Introduction to
Mathematical Logic and Set Theory

Andrei Marcus

September 30, 2022

Contents

Chapter 0

Description of the course

Logic is the study and use of valid reasoning. Logic has two aspects: *informal*, that is, the study of arguments in natural language, and formal, that is, of the study of inferences from the point of view of form, or in other words, the study of abstract rules of deduction. The earliest studies of formal logic are due to Aristotle. When we use abstract symbols in the formal study of inferences, we speak of symbolic logic; usually this it is divided into propositional logic and predicate logic.

Mathematical logic is part of Mathematics and of Logic. Its role is to rigorously define the idea of the truth value of a statement and to explore the application of formal (symbolic) logic methods in different branches of mathematics. Also, mathematical logic deals with the application of mathematical methods and techniques to the study of formal logic.

The development of mathematical logic was strongly motivated by the study of the foundations of mathematics, a study begun in the 19th century, and has important applications in philosophy or linguistics, but also in more recent fields such as computer science (logic programming, artificial intelligence, etc.).

Nowadays, mathematical logic is divided into four subdomains, each focusing on distinct aspects, but obviously the demarcation lines are not strict:

- set theory, which studies abstract collections of objects and the correspondences between them, playing an important role in the foundations of mathematics;

- proof theory, which essentially means the formal analysis of mathematical proofs.

- model theory, which is the formal study of mathematical structures, closely related to abstract algebra;

- recursion theory (or calculability theory), which studies the effective calculability of functions defined on the set of natural numbers, having an important role for the foundations of computer science;

In this introductory course dedicated to first year students from the Faculty of Mathematics and Computer Science we will touch on a small part of the mentioned subjects, often in an informal manner. Also included are some basic topics of Algebra, Arithmetic and Combinatorics closely related to the above, but which usually go beyond the framework of Mathematical Logic.

Chapter 1

PROPOSITIONAL LOGIC

In common language, by a proposition (sentence) we mean a statement about which we can decide whether it is true or false. We can form compound sentences, to which we also associate a value of truth, using words as: and, or, not, if and only if etc. From a mathematical point of view, such a definition is not satisfactory, so a formal approach is required.

1.1 The Formulas of Propositional Logic

Definition 1.1.1 a) **The symbols of propositional logic** are:

1. Parentheses (round brackets): (and).

2. Connectives (symbols of logical operations): $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$.

3. Atomic formulas: $p, q, r, \ldots, x_1, x_2, \ldots$

b) A **propositional formula** is a finite sequence of symbols which satisfy the following rules:

1. Atomic formulas are formulas.

2. If A and B are formulas, then $(\neg A)$, $(A \wedge B)$, $(A \vee B)$, $(A \rightarrow B)$, $(A \leftrightarrow B)$ are also formulas.

3. There are no formulas other than those described above.

Remark 1.1.2 a) In common parlance, the connectives $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ are read as: *non, si, sau, if ... then, if and only if* .

b) Usually, to simplify writing, some parentheses can be omitted by adopting a priority order of the connectives: $\neg$, apoi $\wedge$ and $\vee$, apoi $\rightarrow$ and $\leftrightarrow$. External brackets can also be omitted.

Example 1.1.3 1) The following sequences of symbols are formulas:
$(p \vee q) \rightarrow (r \leftrightarrow (\neg s))$,
$((p \vee (\neg q)) \vee r) \rightarrow (s \wedge (\neg s))$,
$((p \rightarrow q) \rightarrow r) \vee (s \wedge r)$,
$(\neg((p \wedge q) \vee (\neg r))) \rightarrow (t \vee p)$,
$((p \vee q) \rightarrow (p \vee r)) \leftrightarrow ((\neg q) \wedge (\neg p))$.

2) The following sequences of symbols are not formulas:
$p\wedge \rightarrow q$, $p \rightarrow$, $pq \wedge t$, $p \wedge q \vee r$ $p \wedge (q \rightarrow \wedge r)$, $(pq \wedge (r \wedge p \neg q)$.

Definition 1.1.4 a) We say that B **subformula** of the formula A if B is obtained during the construction of A.

b) We talk about a **substitution**, if in the formula A, an atom p or a subformula B is replaced by the formula C (notation: $A(C/p)$, respectively $A(C/B)$).

Example 1.1.5 1) $p \wedge q$, $t \vee p$ are subformulas of the formula $(\neg((p \wedge q) \vee (\neg r))) \rightarrow (t \vee p)$, while $p \rightarrow (t \vee p)$ is not.

2) If $A = (\neg((p\wedge q)\vee(\neg r))) \rightarrow (t\vee p)$, then for $C = r\wedge s$ we have $A(C/p) = (\neg(((r\wedge s)\wedge q)\vee(\neg r))) \rightarrow (t\vee(r\wedge s))$ and $A(C, p \wedge q) = (\neg((r \wedge s) \vee (\neg r))) \rightarrow (t \vee p)$.

1.2 Interpretation of propositional formulas

Definition 1.2.1 Let $V = \{0, 1\}$ be the set **truth values**. Here 0 corresponds to *false*, and 1 corresponds to *true*. A function $f : V^n \to V$ of n variables is called a **truth function**.

A truth function of n variables can be given by a **truth table** with $n + 1$ columns and 2^n rows. The first n columns contain all possible combinations of variables, and the last column contains the corresponding values of the function.

Also, a truth function can be visualized with the help of Euler-Venn diagrams or with the help of relay and switch circuits.

