty set X with a relation $\sim$ that is

(a) reflexive, but not symmetric or transitive;

(b) symmetric, but not reflexive or transitive;

(c) transitive, but not reflexive or symmetric;

[1] It differs from F in Problem 2, because we include $(n, 0)$. Remember that $[(n, m)]$ represents the fraction $\frac{n}{m}$ so now we try to develop rational numbers, where we can divide by 0.

(d) reflexive and symmetric, but not transitive;

(e) reflexive and transitive, but not symmetric;

(f) symmetric and transitive, but not reflexive.

Problem 7. Let X be a nonempty set with a relation $\sim$. Consider the following argument.
Claim: *If $\sim$ is symmetric and transitive, then it is also reflexive.*
Proof. $x \sim y \Rightarrow y \sim x$; also $x \sim y$ and $y \sim x \Rightarrow x \sim x$. Therefore $x \sim x$ for every $x \in X$. □

This argument is at odds with part (f) of the previous problem. Where is the flaw?

Chapter 6

Real numbers

6.1 The first sixteen axioms

As we already mentioned in Section 5.3, in the ancient Greece rational numbers were understood as proportions. They believed that any two quantities are in a certain proportion. In particular, they believed that given two segments A and B, there are natural numbers n and m such that A increased m times and B increased n have the same length. In other words the segments A and B are in the proportion $n : m$ for some natural numbers n and m. However, Hippasus, one of the students of Pythagoras, discovered that there is no proportion between the diagonal of the square and its side. No matter how many times we will increase the length of the diagonal and how many times we will increase the length of a side, the two lengths will never be equal. Pythagoras could not accept this fact and a legend says that Hippasus was thrown overboard and drowned.

Translating into modern language we would say that Hippasus proved that $\sqrt{2}$ is irrational. We know what the rational numbers are and here we have an example of a (real) number that is not rational. This leads to a question: What are the real numbers? How can we define them? In the ancient Greece they did not know how to answer this question and there was no good answer until XIXth century.

Actually, the formal construction of real numbers is much more complicated than the formal construction of the rational numbers. Instead we will list all properties that uniquely determine real numbers. Such properties are called *axioms*. Recall that rational numbers satisfy properties (Q1)-(Q16) listed in Theorems 5.9 and 5.10. These are so called *axioms of an ordered field* (we will discuss it below in details). These axioms do not determine $\mathbb{R}$, because both $\mathbb{Q}$ and $\mathbb{R}$ satisfy them. We will need one more axiom (*axiom of completeness*) that will distinguish $\mathbb{R}$ from $\mathbb{Q}$ and, in fact will determine $\mathbb{R}$ uniquely (up to an isomorphism). These 17 axioms provide a complete set of rules needed in order to use real numbers. All properties of real numbers can be concluded from them. Just like the axioms of the set theory are all what we need in order to use sets. However, we still need to make sure that there is a set with operations of addition and multiplication that satisfies all the 17 axioms. That is, we need to construct the set of real numbers. We have constructed a set of rational numbers satisfying properties (axioms) (Q1)-(Q16) so we know that an ordered field exists, but real numbers satisfy one more axiom (axiom of completeness) and we need to make sure that there is a set satisfying the axiom of completeness. Such a construction was provided in XIXth century by Dedekind and we will briefly describe it in the last Section 6.5. Well, we said a lot of words here making the above description opaque, but don't worry, we will explain everything what was mentioned here.

The first sixteen axioms of real numbers are modelled on properties (Q1)-(Q16).

A *field* is a set $\mathbb{F}$ with operations of addition '+' and multiplication '·' that satisfies the following 10 axioms:

(A1) $x + y = y + x$.

(A2) $x + (y + z) = (x + y) + z$.

(A3) There is an element denoted by 0 such that for every x, $x + 0 = x$.

(A4) For every x there is an element denoted by $-x$ such that $x + (-x) = 0$.

(A5) $x \cdot y = y \cdot x$

(A6) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$

(A7) There is an element denoted by 1 such that $x \cdot 1 = x$.

(A8) For every $x \neq 0$ there is an element denoted by x^{-1} such that $x \cdot (x^{-1}) = 1$.

(A9) $x \cdot (y + z) = x \cdot y + x \cdot z$.

(A10) $1 \neq 0$.

Here we understand that conditions are satisfied for all $x, y, z \in \mathbb{F}$. We did not include quantifiers in the conditions to make them more transparent.

The element 0 is the *neutral* element of addition and 1 is the *neutral* element of the multiplication. Note that despite the fact we use symbols 0 and 1, we *do not* assume that 0 and 1 are integers equal 0 and 1. We just use the same notation, because of similarity of properties to those of integers 0 and 1 Elements $-x$ and x^{-1} are called the *inverse elements of addition* and *multiplication* respectively.

We will also often write xy instead of $x \cdot y$.

Axioms A2 and A6 say that the order in which we add elements and the order in which we multiply elements does not matter and therefore we may write $x + y + z$ and xyz in place of expressions at A2 and A6.

Axiom A3 says that there is an element 0 such that $x + 0 = x$ for all x. However, it does not say that the element 0 is unique. It is unique and that is a theorem. In what follows by **T1, T2**,... we will denote a very tiny theorems that easily follows from the axioms.

T1. *If $\tilde{0}$ has the property that $x + \tilde{0} = x$ for all x, then $\tilde{0} = 0$.*

Indeed,

$$\underbrace{\tilde{0} = \tilde{0} + 0}_{\text{property of } 0 \text{ with } x = \tilde{0}} \overset{(A1)}{=} \underbrace{0 + \tilde{0} = 0}_{\text{property of } \tilde{0} \text{ with } x = 0}.$$

Similarly we prove

T2. *If $\tilde{1}$ has the property that $x \cdot \tilde{1} = x$ for all x, then $\tilde{1} = 1$.*

In the next result we prove that the inverse element of addition $-x$ is unique.

T3. *Let $x \in \mathbb{F}$. If an element $-\tilde{x} \in \mathbb{F}$ has the property that $x + (-\tilde{x}) = 0$, then $-\tilde{x} = -x$.*

Proof. Let $x \in \mathbb{F}$ and assume that $-\tilde{x}$ has the property **(A4')**: $x + (-\tilde{x}) = 0$. Then

$$\begin{aligned} -\tilde{x} &\overset{(A3)}{=} (-\tilde{x}) + 0 \overset{(A1)}{=} 0 + (-\tilde{x}) \overset{(A4)}{=} (x + (-x)) + (-\tilde{x}) \overset{(A1)}{=} ((-x) + x) + (-\tilde{x}) \\ &\overset{(A2)}{=} (-x) + (x + (-\tilde{x})) \overset{(A4')}{=} (-x) + 0 \overset{(A3)}{=} -x \end{aligned}$$

proving that $-\tilde{x} = -x$. □

Similarly we prove uniqueness of x^{-1}.

T4. *Let $0 \neq x \in \mathbb{F}$. If an element $\tilde{x}^{-1} \in \mathbb{F}$ has the property that $x \cdot \tilde{x}^{-1} = 1$, then $\tilde{x}^{-1} = x^{-1}$.*

Note that a set of rational numbers $\mathbb{Q}$, the set of real numbers $\mathbb{R}$ or a set of complex numbers $\mathbb{C}$ are fields, but there are more examples. Much more. Although the properties look like very natural properties of numbers, there are unexpected examples of fields.

Example 6.1. Recall that for $n \in \mathbb{N}$, $n \geq 2$, $\mathbb{Z}_n = \{[0], [1], \ldots, [n-1]\}$ was defined in Example 5.5 as the set of equivalence classes of the congruence relation $a \equiv b \pmod{n}$.

We define in $\mathbb{Z}_n$ addition and multiplication by

$$[a] + [b] = [a + b], \quad [a] \cdot [b] = [ab].$$

First we need to prove that the operations of addition and multiplication are well defined. That is we need to prove:

$$\text{If } ([\tilde{a}] = [a] \text{ and } [\tilde{b}] = [b]), \quad \text{then} \quad ([\tilde{a} + \tilde{b}] = [a + b] \text{ and } [\tilde{a}\tilde{b}] = [ab]).$$

Since $[\tilde{a}] = [a]$ if and only if $\tilde{a} \equiv a \pmod{n}$ and $[\tilde{b}] = [b]$ if and only if $\tilde{b} \equiv b \pmod{n}$, (Theorem 5.6), we have that

$$\tilde{a} = a + kn, \quad \tilde{b} = b + \ell n \quad \text{for some } k, \ell \in \mathbb{Z}.$$

Therefore

$$\tilde{a} + \tilde{b} = a + b + (k + \ell)n, \quad \tilde{a} + \tilde{b} \equiv a + b \pmod{n}, \quad [\tilde{a} + \tilde{b}] = [a + b].$$

Similarly

$$\tilde{a}\tilde{b} = (a+kn)(b+\ell n) = ab + a\ell n + bkn + kln^2 = ab + (a\ell + bk + k\ell n)n,$$

$$\tilde{a}\tilde{b} \equiv ab \pmod{n}, \qquad [\tilde{a}\tilde{b}] = [ab].$$

Note that

$$[a] + [0] = [a], \quad [a] + [-a] = [0], \quad [a] \cdot [1] = [a]$$

and it easily follows that $\mathbb{Z}_n$ with $0 = [0]$, $1 = [1]$ and $-[a] = [-a]$ satisfies all axioms (A1)-(A7), A9 and A10. It turns out however, that the axiom A8 need not be satisfied. The case of axiom A8 is completely answered by the next result.

Theorem 6.2. *$\mathbb{Z}_n$ is a field, that is it satisfies all axioms (A1)-(A10), if and only if $n = p$ is a prime number.*

Proof. We know that $\mathbb{Z}_n$ satisfies all (A1)-(A10) with a possible exception of the axiom A8. Therefore it remians to show that the axiom A8 is satisfied if and only if $n = p$ is a prime number. That is we need to prove two implications: If A8 is satisfied, then n is a prime number and if n is a prime number, then A8 is satisfied.

To prove the first implication it suffices to show that if $n \geq 2$ is not a prime number, then A8 is not satisfied.[1] Assume that $n \geq 2$ is not prime, then $n = k\ell$, for some $k, \ell \in \mathbb{N}$, $k, \ell > 1$. We will show that the element $[k] \in \mathbb{Z}_n$ does not have the multiplicative inverse $[k]^{-1}$ proving that the axiom A8 is not satisfied. Suppose to the contrary that there is $a \in \mathbb{Z}$ such that $[a] = [k]^{-1}$ i.e., $[k][a] = [ka] = [1]$. Then $ak \equiv 1 \pmod{n}$. That is $ak - 1 = mn = m(k\ell)$ for some $m \in \mathbb{Z}$, $1 = k(a - m\ell)$. This however, gives a contradiction, because 1 is not divisible by k.[2]

Now we will prove the second implication. Suppose now that $n = p$ is a prime number. We need to show that the axiom A8 is satisfied, that is, we need to show that for any $[a] \neq [0]$, there is an integer x such that $[x] = [a]^{-1}$, that is $[a][x] = [ax] = [1]$.

We will need the following result from the elementary number theory that we state without proof.

Lemma 6.3 (Bézout's identity)**.** *Let $a, b \in \mathbb{Z} \setminus \{0\}$ be relatively prime integers i.e., integers with the greatest common divisor equal 1. Then there are integers $x, y \in \mathbb{Z}$ such that $ax + by = 1$.*

Since $[a] \neq [0]$, a is not divisible by p. Since p is prime it follows that a and p are relatively prime so there are integers $x, y \in \mathbb{Z}$ such that $ax + py = 1$. However, this equality means that $ax \equiv 1 \pmod{p}$ so $[1] = [ax] = [a][x]$ proving that $[x] = [a]^{-1}$. □

In a field we define the following operations

$$x - y := x + (-y),$$

[1] Because $p \Rightarrow q$ is equivalent to $\neg q \Rightarrow \neg p$, an argument by contrapositive.

[2] The proof presented here has an interesting logical structure. Inside a contrapositive argument we used an argument by contradiction.

$$\frac{x}{y} := x \cdot (y^{-1}), \quad \text{provided } y \neq 0,$$

$$x^2 := x \cdot x,\ x^3 := x^2 \cdot x, \dots$$

$$2 := 1 + 1,\ 3 := 2 + 1,\ 4 := 3 + 1, \dots$$

Example 6.4. Since in $\mathbb{Z}_p$, $0 = [0]$, $1 = [1]$, we have $2 = 1 + 1 = [1] + [1] = [1 + 1] = [2]$, $3 = [3]$ etc. Thus

$$\mathbb{Z}_p = \{[0], [1], \dots, [p-1]\} = \{0, 1, \dots, p-1\}.$$

For example in the field $\mathbb{Z}_5$ we have

$$3 + 4 = [3] + [4] = [7] = [2] = 2, \quad 3 \cdot 2 = [3] \cdot [2] = [3 \cdot 2] = [6] = [1] = 1,\ 3^{-1} = 2.$$

All the properties of operations in a field can be directly deduced from the axioms. For example

Proposition 6.5. *In any field* $(a + b)^2 = a^2 + 2ab + b^2$.

Proof. The result seems obvious and well known, but since elements of fields can be different than real or rational numbers, this fact really requires a careful proof.

$$\begin{aligned}
(a+b)^2 &\overset{\text{def. of } x^2}{=} (a+b)\cdot(a+b) \\
&\overset{\text{(A9)}}{=} (a+b)\cdot a + (a+b)\cdot b \\
&\overset{\text{(A5)}}{=} a\cdot(a+b) + b\cdot(a+b) \\
&\overset{\text{(A9)}}{=} a\cdot a + a\cdot b + b\cdot a + b\cdot b \\
&\overset{\text{def. of } x^2}{=} a^2 + a\cdot b + b\cdot a + b^2 \\
&\overset{\text{(A5)}}{=} a^2 + a\cdot b + a\cdot b + b^2 \\
&\overset{\text{(A7)}}{=} a^2 + (a\cdot b)\cdot 1 + (a\cdot b)\cdot 1 + b^2 \\
&\overset{\text{(A9)}}{=} a^2 + (a\cdot b)\cdot(1+1) + b^2 \\
&\overset{\text{def. of } 2}{=} a^2 + (a\cdot b)\cdot 2 + b^2 \\
&\overset{\text{(A5)}}{=} a^2 + 2\cdot(a\cdot b) + b^2 \\
&= a^2 + 2ab + b^2.
\end{aligned}$$

The proof is complete. □

The above proof is quite long and very formal. For a while we will write such formal proofs, just to realize how to use axioms. In fact all standard rules of addition and multiplication are true in any field and, in practice, we work with them just like in the case of real numbers without paying much attention how to verify our computations directly from the axioms.

In addition to operations of addition and multiplication, the fields of rational and real numbers are equipped with the relation $\leq$ of being less than or equal to. This relation is not present in fields $\mathbb{C}$ or $\mathbb{Z}_p$, p-prime. The relation $\leq$ has to satisfy the following axioms (A11)-(A16) listed below.

A field $\mathbb{F}$ (i.e. a set with operations '+' and '$\cdot$' satisfying axioms (A1)-(A10)) equipped with a relation $\leq$ is called an *ordered field* if

(A11) $x \leq x$.

(A12) $x \leq y \wedge y \leq x \Rightarrow x = y$.

(A13) $x \leq y \wedge y \leq z \Rightarrow x \leq z$.

(A14) For every x, y either $x \leq y$ or $y \leq x$.

(A15) $x \leq y \Rightarrow x + z \leq y + z$

(A16) $0 \leq x \wedge 0 \leq y \Rightarrow 0 \leq xy$.

We also write $x \geq y$ if $y \leq x$.

Note that $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Z}_p$, where p is a prime number are fields, and $\mathbb{Q}$, $\mathbb{R}$ are ordered fields. There are however, other ordered fields than just $\mathbb{Q}$ and $\mathbb{R}$ and need one more axiom that will uniquely identify $\mathbb{R}$ among all ordered fields. We will discuss it in the next section.

In any ordered field we introduce more notation. For example we say that

$$x < y \text{ if } x \leq y \text{ and } x \neq y,$$

and we also write $y > x$ if $x < y$. Then for any two x, y exactly one of the following conditions is satisfied

$$x < y,\ y < x \text{ or } x = y.$$

To see how to use formal language of axioms we will prove some additional properties.

T5. $x \cdot 0 = 0 \cdot x = 0$.

Indeed,

$$\begin{aligned} 0 &\overset{(A4)}{=} x \cdot 0 + (-(x \cdot 0)) \overset{(A3)}{=} x \cdot (0 + 0) + (-(x \cdot 0)) \overset{(A9)}{=} (x \cdot 0 + x \cdot 0) + (-(x \cdot 0)) \\ &\overset{(A2)}{=} x \cdot 0 + (x \cdot 0 + (-(x \cdot 0)) \overset{(A4)}{=} x \cdot 0 + 0 \overset{(A3)}{=} x \cdot 0. \end{aligned}$$

We proved that $x \cdot 0 = 0$. Hence $0 \cdot x = x \cdot 0 = 0$ by A5.

T6. $-(-x) = x$.

Indeed, $(-x) + x \overset{(A1)}{=} x + (-x) \overset{(A4)}{=} 0$ so by T3, x is the unique additive inverse of $-x$ and hence $x = -(-x)$. We could also argue as follows:[3]

$$\begin{aligned} -(-x) &\overset{(A3)}{=} (-(-x)) + 0 \overset{(A4)}{=} (-(-x)) + (x + (-x)) \overset{(A1)}{=} (x + (-x)) + (-(-x)) \\ &\overset{(A2)}{=} x + ((-x) + (-(-x))) \overset{(A4)}{=} x + 0 \overset{(A3)}{=} x. \end{aligned}$$

[3]We already have one proof so we do not need another one, but we just want to practice how to use axioms.

T7. $(-x) \cdot y = x \cdot (-y) = -(x \cdot y)$.

By T5 we have $0 = y \cdot 0 = y \cdot (x + (-x)) = y \cdot x + y \cdot (-x) = x \cdot y + (-x) \cdot y$. Since $x \cdot y + (-x) \cdot y = 0$, by T3, $(-x) \cdot y$ is the unique additive inverse of $x \cdot y$ proving that $(-x) \cdot y = -(x \cdot y)$. Similarly we prove that $x \cdot (-y) = -(x \cdot y)$.

T8. $(-x) \cdot (-y) = x \cdot y$.

Indeed, $(-x) \cdot (-y) \overset{(T7)}{=} -(x \cdot (-y)) \overset{(T7)}{=} -(-(x \cdot y)) \overset{(T6)}{=} x \cdot y$.

T9. If $x > 0$ and $y > 0$, then $x + y > 0$.

Since $x \geq 0$, axiom A15 gives

$$x + y \geq 0 + y = y \geq 0,$$

so $x + y \geq 0$ by axiom A13. It remains to show that $x + y \neq 0$. Suppose to the contrary that $x + y = 0$, then the above inequality gives

$$0 = x + y \geq y$$

and hence $0 \geq y$, which is a contradiction with the assumption $y > 0$.

T10. If $x \neq 0$, then $-x \neq 0$.

Suppose to the contrary that $-x = 0$. Then $0 = x + (-x) = x + 0 = x$ which is a contradiction.

T11. If $x > 0$, then $-x < 0$.

Indeed, A11 and A15 give

$$0 \geq 0 = x + (-x) \geq 0 + (-x) = -x, \quad 0 \geq -x.$$

The fact that $x > 0$ implies that $x \neq 0$ so $-x \neq 0$ by T10. Thus $-x \leq 0$ and $-x \neq 0$ yields $-x < 0$.

Similarly we can prove

T12. If $x < 0$, then $-x > 0$.

T13. If $x \geq y$, then $-x \leq -y$.

Indeed, A15 yields

$$-y = (x + (-x)) + (-y) = x + ((-x) + (-y)) \geq y + ((-x) + (-y)) = (y + (-y)) + (-x) = -x.$$

T14. If $x > 0$ and $y > 0$, then $xy > 0$.

Indeed, $xy \geq 0$ by A16 and it remains to prove that $xy \neq 0$. Suppose to the contrary that $xy = 0$. Since $y \neq 0$, by A8, there is y^{-1} such that $y \cdot y^{-1} = 1$ and we have

$$0 = 0 \cdot y^{-1} = (x \cdot y) \cdot y^{-1} = x \cdot (y \cdot y^{-1}) = x \cdot 1 = x$$

which is a contradiction with the assumption $x > 0$.

T15. If $x \neq 0$, then $x^2 > 0$.

If $x > 0$, then $x^2 = x \cdot x > 0$ by T14. If $x < 0$, then $-x > 0$ by T12 so T8 and T14 give $x^2 = (-x) \cdot (-x) > 0$.

T16. $1 > 0$.

Since $1 \neq 0$, by A10, $1 = 1^2 > 0$ by T15.

T17. If $x < 0$ and $y < 0$, then $xy > 0$

Indeed, by T12 $-x > 0$ and $-y > 0$, so T8 and T14 yield

$$x \cdot y = (-x) \cdot (-y) > 0.$$

T18. If $x < 0$, then $y > 0$, then $xy < 0$.

Indeed, $-x > 0$, so

$$-(x \cdot y) = (-x) \cdot y > 0, \qquad x \cdot y < 0.$$

In any ordered field we define the absolute value by

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

Proposition 6.6. *If $\mathbb{F}$ is an ordered field and $x, y \in \mathbb{F}$, then*

T19. $|xy| = |x||y|$.

T20. $|x|^2 = x^2$.

T21. *If $c \geq 0$, then $|x| \leq c$ if and only if $-c \leq x \leq c$.*

T22. $-|x| \leq x \leq |x|$.

T23. $|x + y| \leq |x| + |y|$.

T24. $\big||x| - |y|\big| \leq |x - y|$.

Proof of T19. If either $x = 0$ or $y = 0$, then both sides are equal zero. Therefore we may assume that $x \neq 0$ and $y \neq 0$ and we have to consider four cases: (a) $x > 0, y > 0$; (b) $x > 0, y < 0$; (c) $x < 0, y > 0$; (d) $x < 0, y < 0$. As an illustration we will investigate the case (c) only; other cases are left to the reader. Since $x < 0$ and $y > 0$ we have $xy < 0$ by T18, so

$$|xy| = -(xy) = (-x) \cdot y = |x||y|.$$

Proof of T20. If $x \geq 0$, then $x = |x|$, so $|x|^2 = x^2$. If $x < 0$, then $|x| = -x$ and hence $|x|^2 = (-x) \cdot (-x) = x^2$ by T8.

Proof of T21. To prove "if and only if" condition we need to prove two implications. The implication from left to right ($\Rightarrow$) and the implication from right to left ($\Leftarrow$).

($\Rightarrow$) We need to show that if $|x| \leq c$, then $-c \leq x \leq c$. Remember that we assume $c \geq 0$. We have two cases.

If $x \geq 0$, then $x = |x| \leq c$, $x \leq c$. Since $c \geq 0$, $-c \leq 0$ by T13 and hence $x \geq 0 \geq -c$ gives $x \geq -c$. Now both inequalities that we proved can be put together: $-c \leq x \leq c$.

If $x < 0$, then $-x = |x| \leq c$, so $-x \leq c$ and hence $x \geq -c$ by T13. Since $x \leq 0 \leq c$, we also have $x \leq c$ and hence $-c \leq x \leq c$.

($\Leftarrow$) We need to prove that if $-c \leq x \leq c$, then $|x| \leq c$.

We have $x \leq c$ and the inequality $-c \leq x$ gives $-x \leq c$. Hence $|x| \leq c$, because $|x|$ is equal to x or $-x$.

Proof of T22. Taking $c = |x|$ in T21 we obtain $|x| \leq |x| = c$, so $-|x| \leq x \leq |x|$.

Proof of T23. We have $-|x| \leq x \leq |x|$, $-|y| \leq y \leq |y|$. Applying A15 several times we have

$$-(|x| + |y|) = -|x| - |y| \leq -|x| + y \leq x + y \leq x + |y| \leq |x| + |y|$$

and the inequality $|x + y| \leq |x| + |y|$ follows from T21.