Definition 1.2.2 The most commonly used truth functions are the **fundamental logic operations** corresponding to the five connectives, which we define below using truth tables:

a) **Negation ("non"):** $\neg p$, defined by

p	$\neg p$
0	1
1	0

b) **Conjunction ("and"):** $p \wedge q$, defined by

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

c) **Disjunction ("or"):** $p \vee q$, defined by

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

d) **Implication ("if ... then"):** $p \to q$, defined by

p	q	$p \to q$
0	0	1
0	1	1
1	0	0
1	1	1

e) **Equivalence ("if and only if"):** $p \leftrightarrow q$, defined by

p	q	$p \leftrightarrow q$
0	0	1
0	1	0
1	0	0
1	1	1

Definition 1.2.3 If A is a formula and $\mathcal{A}$ is the set of atomic formulas contained in A, then an **interpretation** of A is a function $v : \mathcal{A} \to V = \{0, 1\}$. The element $v(p) \in V$ is called **truth value** of the atomic formula p.

Let $A = A(p_1, \ldots, p_n)$ be a formula containing the atoms $p_1, \ldots, p_n$, and let v be an interpretation of A. We denote by $\tilde{A} : V^n \to V$ truth function corresponding to A, obtained using the fundamental logical functions. Then **the truth value of the formula** A corresponding to the interpretation v is given by:

$$H_v(A) := \tilde{A}(v(p_1), \ldots, v(p_n)).$$

Example 1.2.4 In the table below we have the interpretations and truth values corresponding to the formula $A = A(p, q) = ((p \vee q) \wedge (\neg p)) \to q$ (highlighting a few subformulas):

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

We will see later that from Theorem 6.4.6 implies the following theorem, which we will use in the exercises below.

Theorem 1.2.5 *Any truth function of* $n \geq 1$ *variables can be expressed only with the help of the fundamental logical operations.*

Example 1.2.6 In addition to the fundamental logical operations, we mention the following truth functions:

1) **Addition and multiplication of modulo** 2, denoted by the symbols $\oplus$ respectively $\odot$.

2) **Sheffer's function (not and):** $p \mid q = \neg(p \wedge q)$.This is true if at most one of p or q is true.

3) **The Webb–Peirce function (neither-nor; not or):** $p \downarrow q = (\neg p) \wedge (\neg q)$. This is true, if both p and q are not true.

4) **Exclusive disjunction (or-or; xor):** $p \oplus q = \neg(p \leftrightarrow q)$. This is true, if exactly one of p or q is true.

Exercise 1 Draw up the truth tables for the functions in the example above.

Exercise 2 Verify with the help of truth tables the following equalities between functions:

1) $\neg p = 1 \oplus p$.

2) $p \wedge q = p \odot q$.

3) $p \vee q = p \oplus q \oplus p \odot q$.

4) $p \rightarrow q = 1 \oplus p \oplus p \odot q$.

5) $p \leftrightarrow q = 1 \oplus p \oplus q$.

Exercise 3 1) Write down all the truth functions of 1 and 2 variables, respectively.

2) How many truth functions of n variables do exist?

Exercise 4 Show that any truth function of $n \geq 1$ variables can be expressed only by negation and conjunction (or only by negation and disjunction. Specifically, check the following equalities:

1) $p \vee q = \neg((\neg p) \wedge (\neg q))$.

2) $p \wedge q = \neg((\neg p) \vee (\neg q))$.

3) $p \rightarrow q = \neg(p \wedge (\neg q)) = \neg p \vee q$.

4) $p \leftrightarrow q = (\neg(p \wedge (\neg q))) \wedge (\neg(q \wedge (\neg p)))$.

5) $p \oplus q = (p \vee q) \wedge (\neg(p \wedge q))$.

Exercise 5 Show that any truth function of $n \geq 1$ variables can be expressed only by negation and implication. Specifically, express the conjunction, disjunction and equivalence using only negation and implication.

Exercise 6 Show that any truth function of $n \geq 1$ variables can be expressed only with the help of the Sheffer function. Specifically, verify the following equalities:

1) $\neg p = p \mid p$.

2) $p \wedge q = (p \mid q) \mid (p \mid q)$.

3) $p \vee q = (p \mid p) \mid (q \mid q)$.

4) $p \rightarrow q = p \mid (q \mid q) = p \mid (p \mid q)$.

Exercise 7 Show that any function of truth can be expressed only with the help of the Webb-Peirce function. Specifically, verify the following equalities:

0) $p \downarrow q = \neg(p \vee q)$.

1) $\neg p = p \downarrow p$.

2) $p \wedge q = (p \downarrow p) \downarrow (q \downarrow q)$.

3) $p \vee q = (p \downarrow q) \downarrow (p \downarrow q)$.

4) $p \rightarrow q = ((p \downarrow p) \downarrow q) \downarrow ((p \downarrow p) \downarrow q)$.

Definition 1.2.7 a) A formula is called **satisfiable** if it has an interpretation for which the truth value is 1.

b) If there is no such interpretation, the formula is called a **contradiction**, and we denote it by $\mathsf{0}$.

c) A formula is called a **tautology**, if for any interpretation the truth value is 1, and we denote it by $\mathsf{1}$.

Definition 1.2.8 We introduce two *relations* between formulas:

a) If the formula $A \rightarrow B$ is a tautology, then we say that formula B **follows** from the formula A, and we denote this by $A \Rightarrow B$.

In mathematical theorems we use the following expressions: *if* A*, then* B*;* A *is a sufficient condition for* B*;* B *is a necessary condition for* A*.*

b) If formula $A \leftrightarrow B$ is a tautology, then we say that A is **equivalent** to B, and we denote this by $A \Leftrightarrow B$.

In mathematical theorems we use the following expressions: A *is a necessary and sufficient condition for* B*;* B *if and only if* A*;* A *is equivalent to* B.

Example 1.2.9 1) For any formula A, the formula $(\neg A) \vee A$ is a tautology, and $(\neg A) \wedge A$ is a contradiction.