Proof of T24. Since $x = (x - y) + y$, T23 gives $|x| \leq |x - y| + |y|$, so

$$|x| - |y| \leq |x - y| \tag{6.1}$$

Replacing x by y and y by x in the above inequality we obtain $|y| - |x| \leq |y - x| = |x - y|$, so

$$-|x - y| \leq |x| - |y|. \tag{6.2}$$

Now (6.1) and (6.2) give

$$-|x - y| \leq |x| - |y| \leq |x - y|$$

and the claim follows from T21. □

T25. If $x > 0$, then $x > x/2 > 0$.

Proof. First observe that

$$\frac{x}{2} + \frac{x}{2} = (x + x) \cdot 2^{-1} = (x \cdot 2) \cdot 2^{-1} = x.$$

To prove that $x/2 > 0$, suppose to the contrary that $x/2 \leq 0$. Then $x = x/2 + x/2 \overset{(A15)}{\leq} 0 + x/2 = x/2 \leq 0$ which contradicts the assumption that $x > 0$.

To prove that $x > x/2$ suppose to the contrary that $x \leq x/2$. Then

$$\frac{x}{2} = \frac{x}{2} + \frac{x}{2} + \left(-\frac{x}{2}\right) = x + \left(-\frac{x}{2}\right) \overset{(A15)}{\leq} \frac{x}{2} + \left(-\frac{x}{2}\right) = 0$$

which contradicts the fact that $x/2 > 0$ that we have just proved. □

Proposition 6.7. *Let $\mathbb{F}$ be an ordered field and $x \in \mathbb{F}$. If for every $\varepsilon > 0$, $x < \varepsilon$, then $x \leq 0$.*

Proof. Suppose to the contrary that $x > 0$. Take $\varepsilon = x/2$. Then $\varepsilon > 0$ and $x > \varepsilon$ by T25 so it is not true that $x < \varepsilon$. Contradiction. □

As we can see, we can prove all basic properties of operations, inequalities and absolute value directly from the axioms, but proving each new property takes a considerable amount of time. We will not continue this investigation and in what follows we will take all basic properties of operations and inequalities for granted without proving carefully how they follow from axioms. However, we have to be aware that it is really important to know that they can be proved directly from axioms, especially because the ordered fields can be very different from familiar fields of numbers.

6.2 Natural numbers, integers and rational numbers

In a field we can define[4] *quasi natural numbers* $\tilde{\mathbb{N}}$, *quasi integers* $\tilde{\mathbb{Z}}$, and *quasi rational numbers* $\tilde{\mathbb{Q}}$ as

$$\tilde{\mathbb{N}} = \{1, 2 = 1 + 1, 3 = 2 + 1, \ldots\},$$
$$\tilde{\mathbb{Z}} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\},$$
$$\tilde{\mathbb{Q}} = \left\{\frac{n}{m}, n, m \in \tilde{\mathbb{Z}},\ m \neq 0\right\}.$$

We call these numbers 'quasi', because for example in the field $\mathbb{Z}_5$, $7 = 2$, $-1 = 6$ or $2 \cdot 3 = 1$ so $2^{-1} = 3$ and hence $3/2 = 3 \cdot 3 = 4$. That means such 'quasi' natural numbers, integers of rational numbers are very different from 'real' ones.

However, in the ordered field we have $1 > 0$ (T16) and hence $n + 1 > n$ for all $n \in \tilde{\mathbb{Z}}$ so all quasi integers are ordered

$$\ldots, -3 < -2 < -1 < 0 < 1 < 2 < 3 < \ldots$$

[4]The names *quasi natural numbers*, *quasi integers*, and *quasi rational numbers* are not official and you will not find them in the literature. We will use this name only in this section and not in the remaining parts the the book.

and unlike in the case of $\mathbb{Z}_p$ all these quasi integers are different so there is a bijection between $\Phi : \mathbb{Z} \to \tilde{\mathbb{Z}}$ that maps $n \in \mathbb{Z}$ to $n \in \tilde{\mathbb{Z}}$. One can prove that

$$\Phi(n+m) = \Phi(n) + \Phi(m) \quad \text{and} \quad \Phi(nm) = \Phi(n)\Phi(m) \tag{6.3}$$

which means, addition and multiplication in $\tilde{\mathbb{Z}}$ is the same as in $\mathbb{Z}$ and we can identify $\tilde{\mathbb{Z}}$ in an ordered field with $\mathbb{Z}$. A bijection $\Phi : \mathbb{Z} \to \tilde{\mathbb{Z}}$ satisfying (6.3) is called an *isomorphism.*

Note that quasi rational numbers $\tilde{\mathbb{Q}}$ follow the rules (6.4) and (6.5)

$$\frac{n}{m} = \frac{k}{\ell} \quad \equiv \quad n\ell = mk \tag{6.4}$$

Indeed,

$$\frac{n}{m} = \frac{k}{\ell} \quad \equiv \quad (n \cdot m^{-1}) \cdot (m\ell) = (k \cdot \ell^{-1}) \cdot (m\ell) \quad \equiv \quad n\ell = mk.$$

In the first equivalence we use the fact that $(m\ell)^{-1}$ exists so that the equality in the middle implies the equality on the left (after we multiply both sides by $(m\ell)^{-1}$). Similarly we prove

$$\frac{n}{m} \cdot \frac{k}{\ell} = \frac{nk}{m\ell} \quad \frac{n}{m} + \frac{k}{\ell} = \frac{n\ell + mk}{m\ell}. \tag{6.5}$$

We will prove the formula for adding fractions leaving the formula for the product of fractions as an exercise. We have

$$\frac{n}{m} + \frac{k}{\ell} = \frac{n\ell + mk}{m\ell} \quad \equiv \quad \left(\frac{n}{m} + \frac{k}{\ell}\right) \cdot (m\ell) = \left(\frac{n\ell + mk}{m\ell}\right) \cdot (m\ell) \;\equiv\; n\ell + km = n\ell + mk$$

Since the last equality is obviously true and all equalities are equivalent, the first one is also true.

Thus we have the same operations in $\tilde{Q}$ as in $\mathbb{Q}$ and one can prove that there is a bijection $\Phi : \mathbb{Q} \to \tilde{\mathbb{Q}}$ such that

$$\Phi(r+s) = \Phi(r) + \Phi(s) \quad \text{and} \quad \Phi(rs) = \Phi(r)\Phi(s) \tag{6.6}$$

A bijection $\Phi : \mathbb{Q} \to \tilde{\mathbb{Q}}$ satisfying (6.6) is called an *isomorphism.*

In other words we can assume that any ordered field contains a (isomorphic) copy of $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$.

6.3 Supremum, infimum and the last axiom

The above list of axioms is not sufficient for the theory of real numbers. Indeed, $\mathbb{Q}$ satisfies all axioms (A1)-(A16) and we know that there are real numbers that are not rational. We still need one axiom. However, it is not very easy to formulate and we need to do some preparations.

Definition. Let $\mathbb{F}$ be an ordered field and let $\emptyset \neq A \subset \mathbb{F}$.

We say that a set A is *bounded from above* if there is $M \in \mathbb{F}$ such that $x \leq M$ for all $x \in A$. Each such M is called an *upper bound* of A.

We say that A is *bounded from below* if there is $m \in \mathbb{F}$ such that $m \leq x$ for all $x \in A$. Each such m is called a *lower bound* of A.

Finally the set A is called *bounded* if it is bounded from above and bounded from below. The set A is called *unbounded* if it is not bounded.

Similarly we can define sets *unbounded from above* and *unbounded from below*.

Example 6.8. *The set* $A = \{x \in \mathbb{Q} : x < 2\}$ *is bounded from above.*

Indeed, 2 is an upper bound and any $M \geq 2$ is an upper bound. However, 2 is the smallest (the least) upper bound of the set A.

Definition. Let $\mathbb{F}$ be an ordered field and let $\emptyset \neq A \subset \mathbb{F}$.

If A is bounded from above, then M is called a *supremum* (or a *least upper bound*) of A if

1. M is an upper bound of A;
2. If a is any upper bound of A, then $a \geq M$.

If A is bounded from below, then m is called an *infimum* (or a *greatest lower bound*) of A if

1. m is a lower bound of A;
2. If b is any lower bound of A, then $b \leq m$.

It easily follows from the definition that a set can have at most one supremum and at most one infimum, so if a set has a supremum of an infimum, then it is uniquely defined. If the supremum or the infimum of a set A exist, it is denoted respectively by

$$\sup A \qquad \text{and} \qquad \inf A.$$

If the set A is unbounded from above we write $\sup A = \infty$ and if it is unbounded from below we write $\inf A = -\infty$.

Example 6.9. *Let* $A \subset \mathbb{Q}$ *be defined by*

$$A = \left\{ \frac{x}{x+1} : x > 0, \ x \in \mathbb{Q} \right\}.$$

Find $\sup A$ *and* $\inf A$ *(in* $\mathbb{Q}$*).*

Solution. Clearly $0 < x/(x+1) < 1$ for all $x > 0$. Hence 0 is a lower bound and 1 is an upper bound of A. We will show that $\sup A = 1$ and $\inf A = 0$. To prove that $\sup A = 1$, i.e., that 1 is the least upper bound of A it suffices to show that if $y < 1$, then y cannot be an upper bound of A, that is there is $x > 0$, $x \in \mathbb{Q}$ such that $x/(x+1) > y$. To find such x we simply take any $x > y/(1-y)$ and one easily checks that this indeed gives $x/(x+1) > y$.[5] Similarly we prove that $\inf A = 0$. We leave details to the reader. □

The following result provides a useful method of verification whether a given upper bound M is the supremum.

Proposition 6.10. *Let $\mathbb{F}$ be an ordered field and let $M \in \mathbb{F}$ be an upper bound of $\emptyset \neq A \subset \mathbb{F}$. Then $M = \sup A$ if and only if*

$$\forall \varepsilon > 0\ \exists x \in A \quad (x > M - \varepsilon).$$

Remark 6.11. A careful reader will see that a similar argument has already been employed in the solution to the above exercise.

Proof. This result is nearly obvious, but let us prove it carefully. Let $M \in \mathbb{F}$ be an upper bound of A. We want to prove

$$(M = \sup A) \quad \equiv \quad (\forall \varepsilon > 0\ \exists x \in A \ \ x > M - \varepsilon).$$

$\Rightarrow$ Suppose to the contrary[6] that $M = \sup A$ and there is $\varepsilon > 0$ such that for all $x \in A$, $x \leq M - \varepsilon$. Then $a = M - \varepsilon < M$ is an upper bound of A which contradicts the fact that M was the least upper bound.

$\Leftarrow$ We know that M is an upper bound and we need to prove that it is the least one. Suppose to the contrary that there is a smaller upper bound $a < M$, i.e., $x \leq a$ for all $x \in A$. Take $\varepsilon = (M-a)/2$. Then $\varepsilon > 0$ and $M - \varepsilon > a$. It follows from the what we assumed that there is $x \in A$ such that $x > M - \varepsilon > a$ and that contradicts the assumption that a is an upper bound. □

Similarly we prove

Proposition 6.12. *Let $\mathbb{F}$ be an ordered field and let $m \in \mathbb{F}$ be a lower bound of $\emptyset \neq A \subset \mathbb{F}$. Then $m = \inf A$ if and only if*

$$\forall \varepsilon > 0\ \exists x \in A \quad (x < m + \varepsilon).$$

If $\mathbb{F}$ is an ordered field and $A \subset \mathbb{F}$ is bounded from above, then it is not necessarily true that A has supremum in $\mathbb{F}$. Here is an example. Let

$$A = \{x \in \mathbb{Q} :\ x > 0 \text{ and } x^2 \leq 2\}.$$

This set is bounded from above, i.e., by 2. Indeed, if $x > 2$, then $x^2 > 4$, so no element with $x > 2$ can belong to A and hence every element of A satisfies $x \leq 2$. It is also not very surprising that if M is the supremum of A, then it must satisfy $M^2 = 2$ (we will provide a rigorous proof of this fact later, see the proof of Theorem 6.14), but there is no rational number M such that $M^2 = 2$, so the set A has no supremum in $\mathbb{Q}$.

[5]The condition $x > y/(y-1)$ is obtained from solving the inequality $x/(x+1) > y$ for x.
[6]Recall that $\neg(p \Rightarrow q) \equiv p \wedge \neg q$.

A fundamental difference between $\mathbb{Q}$ and $\mathbb{R}$ is that every subset of $\mathbb{R}$ bounded from above has supremum.

Definition. We say that an ordered field $\mathbb{F}$ is *complete* if

(A17) every subset $A \subset \mathbb{F}$ bounded from above has supremum.

The axiom A17 is called the *axiom of completeness.*

The next important result shows that there is one, up to an isomorphism, only one complete ordered field.

Theorem 6.13. *There exists a complete ordered field F. Moreover if F_1 and F_2 are complete ordered fields, then they are isomorphis, i.e. there is a bijection $\Phi : F_1 \to F_2$ such that $\Phi(0) = 0$, $\Phi(1) = 1$, $\Phi(x + y) = \Phi(x) + \Phi(y)$, $\Phi(xy) = \Phi(x)\Phi(y)$ and $x \leq y$ if and only if $\Phi(x) \leq \Phi(y)$.*

The proof of the theorem is rather difficult. A construction of a complete ordered field is briefly described in Section 6.5.

Definition. A complete ordered field is called the set of real numbers.

Now we can prove that $\sqrt{2}$ exists. Actually, we will prove a more general result.

Theorem 6.14. *For any $x \geq 0$ and $n \in \mathbb{N}$ there is one and only one $y \geq 0$ such that $y^n = x$.*

This unique number y is denoted by $y = \sqrt[n]{x} = x^{1/n}$. In the proof we will need the following lemma.

Proof. First we will prove that if a nonnegative solution of the equation $y^n = x$ exists, then it is unique i.e., if $y_1^n = y_2^n = x$, then $y_1 = y_2$. If $x = 0$, then clearly $y_1 = y_2 = 0$. Thus assume that $x > 0$ and hence $y_1, y_2 > 0$. By Theorem 4.10 we have

$$0 = y_1^n - y_2^n = (y_1 - y_2)(y_1^{n-1} + y_1^{n-2}y_2 + \ldots + y_2^{n-1}).$$

Since the second factor on the right hand side is positive, the first one must be zero and hence $y_1 = y_2$.

We are left with the proof of the existence of a solution. Assume first that $x > 1$. Consider the set

$$S = \{z > 0 : \, z^n \leq x\}.$$

If $z > x$, then $z^n > x^n > x$ (because $x > 1$) and hence $z \notin S$. Thus all elements $z \in S$ satisfy $z \leq x$ and hence S is bounded from above by x. Since $1 \in S$ we have $S \neq \emptyset$ and thus

$$y = \sup S \in \mathbb{R}.$$

We have

$$y^n < x \quad \text{or} \quad y^n > x \quad \text{or} \quad y^n = x.$$

It suffices to show that the first and the second possibility lead to a contradiction. Then we will necessarily have $y^n = x$.

Suppose that $y^n < x$. If $\varepsilon \in (0, x)$, then

$$(y+\varepsilon)^n - y^n = \varepsilon \cdot \left((y+\varepsilon)^{n-1} + (y+\varepsilon)^{n-2}y + \ldots + y^{n-1}\right) < \varepsilon \cdot n(y+x)^{n-1}.$$

Taking

$$0 < \varepsilon < \frac{x - y^n}{n(y+x)^{n-1}}$$

we have

$$(y+\varepsilon)^n - y^n < x - y^n, \qquad (y+\varepsilon)^n < x$$

and hence $y + \varepsilon \in S$. Thus $y + \varepsilon \leq \sup S = y$ which is a contradiction.

Suppose now that $y^n > x$. For $y > \varepsilon > 0$ we have

$$y^n - (y-\varepsilon)^n = \varepsilon \cdot \left(y^{n-1} + y^{n-2}(y-\varepsilon) + \ldots + (y-\varepsilon)^{n-1}\right) < \varepsilon \cdot ny^{n-1}.$$

Taking

$$0 < \varepsilon < \min\left\{y, \frac{y^n - x}{ny^{n-1}}\right\}$$

we have $y^n - (y-\varepsilon)^n < y^n - x$ i.e., $(y-\varepsilon)^n > x$. Since $x \geq z^n$ for all $z \in S$ we obtain that $(y-\varepsilon)^n > z$ for all $z \in S$ i.e., $y - \varepsilon$ is an upper bound of S which contradicts the fact that y is the least upper bound. The proof in the case $x > 1$ is complete.

If $x = 0$, then $y = 0$. If $x = 1$, then $y = 1$. If $0 < x < 1$, then $1/x > 1$ and hence there is $w > 0$ such that $w^n = 1/x$. Now $y = 1/w$ satisfies $y^n = x$. □

In the definition of the complete ordered field we only discuss the existence of supremum for sets bounded from above, but it turns out that this condition also implies that every set bounded from below has infimum. Indeed, If $A \subset \mathbb{F}$ is bounded from below, then the set $A' = \{-x : x \in A\}$ is bounded from above, so $M = \sup A'$ exists. Then it is easy to see that $-M = \inf A$.

Exercise 6.15. *Provide a rigorous proof that A' is bounded from above and that $-M = \inf A$.*

We have seen that $\mathbb{Q}$ is an ordered field and $\mathbb{Q}$ can be easily defined with the help of integers. Unfortunately $\mathbb{Q}$ is not complete. How do we know that there exists a complete ordered field? We cannot take this fact for granted. It requires a proof. Namely we have to construct an ordered field. The construction and then verification that it is a complete ordered field is actually quite technical and boring. Very boring indeed. Thus we will only provide a sketch.

We start with the field of rational numbers $\mathbb{Q}$ and we need to add all "missing" real numbers to make the field complete. How do we define real numbers that are not rational?

6.4 Natural and rational numbers among real numbers

We know that natural numbers $\mathbb{N}$, integers $\mathbb{Z}$ and rational numbers $\mathbb{Q}$ form subsets of $\mathbb{R}$. The first property that we are going to prove is that for any real number x we can find a natural number that is larger than x.

Theorem 6.16. *The set $\mathbb{N} \subset \mathbb{R}$ is unbounded from above.*

Proof. Suppose that $\mathbb{N}$ is bounded from above. Since it is not empty, $M = \sup \mathbb{N} \in \mathbb{R}$ exists. Thus

$$n \leq M \quad \text{for all } n \in \mathbb{N}.$$

Hence also

$$n + 1 \leq M \quad \text{for all } n \in \mathbb{N},$$

so

$$n \leq M - 1 \quad \text{for all } n \in \mathbb{N},$$

and therefore $M - 1$ is an upper bound of $\mathbb{N}$. This is however, a contradiction with the assumption that M is the least upper bound. □

Theorem 6.17 (Archimedes Postulate)**.** [7] *For any positive real numbers $a > 0$ and $b \in \mathbb{R}$ there is a natural number $n \in \mathbb{N}$ such that $an > b$.*

Proof. By contrary if $na \leq b$ for all $n \in \mathbb{N}$, then $\mathbb{N}$ is bounded from above by b/a which is a contradiction. □

The next result states that rational numbers are dense in $\mathbb{R}$.

Theorem 6.18. *For any real numbers $x < y$ there is a rational number q such that $x < q < y$.*

Proof. Since $y - x > 0$ it follows from the previous result that $n(y - x) > 1$ for some $n \in \mathbb{N}$ and hence $1 + nx < ny$. Another application of Theorem 6.17 gives the existence of $m_1, m_2 \in \mathbb{N}$ such that

$$m_1 > nx \quad \text{and} \quad m_2 > -nx.$$

Hence $1 < nx + m_2 + 1 < m_1 + m_2 + 1$. By Well-Ordering Principle (Theorem 4.13) there is a smallest integer $m_3 \in \mathbb{N}$ such that $1 < nx + m_2 + 1 < m_3$. Since $m_3 - 1 \in \mathbb{N}$ is smaller than m_3 we clearly have

$$m_3 - 1 \leq nx + m_2 + 1 < m_3,$$

i.e.,

$$m - 1 \leq nx < m,$$

where $m = m_3 - m_2 - 1 \in \mathbb{Z}$. Hence

$$nx < m \leq 1 + nx < ny$$

and since $n > 0$ it follows that

$$x < \frac{m}{n} < y\,.$$

[7] In Archimedes' words: Any magnitude when added to itself enough times will exceed any given magnitude.

The proof is complete. □

One can also easily prove that the set of irrational numbers is dense in $\mathbb{R}$. To prove this it suffices to consider numbers of the form $q\sqrt{2}$, where $q \in \mathbb{Q}$. We leave details to the reader.

Theorem 6.19. *For any real numbers $x < y$ there is an irrational number z such that $x < z < y$.*

6.5 Formal construction of real numbers

Sketch of the construction. In the proof we have to construct real numbers out or rational ones. Suppose for a moment that we already know what the real numbers are and we ask to identify all real numbers using rational numbers only. Every real number α defines a set of rational numbers strictly less than than α

$$A_\alpha = \{p \in \mathbb{Q} : p < \alpha\} \tag{6.7}$$

and such a set uniquely determines the real number α. Thus the idea is to identify real numbers with certain subsets of rational numbers. For example the set

$$\{p \in \mathbb{Q} : p^2 < 2\}$$

will determine the real number α with the property $\alpha^2 = 2$.

Observe that the sets A_α have the following properties

1. $\emptyset \neq A_\alpha \neq \mathbb{Q}$.
2. If $p \in A_\alpha$ and $\mathbb{Q} \ni q < p$, then $q \in A_\alpha$.
3. If $p \in A_\alpha$, then there is $p < r \in \mathbb{Q}$ such that $r \in A_\alpha$.

The above reasoning is not a part of proof, because we already assumed that the real numbers exist. However, it provides a motivation for what is discussed below.

Thus we no longer have real numbers, only rational numbers and we want to construct real numbers. The construction is as follows.

We define the set $\mathbb{R}$ to be the family of subsets of $\mathbb{Q}$ called *cuts*. A cut is a subset $\alpha \subset \mathbb{Q}$ such that

1. α is not empty and $\alpha \neq \mathbb{Q}$.
2. If $p \in \alpha$, $q \in \mathbb{Q}$ and $q < p$, then $q \in \alpha$.
3. If $p \in \alpha$, then $p < r$ for some $r \in \alpha$.

This is a reasonable construction, because the set defined in (6.7) has all the properties listed here.

We will denote rational numbers by letters p, q, r,... and cuts by Greek letters α, β, γ,...

Thus we identify the set of all real numbers $\mathbb{R}$ with cuts. Since cuts do not look like numbers, we have to define the operations of addition, multiplication and the inequality for cuts to make them numbers. Moreover we have to identify rational numbers with certain cuts.

If $\alpha, \beta \in \mathbb{R}$, then we define

$$\alpha + \beta = \{p + q : \, p \in \alpha \text{ and } q \in \beta\}\,.$$

Thus $\alpha + \beta$ is a subset of $\mathbb{Q}$ and one is required to prove that it is a cut. We leave it to the reader.

If $\alpha, \beta \in \mathbb{R}$, we say that $\alpha \leq \beta$, if $\alpha \subset \beta$.

Since elements of $\mathbb{R}$ are sets and elements of $\mathbb{Q}$ are numbers we need to identify rational numbers with cuts to have $\mathbb{Q} \subset \mathbb{R}$. Namely any $q \in \mathbb{Q}$ is identified with the cut

$$q^* = \{r \in \mathbb{Q} : \, r < q\}.$$

In particular 0 is identifies with

$$0^* = \{r \in \mathbb{Q} : \, r < 0\}.$$

Now we can also define multiplication, but it is slightly more complicated. For cuts $\alpha \geq 0^*$ and $\beta \geq 0^*$ we can define

$$\alpha \cdot \beta = \{pq : \, p \in \alpha \text{ and } q \in \beta\}$$

This corresponds to the multiplication of nonnegative numbers and one can think how to define multiplication of other real numbers. We leave it to the reader.

Now $\mathbb{R}$ is equipped with the operations "+", "$\cdot$", the inequality "$\leq$" and $\mathbb{Q} \subset \mathbb{R}$. One can prove that $\mathbb{R}$ with the operations and the inequality is an ordered field. We will not show it (too boring). It still remains to prove that $\mathbb{R}$ is complete., i.e., we have to show that any set $\emptyset \neq A \subset \mathbb{R}$ bounded from above has a supremum. Here is a proof.

Elements $\alpha \in A$ are non-empty sets of rational numbers (cuts), and hence the union of all sets $\alpha \in A$,

$$\gamma = \bigcup_{\alpha \in A} \alpha \neq \emptyset$$

is also a non-empty set of rational numbers.

Let $\beta \in \mathbb{R}$ be any upper bound of A, i.e., $\alpha \leq \beta$ for all $\alpha \in A$, i.e., $\alpha \subset \beta$ for all $\alpha \in A$. Hence the union γ of all sets $\alpha \in A$ also satisfies $\gamma \subset \beta$. In particular $\gamma \neq \mathbb{Q}$.

We proved that $\gamma \neq \emptyset$ and $\gamma \neq \mathbb{Q}$. This is the first property of a cut. One can easily prove that γ satisfies the other two properties., i.e., γ is a cut, i.e., $\gamma \in \mathbb{R}$.

We claim that γ is an upper bound of A. Indeed, since γ is the union of all sets $\alpha \in A$, $\alpha \subset \gamma$ for all $\alpha \in A$, i.e., $\alpha \leq \gamma$ for all $\alpha \in A$. On the other hand we already proved that for any upper bound β, $\gamma \subset \beta$, i.e., $\gamma \leq \beta$, so γ is the least upper bound of A. In other words $\gamma = \sup A$. The proof is complete. □

The proof presented above is very abstract and it is difficult to follow the construction. To have a better imagination of what happens here one can draw pictures

on the margin of the proof to interpret the cuts α and then all the steps of the construction will have a clear interpretation with respect to such pictures.

Form now one we will only need to know that $\mathbb{R}$, the complete ordered field that contains $\mathbb{Q}$ exists and we do not have to know how it was constructed. The construction was only to make sure that such a field exists. If we want to use a certain mathematical object that satisfies certain properties (like $\mathbb{R}$) we cannot take the existence of it for granted. Otherwise it could happen that the properties have some permanent logical error and we can arrive to a contradiction like in the case of Russell's paradox. In the case of real numbers the completeness of the field is a highly nontrivial and non obvious property, so we really had to prove the existence such a field.

One more remark is that there are many very different fields however, one can prove that in a certain sense the complete ordered field is unique. All complete ordered fields are isomorphic which roughly speaking means they are identical.

More precisely if $\mathbb{F}_1$ and $\mathbb{F}_2$ are complete ordered fields, then there is a bijection $f : \mathbb{F}_1 \to \mathbb{R}_2$ such that

$$f(x_1 + x_2) = f(x_1) + f(x_2), \quad f(x_1 x_2) = f(x_1) f(x_2),$$

$$x_1 \leq x_2 \text{ if and only if } f(x_1) \leq f(x_2).$$

Such a bijection is called an *isomorphism* of $\mathbb{F}_1$ and $\mathbb{F}_2$. We will not prove this fact. Observe that $f(0) = 0$ and $f(1) = 1$.

Chapter 7

Cardinality

What does it mean that a set A has n elements? That mean that we can label the elements of the set with numbers $1, 2, \ldots, n$. Being more precise that means that there is a bijection

$$f : \{1, 2, \ldots, n\} \to A.$$

It is also easy to see that two finite sets A and B have the same number of elements if and only if there is a bijection $f : A \to B$.

A set A is *finite* if there is a natural number n such that A has n elements. Sets that are not finite are called *infinite* sets.

We say that two sets (finite or infinite) A and B have the same *cardinality* if there is a bijection $f : A \to B$. As we already observed in the case of finite sets they have the same cardinality if and only if they have the same number of elements. Thus the intuition should be that two infinite sets have the same cardinality if they have the same amount of elements.

Proposition 7.1. *If A has the same cardinality as B and B has the same cardinality as C, then A has the same cardinality as C.*

Proof. That is obvious. If $f : A \to B$ is a bijection and $g : B \to C$ is a bijection, then $g \circ f : A \to C$ is a bijection. □

Proposition 7.2. *The set of all natural numbers has the same cardinality as the set of all positive even numbers.*

Proof. One could think that the set of all natural numbers is larger, because it contains all positive even integers but also all positive odd integers, so it should be twice as large. However, the two sets have the same cardinality, because the function

$$\varphi : \{1, 2, 3, 4, \ldots\} \to \{2, 4, 6, 8, \ldots\}, \qquad \varphi(n) = 2n$$

is the bijection. This is one of the antinomies that are persistent when we deal with infinity.
□

We say that a set A is *denumerable* if it has the same cardinality as the set of natural numbers $\mathbb{N}$. A set A is *countable* if it is finite or denumerable. A set A is *uncountable* if it is not countable.

Proposition 7.2 says that the set of positive integers is denumerable.

Proposition 7.3. *$\mathbb{Z}$ has the same cardinality as $\mathbb{N}$, i.e., $\mathbb{Z}$ is denumerable.*

Proof. It suffices to show that we can arrange elements of $\mathbb{Z}$ into a sequence, but that is easy.

$$0, 1, -1, 2, -2, 3, -3, 4, -4, \ldots$$

Then the bijection between $\mathbb{N}$ and $\mathbb{Z}$ is given by

$$\begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 9 & 9 & \ldots \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ 0 & 1 & \text{-1} & 2 & \text{-2} & 3 & \text{-3} & 4 & \text{-4} & \ldots \end{array}$$

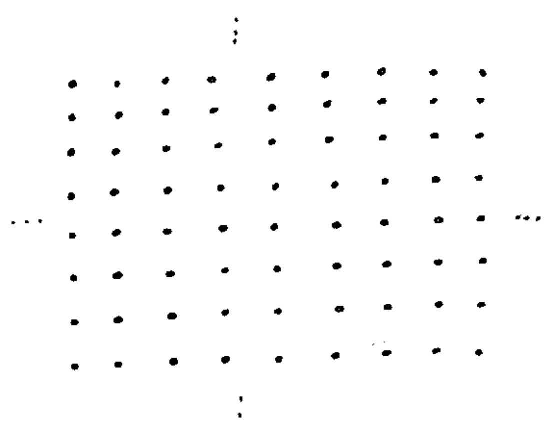
$\square$

We defined the set of natural numbers using the set theory, and me mentioned that it is not difficult to define integers and rational numbers. However, we do not know yet how to define real numbers. We will provide a formal definition in Section 6, but in the next few examples we will assume familiarity with real numbers and decimal expansions.

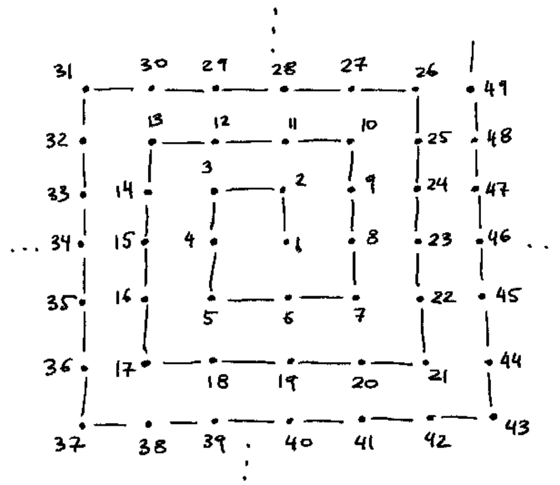
$\mathbb{Z}^2$ is the set of all points in $\mathbb{R}^2$ with both coordinates being integers. It is called *integer lattice*.

Proposition 7.4. *$\mathbb{Z}^2$ is denumerable.*

Proof. The set $\mathbb{Z}^2$ looks as follows



and the next picture shows how to find a bijection between $\mathbb{N}$ and $\mathbb{Z}^2$.



□

Proposition 7.5. *If A is denumerable and $B \subset A$ is an infinite subset, then B is denumerable.*

Proof. The fact that A has the same cardinality as $\mathbb{N}$ means that elements of A can be arranged into a sequence. Indeed, if $f : \mathbb{N} \to A$ is a bijection, and $a_n = f(n)$, then

$$a_1,\, a_2,\, a_3,\, a_4, \ldots$$

is such an arrangement. If $B \subset A$ is an infinite set, then we can arrange the set into an infinite sequence by erasing those elements from the above sequence that do not belong to A. Since elements of B can be arranged into a sequence the set B has the same cardinality as $\mathbb{N}$. □

Proposition 7.6. *The set of rational numbers is denumerable.*

Proof. Each rational number can uniquely be represented in the form $\pm p/q$, where $p, q \in \mathbb{N}$ are relatively prime. Therefore with each rational number we can associate a unique element $(\pm p, q)$ in $\mathbb{Z}^2$. However, not every point in $\mathbb{Z}^2$ is associated with a rational number. For example the point $(2, 4)$ will not be associated with any rational number, because the numbers 2 and 4 are not relatively prime.

Thus there is a bijection between $\mathbb{Q}$ and an infinite subset of $\mathbb{Z}^2$. Since $\mathbb{Z}^2$ has the same cardinality as $\mathbb{N}$, any infinite subset of $\mathbb{Z}^2$ has the same cardinality as $\mathbb{N}$ (Proposition 7.5). In particular $\mathbb{Q}$ has the same cardinality as $\mathbb{N}$. □

Not every set is however, countable.

Theorem 7.7 (Cantor)**.** *The set $\mathbb{R}$ of all real numbers is uncountable.*

Proof. It suffices to prove that the subset $(0, 1) \subset \mathbb{R}$ is uncountable.[1] Suppose $(0, 1)$ is countable, so we can arrange all real numbers from $(0, 1)$ into a sequence. Suppose

$$x_1,\, x_2,\, x_3,\, x_4, \ldots$$

[1] Indeed, countability of $\mathbb{R}$ would imply countability of $(0, 1)$, see Proposition 7.5.

is such a sequence.

Note that the decimal expansion of a real number is not necessarily unique. For example

$$1 = 0.9999\ldots$$

Hence

$$0.001 = 0.0009999\ldots$$

and thus

$$0.237 = 0.2369999\ldots$$

This also shows that every real number $x > 0$ has a decimal expansion in which infinitely many digits are different than zero, and one can prove (we skip the proof) that such an expansion is unique. Let

$$x_1 = 0.a_{11}a_{12}a_{13}a_{14}\ldots$$

$$x_2 = 0.a_{21}a_{22}a_{23}a_{24}\ldots$$

$$x_3 = 0.a_{31}a_{32}a_{33}a_{34}\ldots$$

$$\ldots$$

be the sequence of decimal expansions of numbers $x_1, x_2, x_3, \ldots$ with infinite number of non-zero digits. Here a_{ij} represents jth digit of x_i.

For every n let $b_n \in \{1, \ldots, 8\}$ be a digit different than a_{nn}. Then

$$x = 0.b_1b_2b_3b_4\ldots \in (0, 1)$$

However, x does not appear in the sequence $x_1, x_2, x_3 \ldots$, because the decimal expansion of x differs from that of x_n on nth position. Therefore the sequence $x_1, x_2, x_3, \ldots$ cannot list all real numbers from $(0, 1)$. This contradiction completes the proof. □

It is interesting to investigate what other sets have the same cardinality as $\mathbb{R}$.

Proposition 7.8. *Any finite interval (a, b) has the same cardinality as $\mathbb{R}$.*

Proof. The function $f(x) = \arctan x$ gives a bijection between $(-\pi/2, \pi/2)$ and $\mathbb{R}$. Therefore $(-\pi/2, \pi/2)$ has the same cardinality as $\mathbb{R}$, but (a, b) has the same cardinality as $(-\pi/2, \pi/2)$, because we may easily find a linear function that maps (a, b) onto $(-\pi/2, \pi/2)$ in a bijective way and then the result follows from Proposition 7.1. □

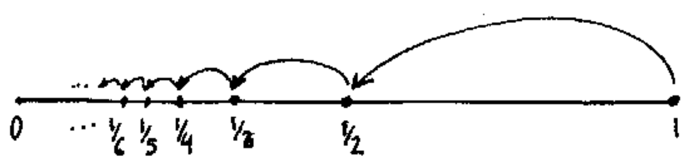
Now a tricky problem.

Example 7.9. *Prove that $(0, 1]$ has the same cardinality as $(0, 1)$.*

Proof. It is not obvious how to start. If we try to build a bijection from $(0, 1]$ to $(0, 1)$ it is not clear what to do with the endpoint 1. Actually it is not possible to construct a continuous bijection. One can prove that every continuous bijection defined on $(0, 1]$ will have an interval with an endpoint included as an image, so it cannot work. Thus we will construct a discontinuous bijection $f : (0,1] \to (0,1)$. We define it by a formula

$$f\left(\frac{1}{n}\right) = \frac{1}{n+1} \quad \text{for } n = 1, 2, 3, \ldots$$

$$f(x) = x \quad \text{for } x \neq \tfrac{1}{n}.$$



We just shift the points of the form $1/n$ to the left and do not move other points. It takes a moment to see that f is a bijection, but if you look at the picture for a while it should be obvious and no further explanations are needed. □

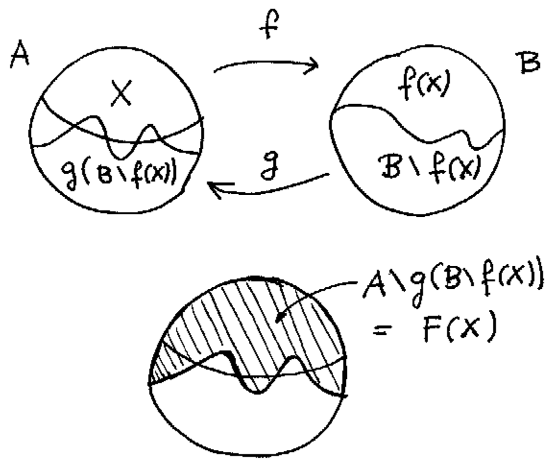
Theorem 7.10 (Cantor-Bernstein). *If A has the same cardinality as a subset $A' \subset B$ and B has the same cardinality as a subset $B' \subset A$, then A has the same cardinality as B.*

Intuitively, the theorem seems clear. Since A has the same cardinality as a subset of B, the amount of elements in the set B should be greater than or equal to the amount of elements in A. Since B has the same cardinality as a subset of A, the amount of elements in A should be greater than or equal to the amount of elements in B, so the amount of elements in sets A and B should be the same. However, since we talk about infinite sets nothing is obvious. The amount of elements in infinite sets is not a number and there is reason to claim that $b \geq a$, $a \geq b$ implies $b = a$, because a and b (amounts of elements in sets A and B) are not numbers. By the assumption of the theorem we have bijections $f : A \to A' \subset B$ and $g : B \to B' \subset A$, but we need to construct a bijection $h : A \to B$ and it is not obvious at all how to do it. The theorem has many non-trivial applications. Consider $A = (0,1)$ and $B = (0,1]$. $f : A \to B$, $f(x) = x$ is a bijection between $A = (0,1)$ and $A' = (0,1) \subset (0,1] = B$ and $g : B \to A$, $g(x) = x/2$ is a bijection between $B = (0,1]$ and $B' = (0,1/2] \subset (0,1) = A$. Thus according to the Cantor-Bernstein theorem the sets $A = (0,1)$ and $B = (0,1]$ have the same cardinality, but we have already seen that proving directly that the two sets have the same cardinality is not an easy task, so the Cantor-Barnstein theorem cannot be easy.

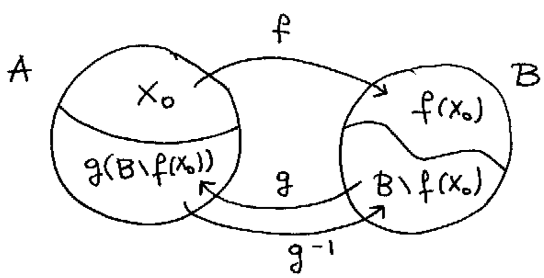
Proof. Let $f : A \to A' \subset B$ and $g : B \to B' \subset A$ be bijections. For a given set $X \subset A$ we define a subset $F(X) \subset A$ by

$$F(X) := A \setminus g(B \setminus f(X)).$$

The construction of the set $F(X)$ is explained on the picture



Suppose that there is a set $X_0 \subset A$ such that $F(X_0) = X_0$, i.e., X_0 is a "fixed point" of F. In this case the situation is as on the picture



and hence the function

$$h(x) = \begin{cases} f(x) & \text{if } x \in X_0, \\ g^{-1}(x) & \text{if } x \notin X_0, \end{cases}$$

defines a bijection of A onto B. Thus it remains to show that there is a set X_0 such that $F(X_0) = X_0$.

Lemma 7.11. *For any family $X_1, X_2, X_3, \ldots$ of subsets of A we have*

$$F\left(\bigcap_{i=1}^{\infty} X_i\right) = \bigcap_{i=1}^{\infty} F(X_i)\,.$$

Proof. First compare the lemma with Exercise **??**. Does the exercise apply here? Is F one-to-one? That is a completely wrong point of view. Observe that F is *not* a function defined on A, because F is not defined for points in A, but for subsets of A, so we cannot really compare the lemma with the exercise. However, as we will see the exercise will be employed in the proof.

We have

$$\begin{aligned}
F\left(\bigcap_{i=1}^{\infty} X_i\right) &= A \setminus g\left(B \setminus f\left(\bigcap_{i=1}^{\infty} X_i\right)\right) \\
&= A \setminus g\left(B \setminus \bigcap_{i=1}^{\infty} f(X_i)\right) \quad \text{(Exercise \textbf{??}, since } f \text{ is one-to-one)} \\
&= A \setminus g\left(\bigcup_{i=1}^{\infty} (B \setminus f(X_i)\right) \quad \text{(De Morgan Law, Proposition 2.7)} \\
&= A \setminus \bigcup_{i=1}^{\infty} g(B \setminus f(X_i)) \quad \text{(Exercise \textbf{??})} \\
&= \bigcap_{i=1}^{\infty} (A \setminus g(B \setminus f(X_i)) \quad \text{(De Morgan Law, Proposition 2.7)} \\
&= \bigcap_{i=1}^{\infty} F(X_i)\,.
\end{aligned}$$

This completes the proof of the lemma. □

Observe that we have a decreasing sequence of sets (Why is it decreasing?)

$$A \supset F(A) \supset F^2(A) \supset F^3(A) \supset \dots,$$

where $F^k(A) = F(F(\dots(F(A))\dots))$. Define

$$X_0 = \underbrace{A}_{X_1} \cap \underbrace{F(A)}_{X_2} \cap \underbrace{F^2(A)}_{X_3} \cap \underbrace{F^3(A)}_{X_4} \cap \dots = \bigcap_{i=1}^{\infty} X_i\,.$$

Observe also that

$$X_0 = A \cap F(A) \cap F^2(A) \cap F^3(A) \cap \dots = F(A) \cap F^2(A) \cap F^3(A) \cap \dots$$

because if B is a subset of A, then $A \cap B = B$. Now the lemma immediately yields

$$F(X_0) = \bigcap_{i=1}^{\infty} F(X_i) = F(A) \cap F^2(A) \cap F^3(A) \cap F^4(A) \cap \dots = X_0$$

and the theorem follows. □

The plane $\mathbb{R}^2$ can be described as the set of all ordered pairs of real numbers

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

The following result is quite surprising.

Theorem 7.12. $\mathbb{R}^2$ *has the same cardinality as* $\mathbb{R}$.

Proof. According to the Cantor-Bernstein theorem it suffices to find a bijection $f : \mathbb{R} \to A \subset \mathbb{R}^2$ and a bijection $g : \mathbb{R}^2 \to B \subset \mathbb{R}$. The first of the two bijections is easy to construct.

$$f(x) = (x, 0)$$

is a bijection of $\mathbb{R}$ onto the x-axis in $\mathbb{R}^2$. The second one is more difficult to construct. In the first step we simplify the problem by observing that $\mathbb{R}^2$ has the same cardinality as the open unit square

$$Q = \{(x, y) : 0 < x, y < 1\} .$$

Indeed, in Proposition 7.8 we constructed a bijection $h : \mathbb{R} \to (0, 1)$ and hence

$$H(x, y) = (h(x), h(y))$$

is a bijection $H : \mathbb{R}^2 \to Q$. Therefore it suffices to construct a bijection $w : Q \to A \subset \mathbb{R}$, because, then $g = w \circ H : \mathbb{R}^2 \to A \subset \mathbb{R}$ will a bijection that we need.

Let $(x, y) \in Q$, i.e., $x \in (0, 1)$ and $y \in (0, 1)$. Consider the unique decimal expansions of x and y with infinite non-zero digits (cf. proof of Theorem 7.7)

$$x = 0.a_1 a_1 a_3 \ldots$$

$$y = 0.b_1 b_2 b_2 \ldots$$

and define

$$H(x, y) = 0.a_1 b_1 a_2 b_2 a_3 b_3 \ldots$$

Namely $H(x, y) \in (0, 1)$ is a number obtained from x and y by mixing the digits of x and y. It is easy to see that the function H is one-to-one and that completes the proof. □

Exercise 7.13. *Prove that $H : Q \to (0, 1)$ is not a surjection.*

When we look at the above proof carefully, we see that the result is not that surprising, after all. Points in $\mathbb{R}^2$ are ordered pairs of numbers (x, y) and hence can be encoded by two sequences of digits (decimal expansions) and points in $\mathbb{R}$ can be encoded by one sequence of digits. Now the main idea in the proof is the observation that two sequences of digits can be encoded in a single sequence of digits – by a suitable mixing of the digits from the two given sequences.

We have seen that $\mathbb{R}$ is uncountable, i.e., it has different cardinality than $\mathbb{N}$. The following results provides another method of constructing infinite sets with different cardinalities.

Theorem 7.14 (Cantor)**.** *There is no bijection between a set A and its power set $P(A)$.*

Recall that $P(A)$ is a family of all subsets of A. If A has n elements, then $P(A)$ has 2^n elements, so clearly there is no bijection, because $2^n > n$. The result is however, not obvious if the sets are infinite.

Proof. Suppose $\varphi : A \to P(A)$ is a bijection. Since $\varphi(a)$ is a subset of A, either a is an element of $\varphi(a)$ or a is not an element of $\varphi(a)$. Let

$$R = \{a \in A : a \notin \varphi(a)\} . \tag{7.1}$$

This set is well defined, since its existence is guaranteed by the Axiom of Specification. R is a subset of A, so $R = \varphi(a_0)$ for some $a_0 \in A$ (because φ is a bijection). We may inquire now whether $a_0 \in R$ or $a_0 \notin R$. If $a_0 \in R$, then a_0 must satisfy the condition from the definition (7.1) of the set R. Hence $a_0 \notin \varphi(a_0) = R$ which is a contradiction. If $a_0 \notin R = \varphi(a_0)$, then the condition from the definition (7.1) is satisfied and hence $a_0 \in R$, a contradiction again. Every possibility leads to a contradiction. That proves that a bijection $\varphi : A \to P(A)$ does not exist. □

Observe that the above proof is remarkably similar to the argument used in Russell's paradox. This time however, it is not a paradox – it is a rigorous mathematical argument.

Cantor's theorem implies, in particular, that the two infinite sets $\mathbb{N}$ and $P(\mathbb{N})$ are not of the same cardinality. That means a collection of all subsets of $\mathbb{N}$ cannot be arranged as a sequence, in other words, the set $P(\mathbb{N})$ is uncountable.

Exercise 7.15. *Prove that $P(\mathbb{N})$ has the same cardinality as $\mathbb{R}$.*

Note that there is a one-to-one function

$$f : A \to P(A), \qquad f(a) = \{a\}$$

but there is no one-to-one function $g : P(A) \to A$, otherwise the sets A and $P(A)$ would have the same cardinality (Cantor-Bernstein). Thus in some sense $P(A)$ has a "larger" amount of elements than A. Then also $P(P(A))$ has "larger" amount of elements than $P(A)$ and so on.