2) A is a contradiction if and only if $\neg A$ is a tautology.

3) A is a tautology if and only if $\neg A$ is a contradiction.

4) If $A = p \wedge (\neg p)$, $B = p \vee (\neg p)$, $C = p \to p$, $D = p \to q$, $E = (\neg p) \vee q$, $F = p \leftrightarrow (\neg p)$, then B and C are tautologies, A and F are contradictions, D and E are satisfiable. Also, these pairs are equivalent.

Remark 1.2.10 Let A be a tautology, p an atomic formula and B a subformula of A. Then for any formula C, $A(C/p)$ is a tautology. If $C \Leftrightarrow B$, then $A(C/B)$ is a tautology.

Theorem 1.2.11 We list some important tautologies. Let A, B, C be propositional formulas.

1) $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$, $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$ (associativity),

2) $A \wedge B \Leftrightarrow B \wedge A$, $A \vee B \Leftrightarrow B \vee A$ (commutativity),

3) $A \wedge (A \vee B) \Leftrightarrow A$, $A \vee (A \wedge B) \Leftrightarrow A$ (absorption),

4) $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$, $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$ (distributivity),

5) $A \wedge A \Leftrightarrow A$, $A \vee A \Leftrightarrow A$ (idempotence),

6) $A \wedge 1 \Leftrightarrow A$, $A \vee 0 \Leftrightarrow A$,

7) $A \wedge 0 \Leftrightarrow 0$, $A \vee 1 \Leftrightarrow 1$,

8) $\neg(\neg A) \Leftrightarrow A$ (the law of double negation),

9) $A \vee (\neg A) \Leftrightarrow 1$ (the law of excluded middle), $A \wedge (\neg A) \Leftrightarrow 0$ (the law of contradiction),

10) $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$, $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$ (De Morgan's laws),

11) $A \leftrightarrow B \Leftrightarrow (A \to B) \wedge (B \to A)$ (the law of equivalence),

12) $A \to B \Leftrightarrow (\neg A) \vee B$ (the law of implication),

13) $A \to B \Leftrightarrow (\neg B) \to (\neg A)$ (the law of contraposition),

14) $(A \wedge B) \to C \Leftrightarrow A \to (B \to C)$ (the law of separation/union of premises (also called the law of exportation/importation)),

15) $A \to (B \to C) \Leftrightarrow B \to (A \to C)$ (the law of permutation of premises).

Exercise 8 Check the tautologies (1) - (15) of the theorem above with the help of truth tables.

1.3 The decision problem

The decision problem in propositional logic means finding an algorithm that determines whether a propositional formula is a *tautology, contradiction, or satisfiable* as well as finding the correct methods of deduction. We will discuss three methods, in principle equivalent: truth tables, normal forms and formal deduction based on deduction schemes.

1.3.1 The method of truth tables

We have already seen this method in the previous paragraph, which is effective in the case of formulas with a small number of atoms.

1.3.2 The method of normal forms

Definition 1.3.1 Let $A = A(x_1, x_2, \ldots, x_n)$ be a propositional formula.

a) A is an **elementary conjunction** if it is a conjunction whose factors are atoms or negations of atoms.

b) A is a **elementary disjunction** if it is a disjunction whose terms are atoms or negations of atoms.

Example 1.3.2 a) The formulas $A = x_1 \wedge \neg x_2 \wedge \neg x_3$, $B = x_1 \wedge x_2 \wedge x_3$, $C = \neg x_1 \wedge \neg x_2 \wedge \neg x_3$ are elementary conjunctions.

b) The formulas $A = x_1 \vee \neg x_2 \vee \neg x_3$, $B = x_1 \vee x_2 \vee x_3$, $C = \neg x_1 \vee \neg x_2 \vee \neg x_3$ are elementary disjunctions.

Definition 1.3.3 a) The formula $A = A(x_1, x_2, \dots, x_n)$ are **conjunctive normal form** (**CNF**), if it is a conjunction of elementary disjunctions, that is:

$$A = A_1 \wedge A_2 \wedge \dots \wedge A_m,$$

where subformula $A_i = A_i(x_1, x_2, \dots, x_n)$ is an elementary disjunction, for any $i = 1, 2, \dots, m$.

b) We say that formula $B = B(x_1, x_2, \dots, x_n)$ are **disjunctive normal form** (**DNF**), if it is a disjunction of elementary conjunctions, that is:

$$B = B_1 \vee B_2 \vee \dots \vee B_m,$$

where the subformula $B_i = B_i(x_1, x_2, \dots, x_n)$ is an elementary conjunction, for any $i = 1, 2, \dots, m$.

Remark 1.3.4 Any propositional formula A is logically equivalent to a CNF, respectively to a DNF (not necessarily uniquely determined). The formula A is brought to a CNF, respectively to a DNF, through a finite sequence of logical equivalences, using the fundamental laws of the propositional logic presented in Theorem 1.2.11, as follows:

1. The formula A is expressed only with the connectives $\neg$, $\wedge$, $\vee$, using the law of implication and the law of equivalence.
2. The negation is passed only on the atoms, using De Morgan's laws and the law of double negation.
3. Conjunctions of disjunctions (for CNF), respectively disjunctions of conjunctions (for DNF) are obtained, using distributivity, absorption, idempotence, commutativity or associativity.

Example 1.3.5 Let $A = \neg x \to x \wedge y$. Applying the above, we get:

$$A = \neg x \to x \wedge y \Leftrightarrow \neg\neg x \vee (x \wedge y) \Leftrightarrow x \vee (x \wedge y)$$

and so we obtained a DNF. Next, we have

$$x \vee (x \wedge y) \Leftrightarrow (x \vee x) \wedge (x \vee y)$$

and we obtain a CNF. Now using idempotence, we have:

$$(x \vee x) \wedge (x \vee y) \Leftrightarrow x \wedge (x \vee y)$$

and we obtain another CNF. Applying the absorption, we have:

$$x \wedge (x \vee y) \Leftrightarrow x,$$

which is yet another CNF, which can also be regarded as a DNF.