Actually one can associate with any set A a *cardinal number* $\mathfrak{n}$ which is something like the "amount" of elements in the set A. Two sets have the same cardinal numbers if they have the same cardinality. If A is a finite set, then its cardinal number is the number of elements in A. If the cardinal number of A is $\mathfrak{n}$ and the cardinal number of B is $\mathfrak{m}$, we write $\mathfrak{n} \geq \mathfrak{m}$ is there is a one-to-one function $f : B \to A$. If $\mathfrak{n} \geq \mathfrak{m}$ and $\mathfrak{m} \geq \mathfrak{n}$, then $\mathfrak{n} = \mathfrak{m}$. Indeed, it is a direct consequence of the Cantor-Bernstein theorem. If $\mathfrak{n} \geq \mathfrak{m}$, but $\mathfrak{n} \neq \mathfrak{m}$ we write $\mathfrak{n} > \mathfrak{m}$.

If A has the cardinal number $\mathfrak{n}$, the cardinal number of $P(A)$ is denoted[2] by $2^{\mathfrak{n}}$. Since $f : A \to P(A)$, $f(a) = \{a\}$ is one-to-one, $2^{\mathfrak{n}} \geq \mathfrak{n}$. Since, according to Cantor's theorem, A and $P(A)$ do not have the same cardinality, $2^{\mathfrak{n}} > \mathfrak{n}$. In particular for any cardinal number we can find a bigger cardinal number. The largest cardinal number does not exist.

Given two cardinal numbers $\mathfrak{n}$ and $\mathfrak{m}$, at most one condition is satisfied

$$\mathfrak{n} < \mathfrak{m}, \qquad \mathfrak{n} = \mathfrak{m}, \qquad \mathfrak{n} > \mathfrak{m}.$$

One can prove that for any two cardinal numbers $\mathfrak{n}$ and $\mathfrak{m}$, $\mathfrak{n} < \mathfrak{m}$, or $\mathfrak{n} = \mathfrak{m}$, or $\mathfrak{n} > \mathfrak{m}$. This result is actually difficult to prove. Indeed, it exactly means that for any two sets A and B there is a one-to-one function $f : A \to B$, or a bijection $f : A \to B$ or a one-to one function $g : B \to A$. How can we assure that such a function exists if we do not know anything about the sets A and B? The proof is based on the Axiom of Choice.

[2]This notation is consistent with the case of finite sets: if A has n elements, the set $P(A)$ has 2^n elements.

Chapter 8

Sequences and limits

In this chapter we will discuss basic result regarding limits of sequences and sums of infinite series.

8.1 Limits of sequences

A sequence $(a_n) = (a_n)_{n=1}^{\infty}$ is a collection of real numbers

$$a_1,\ a_2,\ a_3,\ a_4, \dots \tag{8.1}$$

where the order of elements is important. More precisely we can define a sequence as a function $f : \mathbb{N} \to \mathbb{R}$. Namely, the sequence (8.1) is identified with a function

$$f : \mathbb{N} \to \mathbb{R}, \quad f(n) = a_n.$$

We say that a sequence (a_n) is

increasing if $a_1 \le a_2 \le a_3 \le \dots$,

decreasing if $a_1 \ge a_2 \ge a_3 \ge \dots$,

strictly increasing if $a_1 < a_2 < a_3 < \dots$,

strictly decreasing if $a_1 > a_2 > a_3 > \dots$,

monotone if it is either increasing or decreasing,

strictly monotone if it is either strictly increasing or strictly decreasing.

bounded above if there is $M \in \mathbb{R}$ such that $a_n \le M$ for all n,

bounded below if there is $m \in \mathbb{R}$ such that $a_n \ge m$ for all n,

bounded if there is $M \ge 0$ such that $|a_n| \le M$ for all $n \in \mathbb{N}$.

Clearly, a sequence is bounded if and only if it is bounded above and below.

Roughly speaking we say that a sequence (a_n) converges to $g \in \mathbb{R}$ if the numbers a_n get closer and closer to g as n gets larger and larger. While the intuition behind this description is clear, the notion of convergence has to be given a formal definition.

Let us dissect the intuitive notion of limit more carefully. Suppose that (a_n) converges to $g \in \mathbb{R}$. If $\varepsilon > 0$ is a very small positive number, then for all sufficiently large n, the distance from a_n to g, which equals $|a_n - g|$ will be less than ε i.e. $|a_n - g| < \varepsilon$.

That means if $\varepsilon > 0$ is any number,[1] then we can find n_0 such that for all $n \geq n_0$ we will have $|a_n - g| < \varepsilon$. This leads to the following definition.

Definition. We say that a sequence $(a_n)_{n=1}^{\infty}$ of real numbers *converges* to $g \in \mathbb{R}$ if

$$\forall \varepsilon > 0 \;\; \exists n_0 \in \mathbb{N} \;\; \forall n \geq n_0 \;\; |a_n - g| < \varepsilon\,.$$

We denote this fact by

$$\lim_{n\to\infty} a_n = g \quad \text{or by} \quad a_n \to g \text{ as } n \to \infty.$$

We say that a sequence (a_n) is *convergent* if it converges to some limit $g \in \mathbb{R}$, otherwise the sequence is *divergent*.

Example 8.1. As an exercise in logic, let us write a formal definition of a divergent sequence using quantifiers. A sequence (a_n) is divergent, if for every $g \in \mathbb{R}$ it is not true that (a_n) is convergent to g which can be formally written as

$$\forall g \in \mathbb{R} \;\; \Big(\neg\, (\forall \varepsilon > 0 \;\; \exists n_0 \in \mathbb{N} \;\; \forall n \geq n_0 \;\; |a_n - g| < \varepsilon)\Big)$$

and this is clearly equivalent to

$$\forall g \in \mathbb{R}\; \exists \varepsilon > 0 \;\; \forall n_0 \in \mathbb{N} \;\; \exists n \geq n_0 \;\; |a_n - g| \geq \varepsilon. \tag{8.2}$$

Thus formally we could use (8.2) as a definition of a divergent sequence. However, the expression (8.2) is very complicated and it is much more convenient to define a divergent sequence the way we did before.

In the next four examples we will study limits using directly the definition. Later we will techniques that will greatly simplify the task of finding limits.

Example 8.2. *Prove that* $\lim_{n\to\infty} 1/n = 0$.

Remark 8.3. While this is a very simple example of a limit, the main purpose of it is to learn how a proof of a statement about the limit should be structured.

[1] While we want to think that $\varepsilon > 0$ is a very small number, we put the condition that $\varepsilon > 0$ is any number: if the inequality $|a_n - g| < \varepsilon$ is true for small ε, then it is certainly true for all larger values of ε so no harm is done if we require the condition for all $\varepsilon > 0$.

Proof. We need to prove that

$$\forall \varepsilon > 0 \;\; \exists n_0 \in \mathbb{N} \;\; \forall n \geq n_0 \quad \left| \frac{1}{n} - 0 \right| < \varepsilon \,.$$

Let $\varepsilon > 0$ be given.[2] Let n_0 be a natural number such that $n_0 > 1/\varepsilon$.[3] Then for any $n \geq n_0$ we have[4]

$$n \geq n_0 > \frac{1}{\varepsilon}, \quad n > \frac{1}{\varepsilon}, \quad \frac{1}{n} < \varepsilon, \quad \left| \frac{1}{n} - 0 \right| < \varepsilon.$$

The proof is complete. □

Remark 8.4. One may ask how did we know that $n_0 > 1/\varepsilon$ would be a right choice. In the next two examples we will carefully show the procedure of how to select n_0.

Example 8.5. *Prove that*

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

Finding n_0. Before proving the claim, we will show how to find n_0. Given $\varepsilon > 0$ we want to find $n_0 \in \mathbb{N}$ so that for all $n \geq n_0$, $|n/(n+1) - 1| < \varepsilon$. Thus we start by analyzing the inequality that we want to be true.

$$\left| \frac{n}{n+1} - 1 \right| < \varepsilon, \quad \left| \frac{n-(n+1)}{n+1} \right| < \varepsilon, \quad \frac{1}{n+1} < \varepsilon, \quad n+1 > \frac{1}{\varepsilon}, \quad n > \frac{1}{\varepsilon} - 1.$$

If we choose $n_0 > \frac{1}{\varepsilon} - 1$, then $n \geq n_0$ will certainly satisfy $n > \frac{1}{\varepsilon} - 1$ and by reversing the estimates we will be able to prove the required inequality. This is what we will do below. It is not necessarily to include the above reasoning in the proof since the proof written below will contain all necessary details.

Proof. We need to prove that

$$\forall \varepsilon > 0 \;\; \exists n_0 \in \mathbb{N} \;\; \forall n \geq n_0 \quad \left| \frac{n}{n+1} - 1 \right| < \varepsilon \,.$$

Let $\varepsilon > 0$ be given. Let n_0 be any natural number such that

$$n_0 > \frac{1}{\varepsilon} - 1 \,. \tag{8.3}$$

Then for any $n \geq n_0$ we have

$$n \geq n_0 > \frac{1}{\varepsilon} - 1, \quad n > \frac{1}{\varepsilon} - 1, \quad n+1 > \frac{1}{\varepsilon}, \quad \frac{1}{n+1} < \varepsilon, \quad \left| \frac{n-(n+1)}{n+1} \right| < \varepsilon,$$

$$\left| \frac{n}{n+1} - 1 \right| < \varepsilon \,.$$

The proof is complete. □

[2] We need to prove that the above statement is true for *all* $\varepsilon > 0$ so we choose $\varepsilon > 0$ arbitrarily and then we need to show that $\exists n_0 \in \mathbb{N} \; \forall n \geq n_0 \; |1/n - 0| < \varepsilon$.

[3] Once $\varepsilon > 0$ was chosen, we need to show that there exists $n_0 \in \mathbb{N}$ such that $\forall n \geq n_0 \; |1/n - 0| < \varepsilon$. That means it suffices to show just one value of n_0 such that $\forall n \geq n_0 \; |1/n - 0| < \varepsilon$. Since $\varepsilon > 0$ was selected before selecting n_0, the choice of n_0 may depend on ε.

[4] Since $\varepsilon > 0$ and $n_0 \in \mathbb{N}$ have already been selected, it remains to prove that $\forall n \geq n_0 \; |1/n - 0| < \varepsilon$, and this is what we are doing now.

Remark 8.6. The proof is perfectly fine, even if it does not explain the choice of n_0 in (8.3).[5]

Example 8.7. *Prove that*

$$\lim_{n\to\infty} \frac{3n^2+2}{n^2+2n+1} = 3.$$

Finding n_0. The problem is more difficult thank the two examples discussed above.

We need to find $n_0 \in \mathbb{N}$ such that

$$\left|\frac{3n^2+2}{n^2+2n+1} - 3\right| < \varepsilon \quad \text{for } n \geq n_0.$$

We have[6]

$$\left|\frac{3n^2+2}{n^2+2n+1} - 3\right| = \left|\frac{3n^2+2-3n^2-6n-3}{n^2+2n+1}\right| = \frac{6n+1}{n^2+2n+1} < \frac{6n+1}{n^2} \leq \frac{7n}{n^2} = \frac{7}{n} < \varepsilon.$$

$$\frac{7}{n} < \varepsilon, \quad n > \frac{7}{\varepsilon} \quad \text{and we take} \quad n_0 > \frac{7}{\varepsilon}.$$

The 'finding n_0' part is what we write on a scratch paper and we do not have to include it in the actual proof. We write the proof and we put the scratch paper in trash so a person reading the proof will have no idea why we have chosen $n_0 > 7/\varepsilon$.

Proof. Let $\varepsilon > 0$ be given. Let $n_0 \in \mathbb{N}$ be such that $n_0 > 7/\varepsilon$. Then for $n \geq n_0$ we have $n > 7/\varepsilon$, $7/n < \varepsilon$ so

$$\left|\frac{3n^2+2}{n^2+2n+1} - 3\right| = \left|\frac{3n^2+2-3n^2-6n-3}{n^2+2n+1}\right| = \frac{6n+1}{n^2+2n+1} < \frac{6n+1}{n^2} \leq \frac{7n}{n^2} = \frac{7}{n} < \varepsilon,$$

$$\left|\frac{3n^2+2}{n^2+2n+1} - 3\right| < \varepsilon \quad \text{for all } n \geq n_0.$$

The proof is complete. □

In the next example we will show that the sequence $a_n = (-1)^n$ is divergent. Intuitively, it is absolutely obvious: the sequence oscillates between 1 and -1 infinitely many times so there is no single number $g \in \mathbb{R}$ to which (a_n) would converge. While it seems obvious, writing a formal proof is surprisingly difficult.

Example 8.8. *Prove that the sequence $a_n = (-1)^n$ is divergent.*

Proof. Suppose to the contrary that $\lim_{n\to\infty} a_n = g$ for some $g \in \mathbb{R}$. Then

$$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N} \ \forall n \geq n_0 \quad |a_n - g| < \varepsilon.$$

---

[5] If a student asks me how did I know that (8.3) would be a right choice, I often smile and say: *I often hear voices in my head and one whispered to me: 'take $n_0 > 1/\varepsilon - 1$'.*

[6] Some explanations are necessary. We want $\frac{6n+1}{n^2+2n+1}$ to be less than ε, but the expression is complicated so we estimate it from above by a simpler one $7/n$ so if $7/n < \varepsilon$, then also $\frac{6n+1}{n^2+2n+1} < \varepsilon$.

Since it is true for all $\varepsilon > 0$, it is true for $\varepsilon = 1$. That is there is $n_0 \in \mathbb{R}$ such that

$$|a_n - g| < 1 \quad \text{for all } n \geq n_0.$$

Taking $n \geq n_0$ even, we have

$$|1 - g| < 1.$$

Taking $n \geq n_0$ odd we have

$$|(-1) - g| < 1$$

but the two inequalities lead to a contradiction.[7] Indeed, the triangle inequality $|a + b| \leq |a| + |b|$ yields

$$2 = |1 - (-1)| = |(1 - g) + (g - (-1))| \leq |1 - g| + |g - (-1)| < 1 + 1 = 2,$$

$$2 < 2.$$

The proof is complete. □

Theorem 8.9. *The limit of a sequence is uniquely defined, i.e., if*

$$\lim_{n\to\infty} a_n = g \quad \textit{and} \quad \lim_{n\to\infty} a_n = h$$

then $g = h$*.*

Proof. [8] Although the result seems obvious it requires a proof. Suppose $g \neq h$. Then $\varepsilon = |g - h|/2 > 0$. The definition of the limit with this choice of ε implies that there is n_1 such that

$$|a_n - g| < \varepsilon \quad \text{for all } n \geq n_1$$

and also there is n_2 such that

$$|a_n - h| < \varepsilon \quad \text{for all } n \geq n_2.$$

Hence for all $n \geq \max\{n_1, n_2\}$ we have

$$|g - h| \leq |a_n - g| + |a_n - h| < \varepsilon + \varepsilon = |g - h|$$

which is an obvious contradiction. □

Example 8.10. As we know $\lim_{n\to\infty} a_n = g$, $g \in \mathbb{R}$ means that

$$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N} \ \forall n \geq n_0 \quad |a_n - g| < \varepsilon. \tag{8.4}$$

Given a sequence (a_n) convergent to $g \in \mathbb{R}$, is the following condition also true?

$$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N} \ \forall n \geq n_0 \quad |a_n - g| < \frac{\varepsilon}{3}. \tag{8.5}$$

[7] Geometrically speaking, since the distance between 1 and -1 equals 2, there is no $g \in \mathbb{R}$ such that the distance of g to 1 and to -1 is less than 1.

[8] Note that the argument used in this proof is very similar to the one used in Example 8.8.

Yes, it is! Let $\eta > 0$ be given. Then $\varepsilon = \eta/3 > 0$. Hence according to (8.4), there is $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ $|a_n - g| < \varepsilon = \eta/3$ so we proved that

$$\forall \eta > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 \quad |a_n - g| < \frac{\eta}{3}.$$

Here η is used to denote any positive number and instead of η we can use symbol ε which shows that (8.5) is true. For the same reason the following condition (and many others) is true.

$$\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 \quad |a_n - g| < \varepsilon^2.$$

Example 8.11. *If* $\lim_{n\to\infty} a_n = g \in \mathbb{R}$, *then* $\lim_{n\to\infty} |a_n| = |g|$.

Proof. Recall that $||a| - |b|| \leq |a - b|$ (Theorem 6.6). We need to prove

$$\forall \varepsilon > 0\ \exists n_0\ \forall n \geq n_0 \quad \big||a_n| - |g|\big| < \varepsilon, \tag{8.6}$$

and we know that

$$\forall \varepsilon > 0\ \exists n_0\ \forall n \geq n_0 \quad |a_n - g| < \varepsilon, \tag{8.7}$$

so we prove (8.6) as follows. Let $\varepsilon > 0$ be given. Let n_0 be such that for $n \geq n_0$, $|a_n - g| < \varepsilon$ (such n_0 exists by (8.7)). Then for $n \geq n_0$ we have

$$\big||a_n| - |g|\big| \leq |a_n - g| < \varepsilon,$$

$$\big||a_n| - |g|\big| < \varepsilon.$$

That completes the proof of (8.6). □

Example 8.12. If $a_n = (-1)^n$, then the sequence (a_n) diverges, but $\lim_{n\to\infty} |a_n| = 1$.

However we have

Proposition 8.13. $\lim_{n\to\infty} a_n = 0$ *if and only if* $\lim_{n\to\infty} |a_n| = 0$.

Proof. $\lim_{n\to\infty} a_n = 0$ means that

$$\forall \varepsilon > 0\ \exists n_0\ \forall n \geq n_0 \quad \underbrace{|a_n - 0| < \varepsilon}_{|a_n| < \varepsilon} \tag{8.8}$$

while $\lim_{n\to\infty} |a_n| = 0$ means that

$$\forall \varepsilon > 0\ \exists n_0\ \forall n \geq n_0 \quad \underbrace{||a_n| - 0| < \varepsilon}_{|a_n| < \varepsilon} \tag{8.9}$$

so (8.8) and (8.9) are trivially equivalent. □

Theorem 8.14. *Convergent sequences are bounded.*

Proof. Assume that (a_n) is convergent, $\lim_{n\to\infty} a_n = g$. Taking $\varepsilon = 1$ in the definition of the limit we see that there is n_0 such that

$$|a_n - g| < 1 \quad \text{for all } n \geq n_0.$$

Now

$$|a_n| = |(a_n - g) + g| \leq |a_n - g| + |g| \leq 1 + |g| \quad \text{for all } n \geq n_0.$$

Hence

$$|a_n| \leq \underbrace{1 + |g| + |a_1| + |a_2| + \ldots + |a_{n_0-1}|}_{M}$$

for all n. □

Example 8.15. *Prove that the sequence (a_n), $a_n = n + n(-1)^n$ is divergent.*

Proof. Let $M > 0$ be given. If $n = 2k$, $k > M/4$, then $a_n = 4k > M$ so the sequnece is unbounded above and therefore, divergent (by Theorem 8.14). □

The following result is of fundamental importance.

Theorem 8.16. *If*

$$\lim_{n\to\infty} a_n = a \in \mathbb{R} \quad \textit{and} \quad \lim_{n\to\infty} b_n = b \in \mathbb{R},$$

then

1. $\lim_{n\to\infty}(a_n + b_n) = a + b$,
2. $\lim_{n\to\infty} ca_n = ca$ *for any* $c \in \mathbb{R}$,
3. $\lim_{n\to\infty} a_n b_n = ab$,
4. $\lim_{n\to\infty} a_n/b_n = a/b$, *provided* $b_n \neq 0$ *for all* n *and* $b \neq 0$.

Proof. 1. Given $\varepsilon > 0$ there is n_1 such that

$$|a_n - a| < \frac{\varepsilon}{2} \quad \text{for all } n \geq n_1$$

and there is n_2 such that

$$|b_n - b| < \frac{\varepsilon}{2} \quad \text{for all } n \geq n_2.$$

Hence for all $n \geq n_0 = \max\{n_1, n_2\}$ we have

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Similarly we treat the case of $a_n - b_n$.

2. If $c = 0$, then $ca_n = 0 \to 0 = ca$ so the result is obvious. Thus we may assume that $c \neq 0$. Given $\varepsilon > 0$, there is n_0 such that for $n \geq n_0$ we have $|a_n - a| < \varepsilon/|c|$ so

$$|ca_n - ca| = |c|\,|a_n - a| < \varepsilon \quad \text{for } n \geq n_0.$$

3. Given $\varepsilon > 0$ we need to prove that there is n_0 such that

$$|a_n b_n - ab| < \varepsilon \quad \text{for all } n \geq n_0.$$

First we need to do an analysis to find an appropriate n_0. We have[9]

$$|a_n b_n - ab| = |a_n b_n - a_n b + a_n b - ab| \leq |a_n||b_n - b| + |a_n - a||b| = \heartsuit.$$

The sequence (a_n) is bounded as convergent (Theorem 8.14). Hence there is $M > 0$ such that

$$|a_n| < M \quad \text{for all } n.$$

Taking M larger, if necessary, we can also assume that $|b| < M$. Then we have

$$\heartsuit \leq M|b_n - b| + |a_n - a|M.$$

Since $a_n \to a$ and $b_n \to b$ as $n \to \infty$, there is n_0 such that

$$(8.10) \qquad |a_n - a| < \frac{\varepsilon}{2M} \quad \text{and} \quad |b_n - b| < \frac{\varepsilon}{2M} \quad \text{for all } n \geq n_0.$$

Now we can complete the proof. Fix $M > 0$ such that $|b| < M$ and $|a_n| < M$ for all n. Let $\varepsilon > 0$ be given. Let n_0 be such that (8.10) holds. Then for $n \geq n_0$ we have

$$|a_n b_n - ab| \leq M|b_n - b| + |a_n - a|M < M\frac{\varepsilon}{2M} + \frac{\varepsilon}{2M}M = \varepsilon\,.$$

The proof is complete.

4. It suffices to show that $1/b_n \to 1/b$ as $n \to \infty$. The rest will follow from 2. Since $b_n \to b \neq 0$, there is n_1 such that

$$|b_n - b| < \frac{|b|}{2} \quad \text{for } n \geq n_1$$

and hence

$$|b_n| > \frac{|b|}{2} \quad \text{for } n \geq n_1.$$

For $n \geq n_1$ we have

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b_n - b|}{|b_n b|} \leq \frac{2|b - b_n|}{|b|^2}\,.$$

Let $\varepsilon > 0$. Let $n_0 \geq n_1$ be such that

$$|b - b_n| < \frac{\varepsilon|b|^2}{2} \quad \text{for } n \geq n_0.$$

Then for $n \geq n_0$ we have

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| \leq \frac{2|b - b_n|}{|b|^2} < \frac{2}{|b|^2} \cdot \frac{\varepsilon|b|^2}{2} = \varepsilon.$$

The proof is complete. □

[9] We split $a_n b_n - ab$ into two expressions, and we applied the triangle inequality. This allows us to estimate a_n and b_n separately. It is important to understand this trick. Similar tricks will be used over and over again.

The above result can easily be generalized as follows. If $k \in \mathbb{N}$ and (a_{1n}), (a_{2n}),..., (a_{kn}) are convergent sequences, then

$$\lim_{n\to\infty}(a_{1n}+a_{2n}+\ldots+a_{kn}) = \lim_{n\to\infty}a_{1n}+\lim_{n\to\infty}a_{2n}+\ldots+\lim_{n\to\infty}a_{kn} \tag{8.11}$$

$$\lim_{n\to\infty}(a_{1n}\cdot a_{2n}\cdot\ldots\cdot a_{kn}) = \lim_{n\to\infty}a_{1n}\cdot\lim_{n\to\infty}a_{2n}\cdot\ldots\cdot\lim_{n\to\infty}a_{kn}$$

Example 8.17. Applying the above result we have

$$1=\lim_{n\to\infty}1=\lim_{n\to\infty}\underbrace{\frac{1}{n}+\frac{1}{n}+\ldots+\frac{1}{n}}_{n}=\underbrace{\lim_{n\to\infty}\frac{1}{n}+\lim_{n\to\infty}\frac{1}{n}+\ldots+\lim_{n\to\infty}\frac{1}{n}}_{n}=0+0+\ldots+0=0.$$

Where did we make a mistake? The formula (8.11) is true if k is a fixed integer, while here we have n terms with n approaching to infinity. Be careful!