Remark 1.3.6 The method of normal forms is applied as follows. Let $C = C(x_1, x_2, \dots, x_n)$ be a propositional formula, and let $A = A_1 \wedge A_2 \wedge \dots \wedge A_m$ be a CNF, respectively $B = B_1 \vee B_2 \vee \dots \vee B_m$ a DNF to which C is logically equivalent. Then:

a) C is a tautology if and only if in the CNF A, for any $i = 1, 2, \dots, m$, A_i contains at least one atom together with its negation;

b) C is a contradiction if and only if in the DNF B, for any $i = 1, 2, \dots, m$, B_i contains at least one atom together with its negation.

Example 1.3.7 Let's solve the decision problem by the normal forms method.

a) Let $C = x \wedge \neg y \to x$. We bring C to a normal form:

$$C = x \wedge \neg y \to x \Leftrightarrow \neg(x \wedge \neg y) \vee x \Leftrightarrow (\neg x \vee \neg\neg y) \vee x \Leftrightarrow (\neg x \vee y) \vee x \Leftrightarrow \neg x \vee y \vee x.$$

We got the formula $A = \neg x \vee y \vee x$, which may be regarded both as a CNF, and a DNF. Regarding A as a CNF with a single factor, x appears together with its negation $\neg x$, hence φ is a tautology.

b) Let $C = \neg x \wedge (\neg x \vee y \to x)$. We bring C to a normal form:

$$\begin{aligned} C = \neg x \wedge (\neg x \vee y \to x) \Leftrightarrow \neg x \wedge (\neg(\neg x \vee y) \vee x) \Leftrightarrow \neg x \wedge ((\neg\neg x \wedge \neg y) \vee x) \Leftrightarrow \\ \Leftrightarrow \neg x \wedge ((x \wedge \neg y) \vee x) \Leftrightarrow (\neg x \wedge x \wedge \neg y) \vee (\neg x \wedge x) \end{aligned}$$

We got the DNF $B = (\neg x \wedge x \wedge \neg y) \vee (\neg x \wedge x)$. In every factor of B we have the atom x together with its negation $\neg x$, hence C is a contradiction.

c) Let $C = (x \to y) \wedge (y \to z)$. We bring C to a normal form:

$$C = (x \to y) \wedge (y \to z) \Leftrightarrow (\neg x \vee y) \wedge (\neg y \vee z).$$

We got a CNF $A = (\neg x \vee y) \wedge (\neg y \vee z)$, and we see that C is not tautology. We also find a DNF:

$$A = (\neg x \vee y) \wedge (\neg y \vee z) \Leftrightarrow (\neg x \wedge \neg y) \vee (\neg x \wedge z) \vee (y \wedge \neg y) \vee (y \wedge z).$$

We got the DNF $B = (\neg x \wedge \neg y) \vee (\neg x \wedge z) \vee (y \wedge \neg y) \vee (y \wedge z)$, from which we read that C is not a contradiction, hence C is a satisfiable formula.

Exercise 9 Bring to the normal conjunctive form and to the normal disjunctive form and solve the decision problem for the formulas:
1) $((x \to y) \to (z \to \neg x)) \to (\neg y \to \neg z)$.
2) $((((x \to y) \to \neg x) \to \neg y) \to \neg z) \to z$.
3) $(x \to (y \to z)) \to ((x \to \neg z) \to (x \to \neg y))$.
4) $(\neg x \to \neg y) \to ((y \wedge z) \to (x \wedge z))$.
5) $((x \to y) \to \neg x) \to (x \to (y \wedge x))$.
6) $\neg((x \wedge y) \to \neg x) \wedge \neg((x \wedge y) \to \neg y)$.
7) $(z \to x) \to (\neg(y \vee z) \to x)$.
8) $\neg((x \wedge y) \to x) \vee (x \wedge (y \vee z))$.
9) $\neg(x \wedge (y \vee z)) \to \neg((x \wedge y) \vee z)$.

1.3.3 Rules of inference

Definition 1.3.8 We say that the propositional formula B is a **consequence** of the set of formulas $\Sigma = \{A_1, \dots, A_n\}$ (where $n \geq 0$), if every interpretation which makes $A_1, \dots, A_n$ true, also makes the formula B true.

We denote this by

$$A_1, \dots, A_n \models B \qquad \text{or} \qquad \Sigma \models B \qquad \text{or} \qquad \frac{A_1, \dots, A_n}{B}$$

and we call it an **inference rule**. The formulas $A_1, \dots, A_n$ are called **premises**, and B is called the **conclusion**.

It is clear by the definition that we have $A_1, \dots, A_n \models B$ exactly when the formula

$$A_1 \wedge \cdots \wedge A_n \to B$$

is a tautology, that is, the relation $A_1 \wedge \cdots \wedge A_n \Rightarrow B$ holds.

More generally, if $\Gamma = \{B_1, \dots, B_m\}$ is a set of formulas, then we denote $\Sigma \models \Gamma$ if $\Sigma \models B_j$ for any $j = 1, \dots, m$.

Remark 1.3.9 1) If in particular $n = 0$ (that is, $\Sigma = \emptyset$), then this means that B is a tautology (respectively every formula from Γ is a tautology).

2) We have $\Sigma \models \Gamma$ if and only if the formula $(A_1 \wedge \cdots \wedge A_n) \to (B_1 \wedge \cdots \wedge B_m)$ is a tautology.