Theorem 8.18. *If $a_n \le b_n \le c_n$ and*

$$\lim_{n\to\infty}a_n=\lim_{n\to\infty}c_n=g\in\mathbb{R},$$

then

$$\lim_{n\to\infty}b_n=g\,.$$

Proof. Let $\varepsilon > 0$ be arbitrary. Then there are n_1 and n_2 such that

$$|c_n-g|<\varepsilon \quad \text{for } n\ge n_1,$$

$$|a_n-g|<\varepsilon \quad \text{for } n\ge n_2.$$

Hence for $n \ge n_0 = \max\{n_1, n_2\}$ we have

$$g-\varepsilon<a_n\le b_n\le c_n<g+\varepsilon,$$

$$g-\varepsilon<b_n<g+\varepsilon,$$
$$-\varepsilon<b_n-g<\varepsilon,$$
$$|b_n-g|<\varepsilon.$$

The proof is complete. □

Theorem 8.19. *For $a > 0$, $\lim_{n\to\infty}\sqrt[n]{a}=1$.*

Proof. We assume first that $a > 1$. Let $x_n = \sqrt[n]{a}-1$. Applying Bernoulli's inequality (Proposition 4.6) we have

$$a=(1+x_n)^n\ge 1+nx_n$$
$$0<x_n\le\frac{a-1}{n}\,.$$

Since the left hand side converges to zero (as a constant sequence) and the right hand side converges to zero we conclude from the previous theorem that $\sqrt[n]{a}-1=x_n\to 0$, and hence $\sqrt[n]{a}\to 1$.

If $a = 1$, then $\sqrt[n]{a} = 1 \to 1$. If $0 < a < 1$, then $1/a > 1$, so $\sqrt[n]{1/a} \to 1$ and thus

$$\sqrt[n]{a} = \frac{1}{\sqrt[n]{1/a}} \to 1.$$

The proof is complete. □

Exercise 8.20. *Find the limit* $\lim_{n\to\infty} \sqrt[n]{3^n + 5^n}$.

Solution. We have

$$5 = \sqrt[n]{5^n} \le \sqrt[n]{3^n + 5^n} \le \sqrt[n]{2 \cdot 5^n} = \sqrt[n]{2} \cdot 5.$$

Since the left hand side and the right hand side both converge to 5 we conclude that

$$\lim_{n\to\infty} \sqrt[n]{3^n + 5^n} = 5\,.$$

□

Theorem 8.21. $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

Proof. Let $x_n = \sqrt[n]{n} - 1$. Then $x_n \ge 0$. Applying the binomial formula we have

$$n = (x_n + 1)^n \ge \binom{n}{2} x_n^2 \cdot 1^{n-2} = \frac{n(n-1)}{2}\, x_n^2\,,$$

$$0 \le x_n \le \sqrt{\frac{2}{n-1}}\,.$$

Since the left hand side and the right hand side both converge to zero (Problem 1) we get $\sqrt[n]{n} - 1 = x_n \to 0$, $\sqrt[n]{n} \to 1$. □

Proposition 8.22. *If* $a > 1$, *then* $\lim_{n\to\infty} a^{-n} = 0$. *If* $|a| < 1$, *then* $\lim_{n\to\infty} a^n = 0$.

Proof. If $a > 1$, Bernoulli's inequality, Proposition 4.6, yields

$$a^n = (1 + (a-1))^n \ge 1 + n(a-1)$$

and hence

$$0 < a^{-n} \le \frac{1}{1 + n(a-1)} \to 0\,.$$

If $a = 0$, then $a^n = 0 \to 0$ and if $0 < |a| < 1$, then $1/|a| > 1$ and hence $|a^n| = (1/|a|)^{-n} \to 0$, so $a^n \to 0$. □

Definition. We say that a sequence (a_n) *diverges to* $+\infty$ if[10]

$$\forall M > 0\ \exists n_0\ \forall n \ge n_0\ \ a_n > M.$$

[10] In this case the limit exists $\lim_{n\to\infty} a_n = +\infty$ or $(-\infty)$, but we still say that the sequence (a_n) diverges.

Then we write

$$\lim_{n\to\infty} a_n = +\infty \quad \text{or} \quad a_n \to +\infty \text{ as } n \to \infty.$$

We say that a sequence (a_n) *diverges to* $-\infty$ if

$$\forall M > 0\ \exists n_0\ \forall n \geq n_0 \ \ a_n < -M.$$

Then we write

$$\lim_{n\to\infty} a_n = -\infty \quad \text{or} \quad a_n \to -\infty \text{ as } n \to \infty.$$

Theorem 8.23. *Let $a_n > 0$. Then $a_n \to 0$ if and only if $1/a_n \to \infty$.*

Proof. Let $a_n > 0$, $a_n \to 0$. Let $M > 0$ be given. Let n_0 be such that $a_n < 1/M$ for $n \geq n_0$. Then $1/a_n > M$ for $n \geq n_0$ and hence $1/a_n \to +\infty$. Similarly we prove the other implication. □

The above result and Proposition 8.22 give

Corollary 8.24. *For $a > 1$, $\lim_{n\to\infty} a^n = +\infty$.*

Example 8.25. $\lim_{n\to\infty} \sqrt{n} = +\infty$.

Proof. We need to prove that

$$\forall M > 0\ \exists n_0\ \forall n \geq n_0 \quad \sqrt{n} > M.$$

Given $M > 0$, let n_0 be such that $n_0 > M^2$. Then for all $n \geq n_0$ we have $n \geq n_0 > M^2$, $n > M^2$, $\sqrt{n} > M$. □

Example 8.26. $\lim_{n\to\infty}(\sqrt{n+1} - \sqrt{n}) = 0$.

Proof.

$$\sqrt{n+1} - \sqrt{n} = \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}} \to 0.$$

It is easy to show that $\sqrt{n+1} + \sqrt{n} \to \infty$ as $n \to \infty$ so $\frac{1}{\sqrt{n+1}+\sqrt{n}} \to 0$ by Theorem 8.23. □

Definition. The *extended real line* $\bar{\mathbb{R}}$ is defined as

$$\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}.$$

Theorem 8.27. *If $\lim_{n\to\infty} a_n = g \in \bar{\mathbb{R}}$, then*

$$\lim_{n\to\infty} \frac{a_1 + a_2 + \ldots + a_n}{n} = g.$$

Theorem 8.28. *If $\lim_{n\to\infty} a_n = g \in \bar{\mathbb{R}}$ and $a_n > 0$ for all n, then*

$$\lim_{n\to\infty} \sqrt[n]{a_1 a_2 \cdots a_n} = g.$$

We will only prove the first theorem. The proof of the second theorem is similar.[11]

Proof. We will prove the theorem under the additional assumption that $g \in \mathbb{R}$. Since $a_n - g \to 0$, the sequence $a_n - g$ is bounded i.e., there is $M > 0$ such that

$$|a_n - g| \leq M \quad \text{for all } n.$$

Given $\varepsilon > 0$, let n_0 be such that

$$|a_n - g| < \frac{\varepsilon}{2} \quad \text{for } n \geq n_0.$$

We need to estimate the arithmetic mean with the help of the above inequality. Observe that the inequality applies only to $n \geq n_0$. Therefore we need to split the arithmetic mean into two parts, where the first part will contain $a_1, \ldots, a_{n_0-1}$ and the second one $a_{n_0}, \ldots, a_n$. Then we will estimate the two parts separately.

For $n \geq n_0$ we have

$$\begin{aligned} \left| \frac{a_1 + \ldots + a_n}{n} - g \right| &= \left| \frac{(a_1 - g) + \ldots + (a_n - g)}{n} \right| \\ &\leq \frac{|a_1 - g| + \ldots + |a_{n_0-1} - g|}{n} + \frac{|a_{n_0} - g| + \ldots + |a_n - g|}{n} \\ &\leq \frac{(n_0 - 1)M}{n} + \frac{(n - n_0 + 1)\varepsilon/2}{n} \leq \frac{n_0 M}{n} + \frac{\varepsilon}{2}\,. \end{aligned}$$

Let $n_1 > n_0$ be such that $n_0 M/n < \varepsilon/2$ for $n > n_1$. The for $n > n_1$

$$\left| \frac{a_1 + \ldots + a_n}{n} - g \right| < \varepsilon.$$

The proof is complete. □

As an immediate consequence of the above results we obtain.

Example 8.29. $\displaystyle\lim_{n\to\infty} \frac{1 + \sqrt{2} + \sqrt[3]{3} + \ldots + \sqrt[n]{n}}{n} = 1.$

Example 8.30. If $a_1 = 1$, $a_2 = 2/1, \ldots$, $a_n = n/(n-1)$, then $a_n \to 1$ and hence

$$\sqrt[n]{n} = \sqrt[n]{1 \cdot \frac{2}{1} \cdot \frac{3}{2} \cdots \frac{n}{n-1}} \to 1\,.$$

This proof of the fact that $\sqrt[n]{n} \to 1$ is much more complicated than the one we obtained earlier, but it is still a nice illustration of Theorem 8.28.

More generally we have.

[11]It can actually be concluded from the result about arithmetic means with the help of continuity of the logarithm.

Theorem 8.31. *If* $a_n > 0$ *and all* n *and* $\lim_{n\to\infty} a_{n+1}/a_n = a \in \bar{\mathbb{R}}$, *then* $\lim_{n\to\infty} \sqrt[n]{a_n} = a$.

Proof. Since $a_{n+1}/a_n \to a$ we see that also the the following sequence converges to a

$$1, \frac{a_2}{a_1}, \frac{a_3}{a_2}, \ldots, \frac{a_n}{a_{n-1}}, \ldots \longrightarrow a.$$

Therefore

$$\frac{\sqrt[n]{a_n}}{\sqrt[n]{a_1}} = \sqrt[n]{1 \cdot \frac{a_2}{a_1} \cdot \frac{a_3}{a_2} \cdot \frac{a_n}{a_{n-1}}} \to a\,.$$

Since $\sqrt[n]{a_1} \to 1$, we conclude that $\sqrt[n]{a_n} \to a$. □

Example 8.32. $a_n = n \to \infty$ and hence

$$\lim_{n\to\infty} \sqrt[n]{n!} = \lim_{n\to\infty} \sqrt[n]{a_1 a_2 \cdots a_n} = \infty.$$

8.2 Monotone sequences

The following important result stems from the fact that the ordered field of real numbers is complete.

Theorem 8.33. *Every sequence that is increasing and bounded from above is convergent.*

Proof. Let $A = \{a_1, a_2, a_3, \ldots\}$ be the set of all values of the sequence (a_n). Since the set is bounded from above

$$g = \sup A = \sup\{a_1, a_2, a_3, \ldots\}$$

exists and belongs to $\mathbb{R}$. Note that $a_n \leq g$ for all n, because g is an upper bound of A. Since g is the least upper bound, for every $\varepsilon > 0$, $g - \varepsilon$ is not an upper bound, so there is n_0 such that $a_{n_0} > g - \varepsilon$. Hence for all $n \geq n_0$

$$g - \varepsilon < a_{n_0} \leq a_n \leq g$$

and thus

$$|a_n - g| < \varepsilon\,.$$

We proved that for any $\varepsilon > 0$ there is n_0 such that for all $n \geq n_0$, $|a_n - g| < \varepsilon$ i.e., we proved that

$$\lim_{n\to\infty} a_n = g.$$

The proof is complete. □

By a similar argument one can prove.

Theorem 8.34. *Every sequence that is decreasing and bounded from below is convergent.*

8.3 Subsequences

If (a_n) is a sequence and $n_1 < n_2 < n_3 < \dots$ are positive integers, then the sequence

$$(b_k), \quad b_k = a_{n_k}$$

is called a *subsequence* of (a_n). The sequence $(a_{n_k})_{k=1}^{\infty}$ is simply obtained from (a_n) by selecting infinitely many terms with indices in the increasing order.

Suppose that $\lim_{n\to\infty} a_n = g \in \bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Then clearly $\lim_{k\to\infty} a_{n_k} = g$. For example if $n_k = k+1$, then $a_{n_k} = a_{k+1}$ and hence

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = g.$$

If $n_k = 2k$, then $a_{n_k} = a_{2k}$ and hence

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{2n} = g.$$

However, form convergence of a subsequence we cannot conclude convergence of the sequence.

Example 8.35. $a_n = (-1)^n$ is divergent, although both subsequences $(a_{2n})_n$, $a_{2n} = 1$ and $(a_{2n+1})_n$, $a_{2n+1} = -1$ are convergent.

However, we have.

Proposition 8.36. *If* $\lim_{n\to\infty} a_{2n} = \lim_{n\to\infty} a_{2n+1} = g \in \bar{\mathbb{R}}$, *then* $\lim_{n\to\infty} a_n = g$.

Proof. We will prove the result under the additional assumption that $g \in \mathbb{R}$. For any $\varepsilon > 0$ there is n_1 such that for $n \geq n_1$

$$|a_{2n} - g| < \varepsilon \quad \text{and} \quad |a_{2n+1} - g| < \varepsilon.$$

Hence for $n_0 = 2n_1 + 1$ and $n \geq n_0$ we have

$$|a_n - g| < \varepsilon,$$

because every integer $n \geq n_0 = 2n_1 + 1$ is of the form $n = 2k$ or $n = 2k+1$ for some $k \geq n_1$. □

Here is a nice and quite typical application of subsequences.

Example 8.37. *Find the limit of the sequence* (x_n) *which is defined as follows:* $x_1 > 0$ *is any number and*

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right).$$

Proof. Suppose for a moment that we know that x_n converges to a positive limit $g \in (0, \infty)$. Then

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \frac{1}{2}\left(x_n + \frac{1}{x_n}\right),$$

so

$$g = \frac{1}{2}\left(g + \frac{1}{g}\right),$$

$$2g = g + \frac{1}{g}, \quad g = \frac{1}{g}, \quad g^2 = 1$$

and hence $g = 1$, because $g > 0$. Therefore it remains to prove that the sequence (x_n) has a positive and finite limit. To this end it suffices show that starting from $n = 2$ the sequence is decreasing and bounded from below by 1 (see Theorem 8.34).

- Bounded from below by 1 for $n \geq 2$:

$$\begin{aligned} x_n \geq 1 \quad &\equiv \\ \frac{1}{2}\left(x_{n-1} + \frac{1}{x_{n-1}}\right) \geq 1 \quad &\equiv \\ x_{n-1} + \frac{1}{x_{n-1}} \geq 2 \quad &\equiv \quad \text{(it is easy to see that all } x_k > 0 \text{ for all } k) \\ x_{n-1}^2 + 1 \geq 2x_{n-1} \quad &\equiv \\ (x_{n-1} - 1)^2 \geq 0. & \end{aligned}$$

Since the last inequality is obviously true, the first inequality is also true as equivalent.

- Decreasing starting from $n \geq 2$:

$$\begin{aligned} x_{n+1} \leq x_n \quad &\equiv \\ \frac{1}{2}\left(x_n + \frac{1}{x_n}\right) \leq x_n \quad &\equiv \\ x_n + \frac{1}{x_n} \leq 2x_n \quad &\equiv \\ \frac{1}{x_n} \leq x_n \quad &\equiv \\ x_n^2 \geq 1 \quad &\equiv \\ x_n \geq 1. & \end{aligned}$$

Since the last inequality is true for $n \geq 2$ (as proved earlier), the first inequality is true for $n \geq 2$ as equivalent. □

Proposition 8.38. *A sequence (a_n) does not converge to $g \in \mathbb{R}$ if and only if there is $\varepsilon > 0$ and a subsequence $(a_{n_k})_k$ such that $|a_{n_k} - g| \geq \varepsilon$.*

Proof. A sequence (a_n) does not converge to $g \in \mathbb{R}$ if and only if the following negation is true:

$$\neg\left(\forall \varepsilon > 0\ \exists N\ \forall n > N \quad |a_n - g| < \varepsilon\right),$$

that is

$$\exists\varepsilon > 0\ \forall N\ \exists n > N \quad |a_n - g| \geq \varepsilon. \tag{8.12}$$

Fix such $\varepsilon > 0$ and take $N = 1$. It follows from (8.12) that there is $n_1 > N = 1$ such that $|a_{n_1} - g| \geq \varepsilon$. Take now $N = n_1$ (ε remains the same through the rest of the proof). Then it follows from (8.12) that there is $n_n > N = n_1$ such that $|a_{n_2} - g| \geq \varepsilon$. Take now $N = n_2$ and again we can find $n_3 > n_2$ such that $|a_{n_3} - g| \geq \varepsilon$. Continuing this procedure ad infinitum we obtain a subsequence $(a_{n_k})_k$, $n_1 < n_2 < \ldots$ such that $|a_{n_k} - g| \geq \varepsilon$ for all $k \in \mathbb{N}$. □

As an application we will prove the following useful result.[12]

Theorem 8.39. $\lim_{n\to\infty} a_n = g \in \bar{\mathbb{R}}$ *if and only if every subsequence of* (a_n) *has a subsequence with the limit equal* g*.*

Proof. We will prove the result under the assumption that $g \in \mathbb{R}$, leaving the case of $g = \pm\infty$ to the reader.

If (a_n) converges to g, then every subsequence of a subsequence converges to g. That is nearly obvious.

To prove the other implication suppose to the contrary that every subsequence has a subsequence convergent to $g \in \mathbb{R}$ and (a_n) does not converge to g. According to Proposition 8.38, there is $\varepsilon > 0$ and a subsequence $(a_{n_k})_k$ such that $|a_{n_k} - g| \geq \varepsilon$ for all $k \in \mathbb{N}$. Hence $(a_{n_k})_k$ cannot have a subsequence convergent to g which contradicts our assumption. The proof is complete. □

8.4 Series

An important class of examples of sequences is provided by series.

Definition. If $s_n = a_1 + a_2 + \ldots + a_n$ and

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} (a_1 + a_2 + \ldots + a_n) = g,$$

then we write

$$a_1 + a_2 + a_3 + \ldots = \sum_{n=1}^{\infty} a_n = g\,.$$

We call the sequence (s_n) a *series* and the elements s_n *partial sums* of the series. The limit g is called the sum of the series. We call the series *convergent* if (s_n) is convergent. If $g = +\infty$ or $g = -\infty$, then we say that the series is *divergent* to $+\infty$ or $-\infty$.

The following result provides a necessary, but not sufficient condition for the convergence of the series.

[12] It is nearly obvious that $a_n \to g$ if and only if every subsequence of (a_n) has limit equal g. However, the theorem is deeper than that since we do not require subsequences to have limits equal g, but instead we have a weaker condition that each subsequence has a subsequence with the limit equal g.

Theorem 8.40. *If the series $\sum_{n=1}^{\infty} a_n$ is convergent, then $a_n \to 0$.*

Proof. Suppose the series is convergent to g, i.e., $s_n \to g$. Then also $s_{n-1} \to g$ (why?) and hence $a_n = s_n - s_{n-1} \to g - g = 0$. □

Theorem 8.41. *If $|q| < 1$, then*

$$\sum_{n=0}^{\infty} q^n = 1 + q + q^2 + \ldots = \frac{1}{1-q}\,.$$

If $|q| \geq 1$, then the series $\sum_{n=1}^{\infty} q^n$ is divergent.

Proof. If $|g| < 1$, then (see Problem 3 in Chapter 4)

$$1 + q + q^2 + \ldots + q^n = \frac{1 - q^{n+1}}{1-q} \to \frac{1}{1-q}$$

by Proposition 8.22. If $|q| \geq 1$, then q^n does not converge to zero, so the series is divergent (see the previous theorem). □

Theorem 8.42.

$$\sum_{n=1}^{\infty} \frac{1}{n} = +\infty.$$

Proof. Since the sequence of partial sums is increasing it suffices to show that it is not bounded. This is however, impossible, because according to Example 4.7

$$s_{3n+1} - s_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{3n+1} > 1$$

and hence

$$s_4 - s_1 > 1, \quad s_{13} - s_1 = (s_{13} - s_4) + (s_4 - s_1) > 2, \quad s_{40} - s_1 = (s_{40} - s_{13}) + (s_{13} - s_1) > 3 \quad \text{etc.}$$

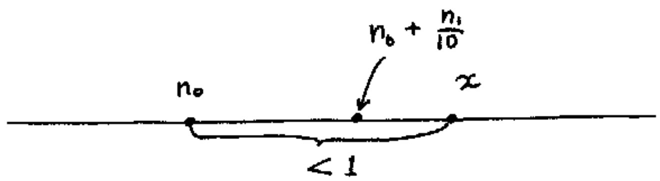
□

Series will be carefully investigated in the Section 10.3.

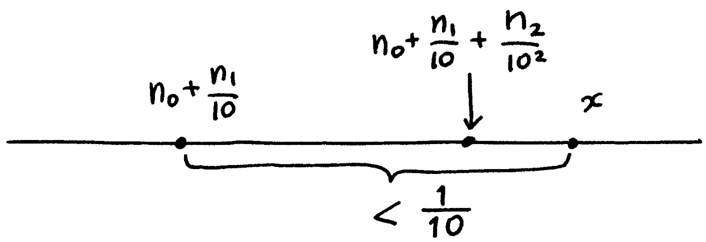
8.5 Decimal expansion

In order to understand the decimal expansion from the perspective of limits of sequences we need to know that bounded and monotone sequences are convergent.

The notion of the limit of a sequence can be used to explain the meaning of the infinite decimal expansion of a real number. Let x_0 be an arbitrary real number. Let n_0 be the largest integer less than or equal to x.



Let n_1 be the largest integer such that $n_0 + n_1/10 \leq x$. Since $n_0 + 10/10 = n_0 + 1 > x$ we conclude that $n_1 \in \{0, 1, 2, \ldots, 9\}$, i.e., n_1 is a digit.



Let n_2 be the largest integer such that $n_0 + n_1/10 + n_2/10^2 \leq x$. Since

$$n_0 + \frac{n_1}{10} + \frac{10}{10^2} = n_0 + \frac{n_1}{10} + \frac{1}{10} > x$$

we conclude that $n_2 \in \{0, 1, 2, \ldots, 9\}$ etc.

The sequence

$$n_0,\ n_0 + \frac{n_1}{10},\ n_0 + \frac{n_1}{10} + \frac{n_2}{10^2}, \ldots$$

is increasing and it is easy to see that it converges to x. Indeed, if $\varepsilon > 0$, then $10^{-N} < \varepsilon$ for some N and hence for $k \geq N$

$$x - \varepsilon < x - \frac{1}{10^N} \leq x - \frac{1}{10^k} < \underbrace{n_0 + \frac{n_1}{10} + \ldots + \frac{n_k}{10^k}}_{a_k} \leq x, \qquad |a_k - x| < \varepsilon\,,$$

which means

$$\lim_{k\to\infty} a_k = x.$$

The fact that

$$x = \lim_{k\to\infty} \left(n_0 + \frac{n_1}{10} + \ldots + \frac{n_k}{10^k} \right) = \sum_{k=0}^{\infty} \frac{n_k}{10^k}$$

where $n_0 \in \mathbb{Z}$ and $n_i \in \{0, 1, 2, \ldots, 9\}$ is simply denoted by

(8.13) $$x = n_0.n_1 n_2 n_3 \ldots$$

and the right hand side of (8.13) is called the *decimal expansion* of x.

Thus we proved that every real number has a decimal expansion. Note also that the decimal expansion of a real number is not unique. Indeed,

$$a_n = 0 + \frac{9}{10} + \frac{9}{10^2} + \ldots + \frac{9}{10^k} = 1 - \frac{1}{10^k} \to 1 \quad \text{as } n \to \infty$$

and hence

$$1 = 0.99999999\ldots \tag{8.14}$$

We already discussed the issue of uniqueness of the decimal expansion when we discussed cardinality of sets, but then we relied of an intuitive understanding of the real number and its decimal expansion so the equality (8.14) could not be properly justified until now.

Observe that in the above argument we did not use completeness of the field of real numbers. For example we could apply the above proof to the field of rational numbers and prove that every rational number has a decimal expansion. However, we will use completeness of $\mathbb{R}$ now. Namely, if we select an arbitrary sequence

$$n_0 \in \mathbb{Z}, \qquad n_i \in \{0, 1, 2, 3, \ldots, 9\},\ i = 1, 2, 3, \ldots$$

then the sequence

$$a_k = n_0 + \frac{n_1}{10} + \frac{n_2}{10^2} + \ldots + \frac{n_k}{10^k}, \quad k = 0, 1, 2, \ldots$$

is increasing and bounded from above by n_0+1. Therefore it is convergent, see Theorem 8.33. The limit is, of course, a real number whose decimal expansion is $n_0.n_1n_2n_3\ldots$ That means, using completeness of $\mathbb{R}$ we proved that every infinite decimal expansion defines a real number.

We used here completeness of $\mathbb{R}$, because it was used in the proof of Theorem 8.33. Indeed, the sequence of rational numbers (approximation of $\sqrt{2}$):

$$1,\ 1.4,\ 1.41,\ 1.414,\ 1.4142,\ 1.41421,\ 1,414213,\ 1.4142135, \ldots$$

is increasing and bounded, but it has no limit in $\mathbb{Q}$.

Remark. We constructed the field of real numbers from the field of rational numbers using cuts. One can also provide a different construction by identification of real numbers with decimal expansions. This is possible, but not as simple as it seems. In the construction we assume familiarity with rational numbers, so numbers 1, 1.4, 1.41, 1.414, 1.4142,... are well defined however, 1.4142135... (decimal expansion of $\sqrt{2}$) can only be understood as a formal expression and we have to define addition[13] and multiplication of such expressions and prove all properties. We would also have to address the problem of non uniqueness of the decimal expansion. After all this construction of real numbers would not be easier than the one that uses cuts.

[13]That is not easy! When we add 23.2987732 to 4.1307921 we start from the last digit, but for infinite expansions we do not have the last digit to start with.

8.6 Examples

Example 8.43. *Find the limit* $\lim\limits_{n\to\infty} n\left(\sqrt[3]{n^3+n}-n\right)$.

Solution. Identity
$$\frac{a^3-b^3}{a^2+ab+b^2}=a-b$$
gives
$$\sqrt[3]{n^3+n}-n=\frac{(n^3+n)-n^3}{\left(\sqrt[3]{n^3+n}\right)^2+n\sqrt[3]{n^3+n}+n^2},$$

$$\begin{aligned} n\left(\sqrt[3]{n^3+n}-n\right) &= \frac{n^2}{\left(\sqrt[3]{n^3+n}\right)^2+n\sqrt[3]{n^3+n}+n^2} \\ &= \frac{1}{\left(\sqrt[3]{1+\frac{1}{n^2}}\right)^2+\sqrt[3]{1+\frac{1}{n^2}}+1}\to\frac{1}{3}. \end{aligned}$$

□

Example 8.44. *Let* $a\in\{1,2,3,\ldots,9\}$ *be a digit. By aaaaa we denote a number whose decimal representation has* 5 *digits a. Find the limit*

$$\lim_{n\to\infty}\frac{a+aa+aaa+\ldots+\overbrace{aaa\ldots a}^{n\ digits}}{10^n}.$$

Solution. We have

$$\begin{aligned} a+aa+\ldots+aaa\ldots a &= a(1+11+111+\ldots+\overbrace{111\ldots1}^{n}) \\ &= a\left(\frac{10-1}{9}+\frac{10^2-1}{9}+\frac{10^3-1}{9}+\ldots+\frac{10^n-1}{9}\right) \\ &= a\Big(\frac{\overbrace{111\ldots10}^{n+1}-n}{9}\Big)=\frac{a}{9}\big(10\cdot\overbrace{111\ldots1}^{n}-n\big) \\ &= \frac{a}{9}\left(10\cdot\frac{10^n-1}{9}-n\right)=\frac{a}{81}\big(10(10^n-1)-9n\big) \end{aligned}$$

and hence
$$\frac{a+aa+\ldots+aaa\ldots a}{10^n}=\frac{a}{81}\left(10\left(1-\frac{1}{10^n}\right)-\frac{9n}{10^n}\right)\to\frac{10a}{81}.$$

□

Example 8.45. *Prove that* $\lim\limits_{n\to\infty}\left(\sqrt[n]{n}-1\right)^n=0.$

Proof. Since $\sqrt[n]{n} \to 1$, $\sqrt[n]{n} - 1 < 1/2$ for $n \geq n_0$ and hence

$$0 < (\sqrt[n]{n} - 1)^n < 1/2^n \to 0.$$

□

Example 8.46. *Find a series whose nth partial sum equals $s_n = (n+1)/n$.*

Solution. $a_1 = s_1 = 2$,

$$a_1 + \ldots + a_{n-1} = s_{n-1} = \frac{n}{n-1}, \qquad a_1 + \ldots + a_n = s_n = \frac{n+1}{n}\,.$$

Hence

$$a_n = s_n - s_{n-1} = \frac{n+1}{n} - \frac{n}{n-1} = \frac{(n+1)(n-1) - n^2}{n(n-1)} = -\frac{1}{n(n-1)}\,,$$

so the series is

$$2 - \frac{1}{2 \cdot 1} - \frac{1}{3 \cdot 2} - \frac{1}{4 \cdot 3} - \frac{1}{5 \cdot 4} - \ldots$$

□

Example 8.47. *Find the sum of the series* $\displaystyle\sum_{n=1}^{\infty} \frac{2n+1}{n^2(n+1)^2}$.

Solution. Since

$$\frac{2n+1}{n^2(n+1)^2} = \frac{1}{n^2} - \frac{1}{(n+1)^2}\,,$$

we have

$$s_n = \left(\frac{1}{1^2} - \frac{1}{2^2}\right) + \left(\frac{1}{2^2} - \frac{1}{3^2}\right) + \ldots + \left(\frac{1}{n^2} - \frac{1}{(n+1)^2}\right) = 1 - \frac{1}{(n+1)^2} \to 1\,.$$

Accordingly

$$\sum_{n=1}^{\infty} \frac{2n+1}{n^2(n+1)^2} = 1\,.$$

□

Example 8.48. *Find the sum* $\displaystyle\sum_{n=1}^{\infty} \sin\frac{n!\pi}{720}$.

Solution. If $n \geq 6$, then

$$n! = \underbrace{1 \cdot 2 \cdots 6}_{720} \cdot 7 \cdots n\,,$$

so $a_n = 0$ for $n \geq 6$. Hence

$$\begin{aligned}
\sum_{n=1}^{\infty} \sin\frac{n!\pi}{720} &= \sin\frac{\pi}{720} + \sin\frac{2\pi}{720} + \sin\frac{6\pi}{720} + \sin\frac{24\pi}{720} + \sin\frac{120\pi}{720} \\
&= \sin\frac{\pi}{720} + \sin\frac{\pi}{360} + \sin\frac{\pi}{120} + \sin\frac{\pi}{30} + \sin\frac{\pi}{6}\,.
\end{aligned}$$

□

Example 8.49. *Prove that* $\displaystyle\sum_{n=1}^{\infty} \frac{n}{3\cdot 5\cdots(2n+1)} = \frac{1}{2}$.

Proof. For $n > 1$ we have

$$\begin{aligned} a_n &= \frac{n}{3\cdot 5\cdots(2n+1)} = \frac{1}{2}\frac{(2n+1)-1}{3\cdot 5\cdots(2n+1)} \\ &= \frac{1}{2}\left(\frac{1}{3\cdot 5\cdots(2n-1)} - \frac{1}{3\cdot 5\cdots(2n+1)}\right) \end{aligned}$$

and hence

$$\begin{aligned} &a_1 + a_2 + \ldots + a_n = \\ &= \underbrace{\frac{1}{3}}_{a_1} + \frac{1}{2}\left(\frac{1}{3} - \frac{1}{3\cdot 5} + \frac{1}{3\cdot 5} - \frac{1}{3\cdot 5\cdot 7} + \ldots + \frac{1}{3\cdot 5\cdots(2n-1)} - \frac{1}{3\cdot 5\cdots(2n+1)}\right) \\ &= \frac{1}{3} + \frac{1}{2}\left(\frac{1}{3} - \frac{1}{3\cdot 5\cdots(2n+1)}\right) \to \frac{1}{3} + \frac{1}{6} = \frac{1}{2}\,. \end{aligned}$$

□

8.7 Problems

Problem 1. Show directly from the definition of the limit that

$$\lim_{n\to\infty} \sqrt{\frac{2}{n-1}} = 0.$$

Chapter 9

Exponential and logarithmic functions

9.1 The definition

Definition. For $n, m \in \mathbb{N}$ and $x > 0$ we define

$$x^{\frac{m}{n}} = \left(x^{\frac{1}{n}}\right)^m \quad \text{and} \quad x^{-\frac{m}{n}} = \frac{1}{x^{\frac{m}{n}}}.$$

Therefore we defined

$$x^q \text{ for any } x > 0 \text{ and } q \in \mathbb{Q}.$$

Observe that a rational number q has many representations as a fraction

$$q = \frac{m}{n} = \frac{km}{kn}.$$

Thus in order to prove that x^q is well defined one needs to show that the definition is independent of the choice of a particular representation i.e.

$$\left(x^{\frac{1}{n}}\right)^m = \left(x^{\frac{1}{kn}}\right)^{km}.$$

This is however, easy and the reader can (and should) check it.

Proposition 9.1. *For $x, y > 0$ and $q \in \mathbb{Q}$ we have $(xy)^q = x^q y^q$.*

Proof. We have

$$\left(x^{\frac{1}{n}} y^{\frac{1}{n}}\right)^n = \underbrace{\left(x^{\frac{1}{n}} y^{\frac{1}{n}}\right) \cdots \left(x^{\frac{1}{n}} y^{\frac{1}{n}}\right)}_{n \text{ times}} = \underbrace{x^{\frac{1}{n}} \cdots x^{\frac{1}{n}}}_{n} \underbrace{y^{\frac{1}{n}} \cdots y^{\frac{1}{n}}}_{n} = \left(x^{\frac{1}{n}}\right)^n \left(y^{\frac{1}{n}}\right)^n = xy.$$

Since $z = x^{1/n} y^{1/n}$ satisfies $z^n = xy$ we have $z = (xy)^{1/n}$ i.e.,

$$x^{\frac{1}{n}} y^{\frac{1}{n}} = (xy)^{\frac{1}{n}}.$$

Thus
$$(xy)^{\frac{m}{n}} = \left((xy)^{\frac{1}{n}}\right)^m = \left(x^{\frac{1}{n}}y^{\frac{1}{n}}\right)^m = \left(x^{\frac{1}{n}}\right)^m \left(y^{\frac{1}{n}}\right)^m = x^{\frac{m}{n}}y^{\frac{m}{n}} .$$
Similarly one can prove that
$$(xy)^{-\frac{m}{n}} = x^{-\frac{m}{n}}y^{-\frac{m}{n}} .$$
The proof is complete. □

Using similar arguments one can prove other well known algebraic properties of the function $f(x) = x^q$ like for example

Proposition 9.2. *If $x > 0$ and $p, q \in \mathbb{Q}$, then $x^p x^q = x^{p+q}$, $(x^p)^q = x^{pq}$.*

We leave the proof as an exercise.

Proposition 9.3. *If $a > 1$, the the function*
$$\mathbb{Q} \ni q \mapsto a^q$$
is strictly increasing.

Proof. We will prove that if $0 < q_1 < q_2$, then $a^{q_1} < a^{q_2}$. The case in which we allow the exponents to be less than or equal zero is left to the reader. Let $0 < q_1 = m_1/n_1 < q_2 = m_2/n_2$. Then $m_1n_2 < m_2n_1$. Since
$$a^{\frac{1}{n_1n_2}} > 1$$
we have
$$a^{q_1} = a^{\frac{m_1}{n_1}} = \left(a^{\frac{1}{n_1n_2}}\right)^{m_1n_2} < \left(a^{\frac{1}{n_1n_2}}\right)^{m_2n_1} = a^{\frac{m_2}{n_2}} = a^{q_2} .$$
The proof is complete. □

If $a = 1$, the function $q \mapsto a^q$ is constant (equal 1) and if $0 < a < 1$, the function $q \mapsto a^q$ is decreasing.

So far we defined x^q for $x > 0$ and rational exponents q only.

Definition. If $a > 1$ and $x \in \mathbb{R}$ we define
$$a^x = \sup\left\{a^q : q \leq x \ \wedge \ q \in \mathbb{Q}\right\} .$$
Clearly the set is nonempty. Since the function $q \mapsto a^q$ is increasing, the set is bounded from above by a^{q_0}, where q_0 is any rational number larger than x. Therefore the supremum exists and is finite.

If $0 < a < 1$, then we define
$$a^x = \frac{1}{\left(\frac{1}{a}\right)^x} .$$
Note that $(1/a)^x$ is well defined, because $1/a > 1$.

One can easily prove that the function $f(x) = a^x$, $x \in \mathbb{R}$ has the same algebraic properties as in the case of rational exponents. For example $(ab)^x = a^x b^x$, $a^{x+y} = a^x a^y$, $(a^x)^y = a^{xy}$

Moreover the function $f(x) = a^x$ is increasing when $a > 1$ and decreasing when $0 < a < 1$. We leave details as an exercise.

For $a > 1$, the function $f(x) = a^x$, $f : \mathbb{R} \to (0, \infty)$ is strictly increasing. Therefore it is one-to-one and hence it has an inverse. The domain of the inverse function is the image of f. The problem is that we did not prove so far that the image of f is $(0, \infty)$. We will do it now.

Proposition 9.4. *Let $f(x) = a^x$, $1 \neq a > 0$, $f : \mathbb{R} \to (0, \infty)$. Then $f(\mathbb{R}) = (0, \infty)$.*

Proof. Assume first that $a > 1$. Let $y \in (0, \infty)$. We have to prove that there is $x \in \mathbb{R}$ such that $a^x = y$. Let

$$A = \{z \in \mathbb{R} : a^z \leq y\}.$$

$A \neq \emptyset$, because $a^{-n} \to 0$ as $n \to \infty$ (Corollary 8.22), so $a^{-n_0} < y$ for some n_0, and hence $-n_0 \in A$. A is bounded from above, because $a^n \to \infty$ as $n \to \infty$ (Corollary 8.24) and hence $a^{n_1} > y$ for some n_1. Thus $a^z > y$ for all $z \geq n_1$, so n_1 is an upper bound of A. Accordingly

$$x := \sup A \in \mathbb{R}.$$

Clearly $x - 1/n$ is not an upper bound of A, so there is $z \in A$ such that $z > x - 1/n$ and hence

$$a^{x-1/n} < a^z \leq y.$$

Since $x + 1/n \notin A$ we also have $a^{x+1/n} > y$. Thus

$$a^x (a^{1/n})^{-1} = a^{x-1/n} < y < a^{x+1/n} = a^x a^{1/n}.$$

Both sequences that bound y converge to a^x as $n \to \infty$. Hence the constant sequence y also converges to a^x i.e., $a^x = y$.

If $0 < a < 1$, then $1/a > 1$, so for every $y \in (0, \infty)$ there is $z \in \mathbb{R}$ such that $(1/a)^z = y$ and hence for $x = -z$ we have $a^x = (1/a)^z = y$. The proof is complete. □

The function $f(x) = a^x$, $f : \mathbb{R} \to (0, \infty)$ is strictly increasing for $a > 1$ and strictly decreasing for $0 < a < 1$. In either case the inverse function $f^{-1}(x)$ exists. Since $f(\mathbb{R}) = (0, \infty)$, the domain of f^{-1} is $(0, \infty)$. The function f^{-1} is strictly increasing when $a > 1$ (why?) and strictly decreasing when $0 < a < 1$ (why?). We denote it by

$$\log_a(x) := f^{-1}(x).$$

We have

$$\log_a(a^x) = f^{-1}(f(x)) = x \quad \text{for all } x \in \mathbb{R}. \tag{9.1}$$

$$a^{(\log_a(x))} = f(f^{-1}(x)) = x \quad \text{for all } x \in (0, \infty), \tag{9.2}$$

Algebraic properties of the logarithm follow easily from the algebraic properties of the exponential function. For example (9.2) yields

$$a^{\log_a xy} = xy = a^{\log_a x} a^{\log_a y} = a^{\log_a x + \log_a y}.$$

Since the function $z \mapsto a^z$ is one-to-one the exponents must be equal i.e.,

$$\log_a xy = \log_a x + \log_a y.$$

All other algebraic properties of the logarithm can be proved in a similar way. We assume that the reader is familiar with those properties and in what follow we will take them for granted.

Theorem 9.5. *If $a > 0$, then $\lim_{n\to\infty} n^{-a} = \lim_{n\to\infty} 1/n^a = 0$.*

Proof. Given $\varepsilon > 0$, for $n > \varepsilon^{-1/a}$, we have

$$\left|\frac{1}{n^a} - 0\right| = \frac{1}{n^a} < \varepsilon$$

and the claim follows. □

Example 9.6. Consider the sequence $a_n = n^{1,000,000}/1.000001^n$. The numerator growths very fast and denominator seems to growth slowly. However, it turns out that eventually the growth of the denominator will be much faster than that of numerator and

$$\lim_{n\to\infty} \frac{n^{1000000}}{1.000001^n} = 0\,.$$

This follows from the next result. Remember this example when you have money in the bank even with a very small interest rate, but you are willing to wait for a sufficiently long time!

Theorem 9.7. *If $p > 0$ and $\alpha \in \mathbb{R}$, then*

$$\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0\,.$$

Proof. For $k > \alpha$, $k > 0$ and $n > 2k$ we have

$$(1+p)^n > \binom{n}{k} p^k \cdot 1^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!}\, p^k > \left(\frac{n}{2}\right)^k \frac{p^k}{k!}\,,$$

because each factor in the product $n(n-1)\cdots(n-k+1)$ is bigger than $n/2$. Hence

$$0 < \frac{n^\alpha}{(1+p)^n} < n^\alpha \left(\frac{2}{n}\right)^k \frac{k!}{p^k} = \frac{2^k k!}{p^k}\, n^{\alpha-k}\,.$$

Since $\alpha - k < 0$, the right hand side converges to zero as $n \to \infty$ (Theorem 9.5), and the claim follows. □

9.2 Number e and natural logarithm

Theorem 9.8. *The sequence $a_n = (1 + 1/n)^n$ is strictly increasing, the sequence $b_n = (1 + 1/n)^{n+1}$ is strictly decreasing and both sequences converge to the same limit. We denote this limit by*

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^{n+1}.$$

Proof. First we will prove that a_n is strictly increasing. To this end it suffices to show that $a_{n+1}/a_n > 1$. We have

$$\begin{aligned}
\frac{a_{n+1}}{a_n} &= \frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{\left(1+\frac{1}{n}\right)^n} = \frac{\left(\frac{n+2}{n+1}\right)^{n+1}}{\left(\frac{n+1}{n}\right)^n} \\
&= \left(\frac{(n+2)n}{(n+1)^2}\right)^n \frac{n+2}{n+1} = \left(\frac{n^2+2n}{n^2+2n+1}\right)^n \frac{n+2}{n+1} \\
&= \left(1-\frac{1}{n^2+2n+1}\right)^n \frac{n+2}{n+1} \\
&\geq \left(1-\frac{n}{n^2+2n+1}\right) \frac{n+2}{n+1} \qquad \text{(Bernoulli)} \\
&= \frac{n^3+3n^2+3n+2}{n^3+3n^2+3n+1} > 1.
\end{aligned}$$

Similarly we can show that the sequence b_n is decreasing (we leave it as an exercise). Since $a_n \leq b_n$ for every n we have

$$2 = a_1 < a_2 < a_3 < \ldots < a_n < \ldots < b_n < b_{n-1} < \ldots < b_1 = 4.$$

Hence a_n is increasing and bounded from above, so convergent. Also b_n is decreasing and bounded from above, so convergent. Clearly $\lim_{n\to\infty} a_n \in (2,4)$, so $\lim_{n\to\infty} a_n \neq 0$ and hence

$$\frac{\lim_{n\to\infty} b_n}{\lim_{n\to\infty} a_n} = \lim_{n\to\infty} \frac{b_n}{a_n} = \lim_{n\to\infty} 1 + \frac{1}{n} = 1, \qquad \lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n.$$

□

One more proof that a_n is increasing. A clever application of the Arithmetic-Geometric mean inequality gives

$$\begin{aligned}
\left(\left(1+\frac{1}{n}\right)^n \cdot 1\right)^{1/(n+1)} &= \sqrt[n+1]{\left(1+\frac{1}{n}\right)\cdots\left(1+\frac{1}{n}\right)\cdot 1} \\
&\leq \frac{\left(1+\frac{1}{n}\right)+\ldots+\left(1+\frac{1}{n}\right)+1}{n+1} = 1+\frac{1}{n+1}\,.
\end{aligned}$$

Hence

$$\left(1+\frac{1}{n}\right)^n \cdot 1 \leq \left(1+\frac{1}{n+1}\right)^{n+1}, \qquad \left(1+\frac{1}{n}\right)^n \leq \left(1+\frac{1}{n+1}\right)^{n+1}.$$

□

Remark. Since $(1+1/n)^n$ is increasing and $(1+1/n)^{n+1}$ is decreasing and e is their common limit, we have that

$$\left(1+\frac{1}{n}\right)^n < e < \left(1+\frac{1}{n}\right)^{n+1}$$

for every n. Taking n large we obtain lower and upper estimate for e. One can prove that

$$e = 2.718281828\ldots$$

Theorem 9.9. $e = \sum_{n=0}^{\infty} \frac{1}{n!}$.

Proof. Let
$$x_n = \left(1 + \frac{1}{n}\right)^n, \quad y_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}.$$
Thus (y_n) is the sequence of the partial sums of the above series. The binomial formula yields
$$\begin{aligned} x_n &= 1^n + \binom{n}{1}1^{n-1}\frac{1}{n} + \binom{n}{2}1^{n-2}\frac{1}{n^2} + \ldots + \binom{n}{n-1}1^1\frac{1}{n^{n-1}} + \binom{n}{n}1^0\frac{1}{n^n} \\ &= 1 + 1 + \frac{n(n-1)}{2!}\frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!}\frac{1}{n^3} + \ldots \\ &+ \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}\frac{1}{n^k} + \ldots + \frac{n(n-1)(n-2)\cdots 1}{n!}\frac{1}{n^n}. \end{aligned}$$
Fix k. For $n \geq k$ we have
$$\begin{aligned} x_n &\geq 1 + 1 + \frac{n(n-1)}{2!}\frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!}\frac{1}{n^3} + \ldots + \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}\frac{1}{n^k} \\ &= 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \ldots + \frac{1}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right). \end{aligned}$$
We fix k and let $n \to \infty$. Then $x_n \to e$ and the right hand side converges to
$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{k!} = y_k.$$
Therefore
$$e \geq y_k \quad \text{for every } k.$$
The sequence (y_k) is increasing and bounded from above by e, so it is convergent and
$$\lim_{k \to \infty} y_k \leq e.$$
On the other hand
$$\begin{aligned} x_n &= 1 + \frac{1}{1!} + \frac{1}{2!}\frac{n(n-1)}{n^2} + \frac{1}{3!}\frac{n(n-1)(n-2)}{n^3} + \ldots + \frac{1}{n!}\frac{n(n-1)\cdots 1}{n^n} \\ &< 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!} = y_n \end{aligned}$$
and hence
$$e = \lim_{n \to \infty} x_n \leq \lim_{n \to \infty} y_n.$$
This and the previous estimate for the limit of y_n yield $\lim_{n \to \infty} y_n = e$ i.e.,
$$\sum_{n=0}^{\infty} \frac{1}{n!} = \lim_{n \to \infty} y_n = e.$$
□

Theorem 9.10. *e is irrational.*

Proof. Let

$$x_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}\,.$$

Then

$$\begin{aligned}
e - x_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \ldots \\
&= \frac{1}{(n+1)!}\left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \frac{1}{(n+2)(n+3)(n+4)} + \ldots\right) \\
&< \frac{1}{(n+1)!}\underbrace{\left(1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \frac{1}{(n+1)^3} + \ldots\right)}_{\text{geometric series}} \\
&= \frac{1}{(n+1)!}\,\frac{1}{1 - \frac{1}{n+1}} = \frac{1}{n!\,n}\,.
\end{aligned}$$

Hence

$$0 < e - x_n < \frac{1}{n!\,n}\,.$$

Suppose that e is a rational number i.e., $e = p/q$ for some $p, q \in \mathbb{N}$. Then

$$0 < e - x_q < \frac{1}{q!\,q}\,,$$

$$0 < \underbrace{eq!}_{\text{integer}} - \underbrace{x_q q!}_{\text{integer}} < \frac{1}{q}\,.$$

Since there are no integers between 0 and $1/q$ we arrived a contradiction. This proves that e cannot be a rational number. □

Definition. The *natural logarithm* is defined by

$$\ln x = \log x = \log_e x.$$

Observe that differently than in high school, $\log x$ is with base e instead of 10.