3) We have the *reflexivity* property$A \models A$, because the formula $A \to A$ is a tautology, by the law of implication and and the law excluded middle. More generally, if $\Gamma \subseteq \Sigma$ are sets of formulas, then $\Sigma \models \Gamma$.

Example 1.3.10 We present below some rules of inference of the Aristotelian classical logic. They can be easily verified with the help of truth tables and are frequently used in proving mathematical theorems. Note that some variants are obtained from others by replacing a formula with its negation.

1. **Classical deductive argument forms.**

   (a) $\frac{A,\ A \to B}{B}$ **(modus ponendo ponens** or briefly **modus ponens (MP))**

   (b) $\frac{\neg A,\ \neg A \to \neg B}{\neg B}$ **(modus tollendo tollens)**

   (c) $\frac{\neg A,\ \neg A \to B}{B}$ **(modus tollendo ponens)**

   (d) $\frac{A,\ A \to \neg B}{\neg B}$ **(modus ponendo tollens)**

2. **Reductio ad absurdum.**

   (a) $\frac{B,\ \neg A \to \neg B}{A}$; $\frac{\neg B,\ \neg A \to B,}{A}$; $\frac{B,\ A \to \neg B}{\neg A}$; $\frac{\neg B,\ A \to B}{\neg A}$.

   (b) $\frac{(\neg A) \to B,\ (\neg A) \to (\neg B)}{A}$; $\frac{A \to B,\ A \to (\neg B)}{\neg A}$.

3. **Contraposition.**

$$\frac{A \to B}{\neg B \to \neg A}$$

4. **Hypothetical syllogism.**

$$\frac{A \to B, B \to C}{A \to C}$$

5. **Disjunctive syllogism.**

$$\frac{A \lor B,\ \neg A}{B}$$

6. **The method of case analysis (proof by exhaustion).**

$$\frac{B \lor C,\ B \to A,\ C \to A}{A}$$

Exercise 10 Verify the validity of the above inference rule with the help of truth tables, respectively using the method of normal forms.

Remark 1.3.11 We present some general properties of the inference rule, which are useful in proving mathematical theorems:

1. If $A_1, \dots, A_n \models B_j$ (for any $j = 1, \dots, m$) and $B_1, \dots, B_m \models C$, then $A_1, \dots, A_n \models C$ (this is the property of *transitivity*, which generalizes the hypothetical syllogism).
2. If $A_1 \models A_2$,…,$A_{n-1} \models A_n$, and $A_n \models A_1$, then the formulas $A_1, \dots, A_n$ are equivalent (this is the **method of cyclic proof**).
3. $\Sigma \cup \{A\} \models B$ if and only if $\Sigma \models A \to B$.

Exercise 11 Prove the above properties.

Remark 1.3.12 Many mathematical proofs become easier if we replace a given rule with an equivalent one.

1. **Direct proof**: we replace $\frac{A}{B \to C}$ with $\frac{A,B}{C}$.
2. **Proof by contraposition**: we replace $\frac{A,\ B}{C}$ with $\frac{A,\ \neg C}{\neg B}$.
3. **Indirect proof (proof by contradiction)**: instead of $\frac{A}{B}$ we show that $A \land (\neg B)$ is a contradiction.

Exercise 12 Prove the equivalence of the above inference rules.

1.3.4 Formal deduction

Another approach to the decision problem is based on the manipulation of symbols starting from a few axioms and inference rules, and does not use interpretation of formulas. We will see that the method of formal deduction is equivalent to that based on truth tables.

1.3.13 We briefly present Hilbert's calculus. (There are also other approaches, such as Gentzen's sequential calculus.) This method starts with the following data:

- several special tautologies, called **the axioms of propositional logic**.

  A1: $A \to (B \to A)$

  A2: $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$

  A3: $((\neg B) \to (\neg A)) \to (((\neg B) \to A) \to B))$, where A, B, C are arbitrary formulas;

- The **Modus Ponens (MP)** inference rule, that is $\frac{A, A \to B}{B}$.

Exercise 13 Check that the formulas **A1**, **A2** and **A3** above are tautologies, by using the method of truth tables, respectively the method of normal forms.

Definition 1.3.14 Let $A_1, \dots, A_n$ $(n \geq 0)$ be propositional formulas. A **deduction** from the formulas $A_1, \dots, A_n$ (called **premises** or **hypotheses**) is a finite sequence $E_1, \dots, E_k$ of formulas such that for any $i = 1, \dots, k$ we have:

(1) E_i is an axiom, or

(2) there is l such that $E_i = A_l$, or

(3) E_i is obtained from E_j, E_l $(j, l < i)$ by using the (MP) rule.

Definition 1.3.15 a) We say that the formula B **deducible** from the formulas $A_1, \dots, A_n$ (notation: $A_1, \dots, A_n \vdash B$), if B is the last term of a deduction from the formulas $A_1, \dots, A_n$. If $n = 0$, then we denote $\vdash B$.

The definition generalizes immediately to the case of two sets Σ and Γ of formulas; we denote $\Sigma \vdash \Gamma$ if $\Sigma \vdash B$ for any $B \in \Gamma$.

b) We say that the set of formulas Σ is **contradictory**, if there is a formula A, such that $\Sigma \vdash A$ and $\Sigma \vdash \neg A$. Otherwise, we say that Σ is **consistent**.

Example 1.3.16 a) Prove that $\vdash A \to A$.

1. $(A \to ((A \to A) \to A)) \to ((A \to (A \to A)) \to (A \to A))$ A2
2. $A \to ((A \to A) \to A)$ A1
3. $(A \to (A \to A)) \to (A \to A)$ 1,2 MP
4. $A \to (A \to A)$ A1
5. $A \to A$ 3,4 MP

b) Prove that $A \to B,\ B \to C \vdash A \to C$.

1. $(B \to C) \to (A \to (B \to C))$ A1
2. $B \to C$ Hypothesis
3. $A \to (B \to C)$ 1,2 MP
4. $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$ A2
5. $(A \to B) \to (A \to C)$ 4,3 MP
6. $A \to B$ Hypothesis
7. $A \to C$ 5,6 MP

c) Prove that $A,\ \neg A \vdash B$.

1. $\neg A$ Hypothesis
2. $(\neg A) \to ((\neg B) \to (\neg A))$ A1
3. $(\neg B) \to (\neg A)$ 1,2 MP
4. A Hypothesis
5. $A \to ((\neg B) \to A)$ A1
6. $(\neg B) \to A$ 4,5 MP
7. $((\neg B) \to (\neg A)) \to (((\neg B) \to A) \to B)$ A3
8. $((\neg B) \to A) \to B$ 3,7 MP
9. B 6,8 MP

We see that this method is not very easy to apply. The following remarks simplify things somewhat.

Remark 1.3.17 a) If $\Sigma \vdash B$ and $\Sigma \vdash B \to C$, then $\Sigma \to C$.

b) If $\Sigma \subseteq \Delta$ and $\Sigma \vdash B$, then $\Delta \vdash B$.

c) If $\Sigma \vdash \Gamma$ and $\Gamma \vdash B$, then $\Sigma \vdash B$.

d) If $\Sigma \vdash B \wedge \neg B$, then $\Sigma \vdash C$ for any formula C .

e) (*Herbrand's Theorem*: $\Sigma \vdash B \to C$ if and only if $\Sigma \cup \{B\} \vdash C$.

Example 1.3.18 To show that $A \to B, B \to C \vdash A \to C$ it is enough to show that $A,\ A \to B,\ B \to C \vdash C$. For this, we have:

1. A Hypothesis
2. $A \to B$ Hypothesis
3. B 1,2 MP
4. $B \to C$ Hypothesis
5. C 3,4 MP.

The following theorem says that the method of deduction is based on truth values ($\Rightarrow$, $\models$) is equivalent to the formal deduction ($\vdash$)). The first implication is easier to prove, the second is difficult.

Theorem 1.3.19 (The Frege–Łukasiewicz completeness theorem) $\Sigma \vdash B$ *if and only if* $\Sigma \models B$.

Chapter 2

FIRST-ORDER LOGIC

We have seen that propositional logic formalizes the use of logical operations *non, and, or, if ... then, if and only if*). First-order logic (also called predicate logic) goes further by introducing *quantifiers*, to formalize the notions of *for all* and *exists*. Thus, first-order logic will be useful for formalizing many more mathematical theories.

In the first-order logic only the variables are quantified, in the second-order logic the predicates (or sets) are also quantified, etc.

2.1 The notion of a predicate

Definition 2.1.1 Let M a set nonempty and let $n \in \mathbb{N}^*$. An **n-ary predicate on the set** M is a subset of the set M^n (that is, an n-ary relation on M).

Remark 2.1.2 In common parlance, an n-ary predicate on the set M is an "open" statement $P(x_1, \dots, x_n)$, in which we can replace the variabless $x_1, \dots, x_n$ with the elements $a_1, \dots a_n \in M$ to obtain the proposition $P(a_1, \dots, a_n)$. In this case,

$$\{(a_1, \dots, a_n) \in M^n \mid P(a_1, \dots, a_n) \text{ is true } \}$$

is an n-ary relation, hence an n-ary predicate on M. However, this approach is not precise enough.

Example 2.1.3 a) "$x + y = z$" is a predicate of 3 variables on $M = \mathbb{R}$.

b) "$x < y$" is a binary predicate on $M = \mathbb{N}$.

c) "$|x| = 1$" is a unary predicate on $M = \mathbb{C}$.

2.2 First-order languages

The symbols and the rules of construction of formulas given below constitute a **first-order language**.

Definition 2.2.1 The symbols of a first-order language $\mathcal{L}$ are the following:

1. Parentheses: (and).
2. Connectives: $\neg, \wedge, \vee, \to, \leftrightarrow$.
3. Quantifiers: $\forall$ (*for any*) and $\exists$ (*exista*).
4. The symbol of equality: $=$.
5. Variables: $x, y, z, \dots$.
6. Constants: $a, b, c, \dots$.
7. Functions (operations): $f, g, \dots$.
8. Predicates: $P, Q, \dots$.

We assume in addition that for every function and for every predicate, the **arity** ≥ 1 is given (that is, the number of its variables sale). The quantifiers may appear only before the variables.

The use of symbols depends on the mathematical theory we want to formalize.

Example 2.2.2 1) *The language* $\mathcal{L}_S$ of set theory uses one binary predicate $\in$ ("belongs to, is an element of").

2) *The language* $\mathcal{L}_G$ of group theory uses the constant 1 (the symbol of the neutral element), taking the inverse is a unary function, and the product is a binary function.

3) *The language* $\mathcal{L}_{\mathbb{N}}$ of the theory of natural numbers uses the constant 0 and three operations $s, +, \cdot$: the successor function s is unary, the addition and the multiplication are binary.