It is not clear at this point why the base e is more important than any other base. It will be transparent later when we will study derivatives, but even now the following result shows a nice and important inequality that is true for the natural logarithm.

Lemma 9.11. $\dfrac{1}{n+1} < \ln\left(1 + \dfrac{1}{n}\right) < \dfrac{1}{n}$ *for $n = 1, 2, 3, \ldots$.*

Proof. The inequality

$$\left(1 + \frac{1}{n}\right)^n < e < \left(1 + \frac{1}{n}\right)^{n+1}$$

implies

$$n \ln\left(1 + \frac{1}{n}\right) < 1 < (n+1)\ln\left(1 + \frac{1}{n}\right)\,.$$

The left inequality gives

$$\ln\left(1+\frac{1}{n}\right) < \frac{1}{n}$$

and the right inequality gives

$$\frac{1}{n+1} < \ln\left(1+\frac{1}{n}\right) .$$

□

Theorem 9.12. *The sequence*

$$a_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} - \ln n$$

is convergent to a finite limit

$$\gamma := \lim_{n\to\infty}\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} - \ln n\right) .$$

Remark. The limit $\gamma = 0.5772156649\ldots$ is called the *Euler constant.* It is not know if γ is rational or not. Good exercise for you!

Proof. We will prove that the sequence is decreasing. To this end it suffices to show that $a_{n+1} - a_n < 0$. We have

$$\begin{aligned}
a_{n+1} - a_n &= \left(1 + \frac{1}{2} + \ldots + \frac{1}{n} + \frac{1}{n+1}\right) - \ln(n+1) - \left(1 + \frac{1}{2} + \ldots + \frac{1}{n}\right) + \ln n \\
&= \frac{1}{n+1} - \ln(n+1) + \ln n \\
&= \frac{1}{n+1} - \ln\left(\frac{n+1}{n}\right) \\
&= \frac{1}{n+1} - \ln\left(1+\frac{1}{n}\right) < 0 ,
\end{aligned}$$

where the last inequality follows from the lemma. Therefore the sequence (a_n) is decreasing. Applying the lemma one more time we have

$$1 > \ln(1+1),\ \frac{1}{2} > \ln\left(1+\frac{1}{2}\right), \ldots, \frac{1}{n} > \ln\left(1+\frac{1}{n}\right) ,$$

and hence

$$\begin{aligned}
a_n &= 1 + \frac{1}{2} + \ldots + \frac{1}{n} - \ln n \\
&> \ln(1+1) + \ln\left(1+\frac{1}{2}\right) + \ldots + \ln\left(1+\frac{1}{n}\right) - \ln n \\
&= \ln 2 + \ln\frac{3}{2} + \ldots + \ln\frac{n+1}{n} - \ln n \\
&= \ln\left(2\cdot\frac{3}{2}\cdot\frac{4}{3}\cdot\frac{5}{4}\cdots\frac{n+1}{n}\right) - \ln n \\
&= \ln(n+1) - \ln n > 0.
\end{aligned}$$

Thus the sequence is decreasing and bounded from below by 0. Hence it is convergent. □

As a corollary we obtain another proof that

$$\sum_{n=1}^{\infty} \frac{1}{n} = +\infty.$$

Indeed, since the sequence of partial sums

$$s_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$$

is increasing it suffices to show that it is not convergent. Suppose it is convergent. Since the sequence in Theorem 9.12 is also convergent, the difference of two sequences i.e., the sequence $s_n - a_n = \ln n$ is also convergent, but it is not, since $\lim_{n\to\infty} \ln n = +\infty$.

9.3 Examples

Example 9.13. *Prove that the sequence* $\sqrt[n]{n}$ *is decreasing starting from* $n = 3$.

Proof. We have

$$n^{1/n} > (n+1)^{1/(n+1)} \;\equiv\; n^{n+1} > (n+1)^n \;\equiv\; n > \frac{(n+1)^n}{n^n} = \left(1 + \frac{1}{n}\right)^n .$$

The last inequality is true for $n \geq 3$, because $n \geq 3 > e > (1 + 1/n)^n$ and hence the first inequality is true for $n \geq 3$ as equivalent. □

Example 9.14. *Find the following limits*

$$(a) \quad \lim_{n\to\infty} \left(\frac{n!}{n^n e^{-n}}\right)^{1/n}, \qquad (b) \quad \lim_{n\to\infty} \left(\frac{(n!)^3}{n^{3n} e^{-n}}\right)^{1/n}.$$

Solution.

(a). Let $a_n = \dfrac{n!}{n^n e^{-n}}$. Then

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{(n+1)!}{(n+1)^{n+1} e^{-(n+1)}} \frac{n^n e^{-n}}{n!} = \frac{n!(n+1)}{(n+1)(n+1)^n e^{-n} e^{-1}} \frac{n^n e^{-n}}{n!} \\ &= \frac{n^n e}{(n+1)^n} = \frac{e}{\left(1 + \frac{1}{n}\right)^n} \to 1 \end{aligned}$$

and hence Theorem 8.31 gives

$$\sqrt[n]{a_n} = \left(\frac{n!}{n^n e^{-n}}\right)^{1/n} \to 1.$$

(b). Let $a_n = \dfrac{(n!)^3}{n^{3n}e^{-n}}$. Then

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{\big((n+1)!\big)^3}{\big((n+1)^{n+1}\big)^3 e^{-(n+1)}}\,\frac{n^{3n}e^{-n}}{(n!)^3} = \frac{(n!)^3(n+1)^3}{(n+1)^3(n+1)^{3n}e^{-n}e^{-1}}\,\frac{n^{3n}e^{-n}}{(n!)^3}\\ &= \frac{n^{3n}e}{(n+1)^{3n}} = \frac{e}{\big((1+\frac{1}{n})^n\big)^3} \to \frac{e}{e^3} = e^{-2}\end{aligned}$$

and hence Theorem 8.31 gives

$$\sqrt[n]{a_n} = \left(\frac{(n!)^3}{n^{3n}e^{-n}}\right)^{1/n} \to e^{-2}\,.$$

□

Chapter 10

Subsequences, series and the Cauchy condition

10.1 The Cauchy condition and the Bolzano-Weierstrass theorem

The following result plays a fundamental role in analysis and toplogy.

Theorem 10.1 (Bolzano-Weierstrass). *Every bounded sequence of real numbers has a convergent subsequence.*

Proof. Let (x_n) be a bounded sequence i.e.,

$$a_1 \leq x_n \leq b_1$$

for some $a_1, b_1 \in \mathbb{R}$ and all $n \in \mathbb{N}$. Divide the interval $[a_1, b_1]$ into two subintervals of equal length:

$$\left[a_1, \frac{a_1 + b_1}{2}\right], \quad \left[\frac{a_1 + b_1}{2}, b_1\right].$$

Obviously infinitely many elements belong to one of the two intervals. Perhaps to both of them, but for sure to at least one of them. Choose such an interval and denote its endpoints by a_2 and b_2. Hence

$$[a_2, b_2] = \left[a_1, \frac{a_1 + b_1}{2}\right] \quad \text{or} \quad [a_2, b_2] = \left[\frac{a_1 + b_1}{2}, b_1\right].$$

Now divide the interval $[a_2, b_2]$ into two subintervals of equal length

$$\left[a_2, \frac{a_2 + b_2}{2}\right], \quad \left[\frac{a_2 + b_2}{2}, b_2\right].$$

Infinitely many elements of the sequence belong to at least one of the two intervals. Denote such an interval by $[a_3, b_3]$ etc. We repeat this procedure infinitely many times.

In this construction we obtain an increasing sequence (a_n) and a decreasing sequence (b_n):

$$a_1 \leq a_2 \leq a_3 \leq \ldots \leq b_3 \leq b_2 \leq b_1.$$

We see that the sequence (a_n) is increasing and bounded from above, so it is convergent. Also the sequence (b_n) is decreasing and bounded form below, so it is also convergent. Since

$$b_n - a_n = \frac{b_1 - a_1}{2^{n-1}} \to 0$$

we conclude that both sequences converge to the same limit

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = g.$$

Now we will show how to select a subsequence (x_{n_k}) such that

$$\lim_{k\to\infty} x_{n_k} = g.$$

Choose $n_1 = 1$. We have $a_1 \leq x_{n_1} \leq b_1$. Infinitely many x_n's satisfy the inequality $a_2 \leq x_n \leq b_2$, so there is $n_2 > n_1$ such that $a_2 \leq x_{n_2} \leq b_2$. Since infinitely many x_n's satisfy the inequality $a_3 \leq x_n \leq b_3$, there is $n_3 > n_2$ such that $a_3 \leq x_{n_3} \leq b_3$ etc.

We constructed an increasing sequence of integers $n_1 < n_2 < n_3 < \ldots$ such that the subsequence (x_{n_k}) satisfies

$$a_k \leq x_{n_k} \leq b_k.$$

Since both the left and the right hand side converge to g we conclude that $x_{n_k} \to g$. □

Definition. We say that (x_n) is a *Cauchy sequence* if

$$\forall \varepsilon > 0\ \exists n_0\ \forall n, m \geq n_0 \quad |x_n - x_m| < \varepsilon.$$

The above condition is called the *Cauchy condition.*

The next result states that a sequence is convergent if and only if it is a Cauchy sequence. This is a very important result and it is easy to understand why. If we want to prove that a sequence (x_n) is convergent using the definition we have to find a number g (limit) such that $|x_n - g| < \varepsilon$ for all sufficiently large n, so the first thing we have to do is to identify the limit. However, in many situations one can prove that a sequence is convergent, even if it is not possible to find the limit. This can be achieved by checking that the sequence satisfies the Cauchy condition – this condition does not involve the limit g and only requires an estimate of $|x_n - x_m|$.

Theorem 10.2. *A sequence (x_n) is convergent if and only if (x_n) is a Cauchy sequence.*

Proof. Since this is "if and only if" condition, we have to prove two implications.

($\Rightarrow$) Suppose that $\lim_{n\to\infty} x_n = g \in \mathbb{R}$. Then, for every $\varepsilon > 0$ we can find n_0 such that

$$|x_n - g| < \frac{\varepsilon}{2} \quad \text{for } n \geq n_0$$

and thus

$$|x_n - x_m| = |x_n - g + g - x_m| \leq |x_n - g| + |x_m - g| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $n, m \geq n_0$.

($\Leftarrow$) Suppose (x_n) is a Cauchy sequence. First we will prove that (x_n) is bounded. The definition of the Cauchy sequence with $\varepsilon = 1$ gives that there is n_0 such that

$$|x_n - x_m| < 1 \quad \text{for all } n, m \geq n_0.$$

In particular we can take $m = n_0$ and hence

$$|x_n - x_{n_0}| < 1 \quad \text{for all } x \geq n_0,$$

so

$$|x_n| < 1 + |x_{n_0}| \quad \text{for all } x \geq n_0.$$

Therefore

$$|x_n| < 1 + |x_{n_0}| + |x_1| + |x_2| + \ldots + |x_{n_0 - 1}| \quad \text{for all } n.$$

Since the sequence (x_n) is bounded, the Bolzano-Weierstrass theorem implies that there is a subsequence (x_{n_k}) convergent to a finite limit

$$x_{n_k} \to g \in \mathbb{R}.$$

We will prove that $\lim_{n\to\infty} x_n = g$. To this end let $\varepsilon > 0$ be given. We need to find N_0 such that

$$|x_n - g| < \varepsilon \quad \text{for all } n \geq N_0.$$

Because (x_n) is a Cauchy sequence, there is N_1 such that

$$|x_n - x_m| < \frac{\varepsilon}{2} \quad \text{for all } n, m \geq N_1.$$

Convergence of x_{n_k} implies that there is $n_{k_0} \geq N_1$ such that

$$|x_{n_{k_0}} - g| < \frac{\varepsilon}{2}\,.$$

Set $N_0 = n_{k_0}$. Then for $n \geq N_0 = n_{k_0} \geq N_1$ we have

$$|x_n - g| \leq |x_n - x_{n_{k_0}}| + |x_{n_{k_0}} - g| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

□

10.2 The upper and the lower limits

For any sequence (a_n) of real numbers we define

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup\{a_n, a_{n+1}, a_{n+2}\ldots\}, \quad \liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf\{a_n, a_{n+1}, a_{n+2}\ldots\}.$$

First observe that if (a_n) is unbounded from above then there is a subsequence divergent to ∞ and $\limsup_{n\to\infty} a_n = \infty$. Similarly, if (a_n) is unbounded from below, then there is a subsequence divergent to $-\infty$ and $\liminf_{n\to\infty} a_n = -\infty$.

Moreover, it is not difficult to prove that

$$\lim_{n\to\infty} a_n = \infty \quad \text{if and only if} \quad \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = +\infty \tag{10.1}$$

and

$$\lim_{n\to\infty} a_n = -\infty \quad \text{if and only if} \quad \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = -\infty. \tag{10.2}$$

We leave details to the reader.

For the rest of this section we will focus on a more interesting case of bounded sequences.

Let (a_n) be a bounded sequence. Since it has a convergent subsequence, the set of possible limits of convergent subsequences of (a_n) is bounded and non-empty.

Theorem 10.3. *Let (a_n) be a bounded sequence. Then there is a convergent subsequence of (a_n) such that its limit is greater than or equal to the limit of any other convergent subsequence of (a_n). In other words $\sup G \in G$.*

Proof. Let $x = \sup G$. It suffices to show that there is a subsequence convergent to x.

Let $k \in \mathbb{N}$. Since $x - 1/k < \sup G$, there is a subsequence with the limit larger than $x - 1/k$ and, in particular, infinitely many a_n's are larger than $x - 1/k$. Therefore, there is n_1 such that $a_{n_1} > x - 1/1$. Since infinitely many a_n's are larger than $x - 1/2$, there is $n_2 > n_1$ such that $a_{n_2} > x - 1/2$. Since there are infinitely many a_n's larger than $x - 1/3$, there is $n_3 > n_2$ such that $a_{n_3} > x - 1/3$ etc. We obtain a subsequence $(a_{n_k})_k$ such that $a_{n_k} > x - 1/k$. The sequence $(a_{n_k})_k$ is bounded and so it has a convergent subsequence $(a_{n_{k_\ell}})_\ell$. Clearly $\lim_{\ell\infty} a_{n_{k_\ell}} \leq x$. Note that $k_\ell \geq \ell$ so

$$a_{n_{k_\ell}} \geq x - \frac{1}{k_\ell} \geq x - \frac{1}{\ell} \geq x - \frac{1}{i} \quad \text{provided } \ell \geq i$$

and hence

$$x \geq \lim_{\ell\to\infty} a_{n_{k_\ell}} \geq x - \frac{1}{i} \quad \text{for all } i \in \mathbb{N}$$

proving that $\lim_{\ell\to\infty} a_{n_{k_\ell}} = x$. □

Note that the sequence

$$b_n = \sup\{a_n, a_{n+1}, a_{n+2}, \ldots\}$$

is bounded and decreasing[1] so it is convergent.

Theorem 10.4. *Let (a_n) be a bounded sequence. Then*

$$\lim_{n\to\infty} \sup\{a_n, a_{n+1}, a_{n+2}, \ldots\} = \sup G.$$

[1] Increasing n we, decrease the set over which we take supremum and hence supremum is decreasing.

Proof. Let (a_{n_k}) be a subsequence convergent to $\sup G$, the existence of which is guaranteed by Theorem 10.3. Fix $i \in \mathbb{N}$. For each $n \in \mathbb{N}$, there is $n_k > n$ such that $a_{n_k} > \sup G - 1/i$ and hence for each n

$$b_n = \sup\{a_n, a_n + 1, a_{n+2} \ldots\} \geq a_{n_k} > \sup F - \frac{1}{i}$$

proving that $\lim_{n\to\infty} b_n \geq \sup G - 1/i$ so $\lim_{n\to\infty} b_n \geq \sup G$. It remains to show that $\lim_{n\to\infty} b_n \leq \sup G$. Suppose to the contrary that $\lim_{n\to\infty} b_n > \sup G$. Then $\lim_{n\to\infty} b_n > \sup G + \varepsilon$ for some $\varepsilon > 0$. Since the sequence b_n is decreasing, for each n, $\sup\{a_n, a_{n+1}, a_{n+2}, \ldots\} > \sup G + \varepsilon$ and so for each n we can find $k \geq n$ such that $a_k > \sup G + \varepsilon$. It easily follows that there is a subsequence (a_{n_k}), $a_{n_k} > \sup G + \varepsilon$ and so (a_{n_k}) has a subsequence convergent to a limit that is greater than or equal to $\sup G + \varepsilon$ which contradicts the definition of the set G. □

Similarly we prove

Theorem 10.5. *Let (a_n) be a bounded sequence. Then there is a convergent subsequence of (a_n) such that its limit is less than or equal to the limit of any other convergent subsequence of (a_n). In other words, $\inf G \in G$.*

Theorem 10.6. *Let (a_n) be a bounded sequence. Then*

$$\lim_{n\to\infty} \inf\{a_n, a_{n+1}, a_{n+2} \ldots\} = \inf G.$$

It follows that is $\limsup_{n\to\infty} a_n$ and $\liminf_{n\to\infty} a_n$ are the lowest and the largest possible limits among all subsequences of (a_n).

Finally we have

Theorem 10.7. *For any sequence (a_n) of real numbers we have*

$$\lim_{n\to\infty} a_n = g \quad \textit{if and only if} \quad \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = g.$$

Proof. The case $g = \pm\infty$ was discussed in (10.1) and (10.2) and left to the reader as an exercise. Thus we assume that $g \in \mathbb{R}$.

If $a_n \to g$, then all the subsequences converge to g so $G = \{g\}$, $\inf G = \sup G = g$ and clearly

$$\liminf_{n\to\infty} a_n = \inf G = \limsup_{n\to\infty} a_n = \sup G = g. \tag{10.3}$$

That proves implication from left to right. To prove the implication from right to left assume that (10.3) is satisfied. Since $g \in \mathbb{R}$, the sequence (a_n) is bounded as otherwise it would have a subsequence divergent to $-\infty$ when not bounded from below (so $\liminf_{n\to\infty} a_n = -\infty$) or a subsequence divergent to $+\infty$ when not bounded from above (so $\limsup_{n\to\infty} a_n = +\infty$).

Since (a_n) is bounded, every subsequence of (a_n) has a convergent subsequence. Since $\sup G = \inf G = \{g\}$ it follows that $G = \{g\}$ so all such convergent subsequences must converge to g and hence $\lim_{n\to\infty} a_n = g$ by Theorem 8.39.

□

10.3 Series

As we know, the series $a_1 + a_1 + a_3 + \ldots$ can be identified with the sequence of partial sums $s_n = a_1 + a_2 + \ldots + a_n$. Therefore the characterization that a sequence is convergent if and only if it is a Cauchy sequence has a direct reformulation for series. The reader will easily check that the corresponding result reads as follows.

Theorem 10.8. *The series $a_1 + a_2 + a_3 + \ldots$ is convergent (to a finite sum) if and only if*

$$\forall \varepsilon > 0\ \exists n_0\ \forall n \geq n_0\ \forall m \geq 0 \quad |a_n + a_{n+1} + \ldots + a_{n+m}| < \varepsilon.$$

The condition formulated in the above theorem is called the *Cauchy condition* for a series.

Theorem 10.9. *If $a_n \geq 0$, $n = 1, 2, 3, \ldots$, then $\sum_{n=1}^{\infty} a_n$ converges if and only if the sequence of partial sums is bounded.*

Proof. It is easy to see that the condition $a_n \geq 0$ is equivalent to the condition that the sequence of partial sums is increasing, and an increasing sequence is convergent if and only if it is bounded. □

The next test for convergence of a series makes use of the Cauchy condition.

Theorem 10.10 (Comparison test)**.**

(a) *Suppose that there is N such that $|a_n| \leq b_n$ for $n \geq N$ and $\sum_{n=1}^{\infty} b_n$ converges. Then the series $\sum_{n=1}^{\infty} a_n$ converges, too.*

(b) *Suppose that there is N such that $a_n \geq b_n \geq 0$ for all $n \geq N$ and $\sum_{n=1}^{\infty} b_n$ diverges (to ∞, of course), then the series $\sum_{n=1}^{\infty} a_n$ diverges, too (to ∞, of course).*

Proof. Part (b) is quite obvious, so we will only prove part (a). Since the series $\sum_{n=1}^{\infty} b_n$ converges, it satisfies the Cauchy condition

$$\forall \varepsilon > 0\ \exists n_0\ \forall n \geq n_0\ \forall m \geq 0 \quad b_n + b_{n+1} + \ldots b_{n+m} < \varepsilon.$$

We did not put absolute value because $b_k \geq 0$. Hence for $n \geq \max\{n_0, N\}$ and $m \geq 0$

$$|a_n + a_{n+1} + \ldots + a_{n+m}| \leq b_n + b_{n+1} + \ldots + b_{n+m} < \varepsilon,$$

so the series $\sum_{n=1}^{\infty} a_n$ satisfies the Cauchy condition and thus it is convergent. □

Definition. We say that a series $\sum_{n=1}^{\infty} a_n$ is *absolutely convergent* if $\sum_{n=1}^{\infty} |a_n|$ is convergent.

Theorem 10.11. *If $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, then $\sum_{n=1}^{\infty} a_n$ is convergent.*

Proof. Observe that $|a_n| \leq |a_n| := b_n$. Since the series $\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} |a_n|$ is convergent, the Comparison Test implies convergence of $\sum_{n=1}^{\infty} a_n$. □

Theorem 10.12 (Cauchy condensation test)**.** *Suppose $a_1 \geq a_2 \geq a_3 \geq \ldots \geq 0$. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{n=0}^{\infty} 2^n a_{2^n}$ converges.*

Proof. Denote

$$s_n = a_1 + a_2 + \ldots + a_n, \qquad t_k = a_1 + 2a_2 + 4a_4 + \ldots + 2^k a_{2^k}\,.$$

For $n < 2^k$ we have

$$\begin{aligned} s_n &= a_1 + \ldots + a_n \le a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \ldots + (a_{2^k} + \ldots + a_{2^{k+1}-1}) \\ &\le a_1 + 2a_2 + 4a_4 + \ldots + 2^k a_{2^k} = t_k. \end{aligned}$$

Now if $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges, the partial sums t_k are bounded, so are the partial sums s_n, and since $a_n \ge 0$ that implies convergence of the series $\sum_{n=1}^{\infty} a_n$.

For $n > 2^k$ we have

$$\begin{aligned} s_n &= a_1 + a_2 + \ldots + a_n \\ &\ge a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \ldots + (a_{2^{k-1}+1} + \ldots + a_{2^k}) \\ &\ge \frac{1}{2}a_1 + a_2 + 2a_4 + 4a_8 + \ldots + 2^{k-1} a_{2^k} = \frac{1}{2} t_k\,. \end{aligned}$$

Now if $\sum_{n=1}^{\infty} a_n$ converges, the partial sums s_n are bounded, so are the partial sums t_k and hence the series $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges. □

Theorem 10.13. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^p}$ *converges for* $p > 1$ *and diverges for* $0 < p \le 1$.

Proof. For $0 < p \le 1$, $1/n^p \ge 1/n$, so the divergence of the series $\sum_{n=1}^{\infty} 1/n^p$ follows from $\sum_{n=1}^{\infty} 1/n = +\infty$. To prove convergence for $p > 1$ we apply the Cauchy condensation test. Let $a_n = 1/n^p$. Then

$$2^n a_{2^n} = 2^n \frac{1}{(2^n)^p} = \frac{1}{(2^{p-1})^n}\,,$$

$$\sum_{n=0}^{\infty} 2^n a_{2^n} = \sum_{n=0}^{\infty} \left(\frac{1}{2^{p-1}}\right)^n$$

converges as a geometric series with $|1/2^{p-1}| < 1$. Hence $\sum_{n=1}^{\infty} a_n$ converges, too. □

Remark. One can prove that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}\,.$$

Theorem 10.14. $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$ *converges for* $p > 1$ *and diverges for* $0 < p \le 1$.

Proof. Let $a_n = 1/(n(\log n)^p)$. Then

$$2^n a_{2^n} = 2^n \frac{1}{2^n (\log 2^n)^p} = \left(\frac{1}{\log 2}\right)^p \frac{1}{n^p}\,.$$

Hence the previous result yields that

$$\sum_{n=1}^{\infty} 2^n a_{2^n} = \left(\frac{1}{\log 2}\right)^p \sum_{n=2}^{\infty} \frac{1}{n^p}$$

converges if and only if $p > 1$ and the theorem follows from the Cauchy condensation test. □

Theorem 10.15 (d'Alembert test)**.**

(a) *If* $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$*, then the series* $\sum_{n=1}^{\infty} a_n$ *converges absolutely.*

(b) *If* $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$*, then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Theorem 10.16 (Cauchy test)**.**

(a) *If* $\lim_{n\to\infty} \sqrt[n]{|a_n|} < 1$*, then the series* $\sum_{n=1}^{\infty} a_n$ *converges absolutely.*

(b) *If* $\lim_{n\to\infty} \sqrt[n]{|a_n|} > 1$*, then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Remark. If

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1 \quad \text{or} \quad \lim_{n\to\infty} \sqrt[n]{|a_n|} = 1,$$

then we cannot conclude convergence or divergence of the series. For example, if $a_n = 1/n$, then the above limits are equal 1 and the series $\sum_{n=1}^{\infty} a_n$ diverges. If $a_n = 1/n^2$, then still the above limits are equal 1, but this time the series $\sum_{n=1}^{\infty} a_n$ converges.

We will prove the d'Alembert test only; the proof for the Cauchy test is similar and left as an exercise.

Proof. If $\lim_{n\to\infty} |a_{n+1}/a_n| < 1$, then there is $0 < q < 1$ and n_0 such that

$$\left|\frac{a_{n+1}}{a_n}\right| < q \quad \text{for } n \geq n_0.$$

For $n \geq n_0$ we have

$$|a_{n+1}| < q|a_n| < q^2|a_{n-1}| < \ldots < q^{n+1-n_0}|a_{n_0}|\,,$$
$$|a_{n+1}| < \left(q^{-n_0}|a_{n_0}|\right) q^{n+1}$$

Replacing $n+1$ by n in this formula we have

$$|a_n| < \left(q^{-n_0}|a_{n_0}|\right) q^n \quad \text{for } n > n_0.$$

Since the series

$$\sum_{n=1}^{\infty} \left(q^{-n_0}|a_{n_0}|\right) q^n = \left(q^{-n_0}|a_{n_0}|\right) \sum_{n=1}^{\infty} q^n$$

converges, the series $\sum_{n=1}^{\infty} a_n$ converges absolutely by the comparison test.

If $\lim_{n\to\infty} |a_{n+1}/a_n| > 1$, there are n_0 and $q > 1$ such that $|a_{n+1}/a_n| > q$ for $n \geq n_0$ and it easily follows that a_n does not converge to zero. Hence the series $\sum_{n=1}^{\infty} a_n$ diverges (see Theorem 8.40). □

Example 10.17. For every $x \in \mathbb{R}$ the series $\sum_{n=0}^{\infty} x^n/n!$ converges absolutely. It is obvious if $x = 0$, so we can assume that $x \neq 0$. If $a_n = x^n/n!$, then $|a_{n+1}/a_n| = |x|/(n+1) \to 0$, so the absolute convergence follows from the d'Alambert test.

Example 10.18. *Investigate convergence of the series* $\displaystyle\sum_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^{(n+1)n}$.

Solution. Let $a_n = \left(\frac{n}{n+1}\right)^{(n+1)n}$. Then

$$\sqrt[n]{a_n} = \left(\frac{n}{n+1}\right)^{n+1} = \frac{1}{\left(\frac{n+1}{n}\right)^{n+1}} = \frac{1}{\left(1+\frac{1}{n}\right)^n} \frac{1}{1+\frac{1}{n}} \to \frac{1}{e} < 1$$

and hence the series converges. □

Theorem 10.19. *Assume that $a_n > 0$, $b_n > 0$ and*

$$\frac{a_{n+1}}{a_n} \leq \frac{b_{n+1}}{b_n} \quad \textit{for all } n \geq n_0.$$

If the series $\sum_{n=1}^{\infty} b_n$ converges, then the series $\sum_{n=1}^{\infty} a_n$ converges, too.

Remark 10.20. If $\lim_{n\to\infty} b_{n+1}/b_n < 1$, then convergence of the series $\sum_{n=1}^{\infty} a_n$ follows from the d'Alambert test. However, if

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{b_{n+1}}{b_n} = 1$$

and we know that the series $\sum_{n=1}^{\infty} b_n$ converges, we still can conclude convergence of $\sum_{n=1}^{\infty} a_n$ even if d'Alambert's test does not apply. We will see examples after we prove the theorem.

Proof. Let $c_n = a_n/b_n$. Then

$$c_{n+1} = \frac{a_{n+1}}{b_{n+1}} \leq \frac{a_n}{b_n} = c_n \quad \text{for } n \geq n_0,$$

so c_n is decreasing starting from $n = n_0$. Hence c_n is bounded, say $c_n \leq M$ for all n. Therefore

$$a_n = c_n b_n \leq M b_n$$

and convergence of the series $\sum_{n=1}^{\infty} M b_n = M \sum_{n=1}^{\infty} b_n$ implies convergence of $\sum_{n=1}^{\infty} a_n$. □

Now we will show two applications of the above result.

Example 10.21. *Investigate convergence of the series* $\displaystyle\sum_{n=1}^{\infty} \frac{n^{n-2}}{e^n n!}$.

Solution. Let $a_n = \dfrac{n^{n-2}}{e^n n!}$. Then

$$\begin{aligned}
\frac{a_{n+1}}{a_n} &= \frac{(n+1)^{n-1}}{e^{n+1}(n+1)!}\,\frac{e^n n!}{n^{n-2}} = \frac{(n+1)^{n-2}(n+1)e^n n!}{e\,e^n(n+1)n!n^{n-2}} \\
&= \frac{\left(1+\frac{1}{n}\right)^{n-2}}{e} = \underbrace{\frac{\left(1+\frac{1}{n}\right)^{n}}{e}}_{<1}\left(1+\frac{1}{n}\right)^{-2} \\
&< \left(\frac{n}{n+1}\right)^2 = \frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}}\,.
\end{aligned}$$

Hence

$$\frac{a_{n+1}}{a_n} < \frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}}\,.$$

Since the series $\sum_{n=1}^{\infty} 1/n^2$ converges, the series $\sum_{n=1}^{\infty} a_n$ converges, too. □

Example 10.22. *Investigate convergence of the series* $\displaystyle\sum_{n=1}^{\infty} \frac{n^n}{e^n n!}$.

Solution. Let $a_n = \dfrac{n^{n}}{e^n n!}$. Then

$$\begin{aligned}
\frac{a_{n+1}}{a_n} &= \frac{(n+1)^{n+1}}{e^{n+1}(n+1)!}\,\frac{e^n n!}{n^{n}} = \frac{(n+1)^{n}(n+1)e^n n!}{e^n e n!(n+1)n^{n}} \\
&= \frac{\left(1+\frac{1}{n}\right)^{n}}{e} > \frac{\left(1+\frac{1}{n}\right)^{n}}{\left(1+\frac{1}{n}\right)^{n+1}} = \frac{1}{1+\frac{1}{n}} \\
&= \frac{n}{n+1} = \frac{\frac{1}{n+1}}{\frac{1}{n}}\,.
\end{aligned}$$

Hence

$$\frac{\frac{1}{n+1}}{\frac{1}{n}} \le \frac{a_{n+1}}{a_n}\,.$$

Suppose that the series $\sum_{n=1}^{\infty} a_n$ converges. The the theorem would give convergence of the series $\sum_{n=1}^{\infty} 1/n$ which is a contradiction. Therefore $\sum_{n=1}^{\infty} a_n$ diverges. □

10.4 Alternating series

Theorem 10.23. *Suppose that*

(a) $|c_1| \ge |c_2| \ge |c_3| \ge \ldots$, $\lim_{n\to\infty} c_n = 0$.

(b) $c_1 \ge 0$, $c_2 \le 0$, $c_3 \ge 0$, $c_4 \le 0, \ldots$

Then the series $\sum_{n=1}^{\infty} c_n$ *is convergent.*

Remark 10.24. A series of the form described in the above theorem is called an *alternating series*, because c_n changes sign.

Proof. Observe that

$$c_{2n+1} + c_{2n+2} \geq 0, \quad \text{because } c_{2n+1} \geq 0 \text{ and } |c_{2n+2}| \leq |c_{2n+1}|$$

and

$$c_{2n} + c_{2n+1} \leq 0, \quad \text{because } c_{2n} \leq 0 \text{ and } |c_{2n+1}| \leq |c_{2n}|.$$

Hence the sequence of partial sums $s_n = a_1 + a_2 + \ldots + a_n$ satisfies

$$s_{2n} \leq s_{2n} + (c_{2n+1} + c_{2n+2}) = s_{2n+2}, \qquad s_{2n-1} \geq s_{2n-1} + (c_{2n} + c_{2n+1}) = s_{2n+2}.$$

Thus $s_2 \leq s_4 \leq s_6 \leq \ldots$ and $s_1 \geq s_3 \geq s_5 \geq \ldots$ Since

$$s_{2n} \leq s_{2n} + \underbrace{c_{2n+1}}_{\geq 0} = s_{2n+1}$$

we have

$$s_2 \leq s_4 \leq s_6 \leq s_8 \leq \ldots \leq s_9 \leq s_7 \leq s_5 \leq s_3 \leq s_1.$$

Thus the sequence $(s_{2n})_{n=1}^{\infty}$ is convergent as increasing and bounded and $(s_{2n+1})_{n=0}^{\infty}$ is convergent as decreasing and bounded. Since

$$|s_{2n+1} - s_{2n}| = |c_{2n+1}| \to 0$$

we conclude that

$$\lim_{n\to\infty} s_{2n} = \lim_{n\to\infty} s_{2n+1}$$

and hence the sequence (s_n) is also convergent (Proposition 8.36). This however, means convergence of the series $\sum_{n=1}^{\infty} c_n$. □

As an immediate application of the result wee see that the series $\sum_{n=1}^{\infty}(-1)^{n+1}/n$ is convergent, but not absolutely convergent. We will find the sum of this series.

Theorem 10.25. $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \ln 2$.

Proof. We need the following lemma.

Lemma 10.26.

$$\lim_{n\to\infty}\left(\frac{1}{n} + \frac{1}{n+1} + \ldots + \frac{1}{2n}\right) = \ln 2\,.$$

Proof. Recall the inequality

$$\frac{1}{n+1} < \ln\left(1 + \frac{1}{n}\right) < \frac{1}{n}\,. \tag{10.4}$$

The left inequality yields

$$\begin{aligned}\frac{1}{n}+\frac{1}{n+1}+\ldots+\frac{1}{2n} &< \ln\left(1+\frac{1}{n-1}\right)+\ln\left(1+\frac{1}{n}\right)+\ldots+\ln\left(1+\frac{1}{2n-1}\right)\\ &= \ln\left(\frac{n}{n-1}\cdot\frac{n+1}{n}\cdot\frac{n+2}{n+1}\cdots\frac{2n}{2n-1}\right)\\ &= \ln\frac{2n}{n-1}\,.\end{aligned}$$

Similarly the right inequality of (10.4) gives

$$\ln\frac{2n+1}{n}<\frac{1}{n}+\frac{1}{n+1}+\ldots+\frac{1}{2n}\,.$$

Hence

$$\ln\frac{2n+1}{n}<\frac{1}{n}+\frac{1}{n+1}+\ldots+\frac{1}{2n}<\ln\frac{2n}{n-1} \tag{10.5}$$

It suffices to prove that[2]

$$\lim_{n\to\infty}\ln\frac{2n+1}{n}=\lim_{n\to\infty}\ln\frac{2n}{n-1}=\ln 2.$$

We have

$$\frac{1}{2n+1}<\ln\left(1+\frac{1}{2n}\right)<\frac{1}{2n}$$

and hence

$$\lim_{n\to\infty}\ln\left(1+\frac{1}{2n}\right)=0.$$

Thus

$$\lim_{n\to\infty}\ln\frac{2n+1}{n}=\lim_{n\to\infty}\left(\ln 2+\ln\left(1+\frac{1}{2n}\right)\right)=\ln 2.$$

Similarly we prove that $\lim_{n\to\infty}\ln 2n/(n-1)=\ln 2$. □

Now we can prove the theorem. We know that the series converges (alternating series), $s_n\to g$, so $s_{2n}\to g$. We have

$$\begin{aligned}s_{2n} &= 1-\frac{1}{2}+\frac{1}{3}-\ldots-\frac{1}{2n}\\ &= 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots+\frac{1}{2n}-2\left(\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2n}\right)\\ &= 1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{2n}-\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}\right)\\ &= \frac{1}{n+1}+\frac{1}{n+2}+\ldots+\frac{1}{2n}\\ &= \left(\frac{1}{n}+\frac{1}{n+1}+\ldots+\frac{1}{2n}\right)-\frac{1}{n}\to\ln 2\end{aligned}$$

and hence $g=\ln 2$. □

[2] This follows immediately from the continuity of the function $\ln x$, but we do not want to refer to continuous functions at this point of the game.

10.5 Multiplication of series

Formally we would like to multiply two series as follows

$$(a_1 + a_2 + a_3 + \ldots)(b_1 + b_2 + b_3 + \ldots) = a_1 b_1 + (a_1 b_2 + a_2 b_1) + (a_1 b_3 + a_2 b_2 + a_3 b_1) + \ldots$$

In the first group $a_1 b_1$ we collect all terms with indices that add up to 2. In the second group $a_1 b_2 + a_2 b_1$ we collect terms with indices that add up to 3. Then terms with indices that add up to 4 and so on. Since we deal with infinite sums we have to rigorously investigate when the above formula is correct. We have

Theorem 10.27 (Cauchy multiplication formula)**.** *If the series $\sum_{n=1}^{\infty} a_n$ converges absolutely and the series $\sum_{n=1}^{\infty} b_n$ converges, then*

$$\left(\sum_{n=1}^{\infty} a_n\right)\left(\sum_{n=1}^{\infty} b_n\right) = \sum_{n=1}^{\infty} c_n\,,$$

where

$$c_1 = a_1 b_1$$

$$c_2 = a_1 b_2 + a_2 b_1$$

$$\ldots$$

$$c_n = a_1 b_n + a_2 b_{n-1} + \ldots + a_n b_1$$

$$\ldots$$

Proof. Let

$$s_n = a_1 + a_2 + \ldots + a_n, \quad t_n = b_1 + b_2 + \ldots + b_n, \quad u_n = c_1 + c_2 + \ldots + c_n.$$

It suffices to prove that

$$\lim_{n\to\infty} s_n t_n - u_n = 0.$$

Indeed, this will readily yield

$$\sum_{n=1}^{\infty} c_n = \lim_{n\to\infty} u_n = \lim_{n\to\infty} s_n t_n = \lim_{n\to\infty} s_n \cdot \lim_{n\to\infty} t_n = \left(\sum_{n=1}^{\infty} a_n\right)\left(\sum_{n=1}^{\infty} b_n\right).$$

Observe that

$$\begin{aligned} u_n &= \sum_{i+j\le n+1} a_i b_j = a_1(b_1 + \ldots + b_n) + a_2(b_1 + \ldots + b_{n-1}) + \ldots + a_n b_1 \\ &= a_1 t_n + a_2 t_{n-1} + \ldots + a_n t_1. \end{aligned}$$

Since the series $\sum_{n=1}^{\infty} b_n$ and $\sum_{n=1}^{\infty} |a_n|$ converge, the sequences of partial sums are bounded i.e., there is $M > 0$ such that

$$|t_n| \le M, \quad |a_1| + \ldots + |a_n| \le M \quad \text{for all } n.$$

In particular $|t_n - t_m| \le 2M$ for all n and m. Moreover it follows from the Cauchy condition that given $\varepsilon > 0$ there is N such that for $n, m \ge N$

$$|t_n - t_m| < \frac{\varepsilon}{3M}, \qquad |a_{N+1}| + \ldots + |a_n| < \frac{\varepsilon}{3M}\,.$$

Now for $n > 2N$ we have

$$\begin{aligned}
|s_n t_n - u_n| &= \Big|\underbrace{(a_1 t_n + a_2 t_n + \ldots + a_n t_n)}_{s_n t_n} - \underbrace{(a_1 t_n + a_2 t_{n-1} + \ldots + a_n t_1)}_{u_n}\Big| \\
&\le \big(|a_1|\,|t_n - t_n| + |a_2|\,|t_n - t_{n-1}| + \ldots + |a_N|\,|t_n - t_{n-N+1}|\big) \\
&+ \big(|a_{N+1}|\,|t_n - t_{n-N}| + \ldots + |a_n|\,|t_n - t_1|\big) \\
&\le \big(|a_1| + \ldots + |a_N|\big)\frac{\varepsilon}{3M} + \big(|a_{N+1}| + \ldots + |a_n|\big)2M \\
&\le M \cdot \frac{\varepsilon}{3M} + \frac{\varepsilon}{3M} \cdot 2M = \varepsilon
\end{aligned}$$

and hence $\lim_{n\to\infty} s_n t_n - u_n = 0$. We used here the fact that

$$|t_n - t_n|,\ |t_n - t_{n-1}|, \ldots, |t_n - t_{n-N+1}| < \frac{\varepsilon}{3M}$$

which is true, because

$$n, n-1, \ldots, n-N+1 \ge N.$$

□

Exercise 10.28. *Use the Cauchy multiplication formula to find the sum of the series*

$$\sum_{n=1}^{\infty} n x^{n-1}, \quad |x| < 1.$$

Solution. The series $\sum_{n=0}^{\infty} x^n$ converges absolutely for $|x| < 1$. Hence

$$\begin{aligned}
\left(\frac{1}{1-x}\right)^2 &= \left(\sum_{n=0}^{\infty} x^n\right)\left(\sum_{n=0}^{\infty} x^n\right) = (1 + x + x^2 + \ldots)(1 + x + x^2 + \ldots) \\
&= 1 + (1\cdot x + x \cdot 1) + (1 \cdot x^2 + x \cdot x + x^2 \cdot 1) + (1 \cdot x^3 + x \cdot x^2 + x^2 \cdot x + x^3 \cdot 1) + \ldots \\
&+ (1 \cdot x^n + x \cdot x^{n-1} + x^2 \cdot x^{n-2} + \ldots + x^n \cdot 1) + \ldots \\
&= 1 + 2x + 3x^2 + 4x^3 + \ldots + (n+1)x^n + \ldots \\
&= \sum_{n=0}^{\infty}(n+1)x^n = \sum_{n=1}^{\infty} n x^{n-1}\,.
\end{aligned}$$

Thus

$$\sum_{n=1}^{\infty} n x^{n-1} = \left(\frac{1}{1-x}\right)^2.$$

□

Example 10.29. As an application of the Cauchy multiplication formula we will prove that

$$\left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{y^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}\,.$$

It follows from the d'Alambert test that both series

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{and} \quad \sum_{n=0}^{\infty} \frac{y^n}{n!}$$

converge absolutely. Hence the Cauchy formula gives

$$\begin{aligned}
&\left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{y^n}{n!}\right) \\
= &\sum_{n=0}^{\infty} \left(\frac{x^0}{0!}\cdot\frac{y^n}{n!} + \frac{x^1}{1!}\cdot\frac{y^{n-1}}{(n-1)!} + \frac{x^2}{2!}\cdot\frac{y^{n-2}}{(n-2)!} + \ldots + \frac{x^n}{n!}\cdot\frac{y^0}{0!}\right) \\
= &\sum_{n=0}^{\infty} \frac{1}{n!}\Big(\underbrace{x^0y^n + \frac{n}{1!}\,xy^{n-1} + \frac{n(n-1)}{2!}\,x^2y^{n-2} + \ldots + \frac{n(n-1)\cdots 1}{n!}\,x^ny^0}_{(x+y)^n}\Big) \\
= &\sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}\,.
\end{aligned}$$

□

10.6 Examples

Exercise 10.30. *Investigate convergence of the series* $\sum_{n=1}^{\infty} \frac{1}{n} \ln\left(1+\frac{1}{n}\right)$.

Solution. The inequality

$$\ln\left(1+\frac{1}{n}\right) < \frac{1}{n}$$

yields

$$0 < \frac{1}{n}\ln\left(1+\frac{1}{n}\right) < \frac{1}{n^2}\,.$$

Since the series $\sum_{n=1}^{\infty}$ converges, the comparison test implies convergence of the series

$$\sum_{n=1}^{\infty} \frac{1}{n}\ln\left(1+\frac{1}{n}\right)\,.$$

□

Exercise 10.31. *Prove that the series* $\sum_{n=1}^{\infty} \left(\sqrt[n]{n}-1\right)^n$ *converges.*

Proof. Let $a_n = \left(\sqrt[n]{n} - 1\right)^n$. Since $\sqrt[n]{|a_n|} = \sqrt[n]{n} - 1 \to 0 < 1$, the series converges by the Cauchy test. □

Exercise 10.32. *Prove that if* $\lim_{n\to\infty} na_n = g > 0$, *then the series* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Proof. Since $na_n \to g > 0$, there is n_0 such that for $n \geq n_0$, $na_n > g/2$, $a_n > g/(2n)$. Divergence of the series

$$\sum_{n=1}^{\infty} \frac{g}{2n} = \frac{g}{2} \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

implies divergence of the series $\sum_{n=1}^{\infty} a_n$ (comparison test). □

Exercise 10.33. *Find a convergent series* $\sum_{n=1}^{\infty} a_n$, $a_n > 0$ *such that* na_n *does not converge to* 0.

Solution. The series $\sum_{n=1}^{\infty} 1/n^2$ converges, but $na_n = 1/n \to 0$, so this is not a good example. However, we can make a trick. Consider the series

$$a_1 + a_2 + a_3 + \ldots = \frac{1}{1^2} + 0 + 0 + \frac{1}{2^2} + 0 + 0 + 0 + 0 + \frac{1}{3^2} + 0 + 0 + 0 + 0 + 0 + 0 + \frac{1}{4^2} + \ldots ,$$

i.e.,

$$a_{n^2} = \frac{1}{n^2}, \qquad a_k = 0 \text{ for } k \neq 1^2, 2^2, 3^2, \ldots$$

This series clearly converges. Suppose that $na_n \to 0$. Then the subsequence $n^2 a_{n^2}$ also converges to zero, so

$$n^2 a_{n^2} = n^2 \cdot \frac{1}{n^2} = 1 \to 0$$

which is a contradiction. Hence na_n does not converge to 0. □

Exercise 10.34. *Investigate convergence of the series* $\displaystyle\sum_{n=2}^{\infty} \frac{1}{(\ln n)^{\ln n}}$.

Solution. We have

$$\frac{1}{(\ln n)^{\ln n}} = \frac{1}{n^{\ln(\ln n)}} < \frac{1}{n^e} \quad \text{for } n > e^{e^e}.$$

Since the series $\sum_{n=1}^{\infty} 1/n^e$ converges, the original series converges, too (comparison test). □