Definition 2.2.3 a) **The terms (the expressions)** of the first-order language $\mathcal{L}$ are finite sequences of symbols which satisfy the rules:

1. Any variable is a term.
2. Any constant is a term.
3. If f is a function of n variables and $t_1, \dots, t_n$ are terms, then $f(t_1, \dots, t_n)$ is aterm. (Often, instead of $f(x, y)$ we denote xfy, for instance, $x + y$.)
4. Other terms do not exist.

b) The **formulas** of the first-order language $\mathcal{L}$ are finite sequences of symbols which satisfy the rules:

1. If P is an n-ary predicate and $t_1, \dots, t_n$ are terms, then $P(t_1, \dots, t_n)$ is a formula.
2. If t_1 and t_2 are terms, then $(t_1 = t_2)$ is a formula.
3. If φ, ψ are formulas, then $(\neg\varphi)$, $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$, $(\varphi \to \psi)$, $(\varphi \leftrightarrow \psi)$ are formulas. (When possible, we will omit some of the parentheses.)
4. If φ is a formula and x is a variable, then $\forall x\varphi$ and $\exists x\varphi$ are formulas. In this case, we say that x is a **quantified (bound)** variable.
5. Other formulas do not exist.

The formulas of type 1 and 2 are called **atomic formulas**.

Definition 2.2.4 Let x be a variable of the language $\mathcal{L}$. We say that x is a **free variable** of the formula φ if:

1. φ is an atomic formula, and x appears in φ.
2. φ has the form $(\neg\alpha)$ and x is a free variable in α.
3. φ is of the form $(\alpha \vee \beta)$ or $(\alpha \wedge \beta)$ or $(\alpha \to \beta)$ or $(\alpha \leftrightarrow \beta)$, and x is free variable in α or in β.
4. φ is of the form $\forall y\alpha$ or $\exists y\alpha$, where y is different from x, and x is a free variable in α.

We say that the variable x is **bound**, if it is not free. A formula in which every variable is bound is called a **closed formula**.

Example 2.2.5 1) In the formula "$\forall x(x = y)$", the variable x is bound, while y is free. The formula "$\forall x\forall y(x \wedge y = y \wedge x)$" is closed.

2) Consider the formula $\forall x((x = y) \wedge (P(x) \to Q(y)))$; then $x = y$, $P(x) \to Q(y)$, $P(x)$ are subformulas, but $\forall x(x = y)$ is not a subformula.

Exercise 14 Let f, g, h symbols of functions of 1, 2 respectively 3 variables, and let P, Q symbols of predicates of 1 respectively 3 variables.

1. Are terms the following words?
   (a) $f(g(x, y))$.
   (b) $g(f(z), h(x, y, z))$.
   (c) $f(g(x), h(x, y, z))$.
2. Are formulas the following words?
   (a) $Q(x, f(x), h(y, z, z))$.
   (b) $(P(x) \to (\forall y)(Q(x, y, z) \wedge P(g(x, y))))$.
   (c) $Q(P(x), f(y), z)$.
   (d) $f(h(x, y, z))$.

Exercise 15 Write down all the subformulas of trhe formulas:
a) $Q(f(x), g(x, y))$;
b) $\exists x Q(x, y) \to \neg(P(g(x, y)) \wedge \forall x P(z))$.

Exercise 16 Describe the set of terms of a first-order language, if we are given:
a) a variable x and a symbol of unary function (of one variable) f;
b) a variable x and a symbol of binary function (of two variables) f;

2.3 The structure of a first-order language. Model

We now give meaning and truth values to formulas of a first-order language.

Definition 2.3.1 A **structure** $\mathcal{M}$ of a first-order language $\mathcal{L}$ consists of the following data:

1. A nonempty set M, which we call **univers**, and we denote it by $|\mathcal{M}|$.
2. To every constant a corresponds an element $\tilde{a} \in M$.
3. To every symbol of n-ary function f corresponds a function $\widetilde{f} : M^n \to M$.
4. To every symbol of n-ary predicate P corresponds an n-ary predicate $\widetilde{P}$ on the set M (that is, a subset $\widetilde{P} \subseteq M^n$).
5. To the symbol of equality corresponds the equality relation on M.

We often simply denote $\widetilde{a}$ by a, $\widetilde{f}$ by f, and $\widetilde{P}$ by P. In what follows, we consider a fixed first-order language $\mathcal{L}$, and a structure $\mathcal{M}$ of $\mathcal{L}$, with $M = |\mathcal{M}|$.

Definition 2.3.2 a) If $\mathcal{V}$ is the set of variables of $\mathcal{L}$, then a function $s : \mathcal{V} \to M$ is called an **interpretation** of the structure $\mathcal{M}$.

b) We define inductively **the value** $H_s^{\mathcal{M}}(t) \in M$ of the term t, corresponding to the interpretation s, as follows:

1. For every variable x, we have $H_s^{\mathcal{M}}(x) = s(x)$.
2. For every constant a, we have $H_s^{\mathcal{M}}(a) = \widetilde{a}$.
3. For every n-ary function f and terms $t_1, \ldots, t_n$ we have
$$H_s^{\mathcal{M}}(f(t_1, \ldots, t_n)) = \widetilde{f}(H_s^{\mathcal{M}}(t_1), \ldots, H_s^{\mathcal{M}}(t_n)).$$

c) We define inductively **the value** $H_s^{\mathcal{M}}(\varphi) \in V = \{0, 1\}$ of the formula φ, corresponding to the interpretation s, as follows:

1. for any n-ary predicate P and any terms $t_1, \ldots, t_n$, $H_s^{\mathcal{M}}(P(t_1, \ldots, t_n)) = 1$ if $(H_s^{\mathcal{M}}(t_1), \ldots, H_s^{\mathcal{M}}(t_n)) \in \widetilde{P}$, otherwise $H_s^{\mathcal{M}}(P(t_1, \ldots, t_n)) = 0$.
2. $H_s^{\mathcal{M}}(t_1 = t_2) = 1$, if $H_s^{\mathcal{M}}(t_1) = H_s^{\mathcal{M}}(t_2)$, otherwise $H_s^{\mathcal{M}}(t_1 = t_2) = 0$.
3. $H_s^{\mathcal{M}}(\neg\varphi) = 1$, if $H_s^{\mathcal{M}}(\varphi) = 0$, otherwise $H_s^{\mathcal{M}}(\neg\varphi) = 0$.
   $H_s^{\mathcal{M}}(\varphi \vee \psi) = 1$, if $H_s^{\mathcal{M}}(\varphi) = 1$ or $H_s^{\mathcal{M}}(\psi) = 1$, otherwise $H_s^{\mathcal{M}}(\varphi \vee \psi) = 0$.
   $H_s^{\mathcal{M}}(\varphi \wedge \psi) = 1$, if $H_s^{\mathcal{M}}(\varphi) = H_s^{\mathcal{M}}(\psi) = 1$, otherwise $H_s^{\mathcal{M}}(\varphi \wedge \psi) = 0$.
   $H_s^{\mathcal{M}}(\varphi \to \psi) = 0$, if $H_s^{\mathcal{M}}(\varphi) = 1$ and $H_s^{\mathcal{M}}(\psi) = 0$, otherwise $H_s^{\mathcal{M}}(\varphi \to \psi) = 1$.
   $H_s^{\mathcal{M}}(\varphi \leftrightarrow \psi) = 1$, if $H_s^{\mathcal{M}}(\varphi) = H_s^{\mathcal{M}}(\psi)$, otherwise $H_s^{\mathcal{M}}(\varphi \leftrightarrow \psi) = 0$.
4. Consider the function (interpretation)
$$s(x|m) : \mathcal{V} \to M, \qquad s(x|m)(y) = \begin{cases} s(y), & \text{if } y \neq x \\ m & \text{if } y = x \end{cases}.$$
   Then:
   $H_s^{\mathcal{M}}(\forall x\varphi) = 1$ if and only if for any $m \in M$ we have $H_{s(x|m)}^{\mathcal{M}}(\varphi) = 1$.
   $H_s^{\mathcal{M}}(\exists x\varphi) = 1$ if and only if there exists $m \in M$ such that $H_{s(x|m)}^{\mathcal{M}}(\varphi) = 1$.

Definition 2.3.3 a) We say that $\mathcal{M}$ is **a model** of φ (or that $\mathcal{M}$ **satisfies** φ), if $H_s^{\mathcal{M}}(\varphi) = 1$ for any interpretation s of $\mathcal{M}$. Notation: $\mathcal{M} \models \varphi$.

We say that $\mathcal{M}$ is **a model** for the set Γ of formulas (or that $\mathcal{M}$ **satisfies** Γ), if $\mathcal{M} \models \gamma$ for any $\gamma \in \Gamma$. Notation: $\mathcal{M} \models \Gamma$.

One can show by induction that:

Theorem 2.3.4 1) *If the interpretations* s *and* r *coincide on the variables which occur in the term* t*, then* $H_s^{\mathcal{M}}(t) = H_r^{\mathcal{M}}(t)$.

2) *If* s *and* r *coincide on the free variables which occur in the formula* φ*, then* $H_s^{\mathcal{M}}(\varphi) = H_r^{\mathcal{M}}(\varphi)$.

Corollary 2.3.5 *If* σ *is a closed formula, then* $\mathcal{M} \models \sigma$ *if and only if there exists an interpretation* s *such that* $H_s^{\mathcal{M}}(\sigma) = 1$. (Hence the value of a closed formula is independent of the fixed interpretation of the structure.)

Definition 2.3.6 a) A formula φ is a **tautology** (**identically true**), if every structure $\mathcal{M}$ is a model of φ. The formula φ is called a **contradiction**, if $\neg\varphi$ is a tautology.

b) If the formula $\varphi \to \psi$ is tautology, then we say that ψ **follows** from φ, and we denote $\varphi \Rightarrow \psi$.

c) If the formula $\varphi \leftrightarrow \psi$ is a tautology, then we say that φ is **equivalent** to ψ, and we denote $\varphi \Leftrightarrow \psi$.

Example 2.3.7 1) For any formula φ we have that $\varphi \to \varphi$ is a tautology, while $\neg(\varphi \to \varphi)$ is a contradiction.

2) $\forall y(y = y)$ is a tautology, while $\exists y(\neg(y = y))$ is a contradiction.

3) If φ is a tautology, then every generalization $\forall x_1 \ldots \forall x_n \varphi$ is a tautology.

Remark 2.3.8 a) φ is a contradiction if and only if for any structure $\mathcal{M}$ and any interpretation $s : \mathcal{V} \to |\mathcal{M}|$, we have $H_s^{\mathcal{M}}(\varphi) = 0$.

b) If φ is a contradiction, then it does not have a model. The converse holds only for closed formulas.

c) $\varphi \Rightarrow \psi$ if and only if for any structure $\mathcal{M}$ and for any interpretation $s : \mathcal{V} \to |\mathcal{M}|$, if s satisfies on φ, then it also satisfies ψ.

d) If $\varphi \Rightarrow \psi$, then every model $\mathcal{M}$ of φ is also a model of ψ. The converse holds only for closed formulas.

e) $\varphi \Leftrightarrow \psi$ if and only if for any structure $\mathcal{M}$ and for any interpretation $s : \mathcal{V} \to |\mathcal{M}|$, s satisfies φ if and only if it satisfies ψ.

f) If $\varphi \Leftrightarrow \psi$, then φ has exactly the same models as ψ. The converse holds only for closed formulas.

g) Alonzo Church proved in 1936 that a general decision procedure cannot be given for a first-order language.

2.3.9 We present several **important tautologies**, which will be used in the proofs from the next chapters. Let A, B and C be formulas of the first-order language $\mathcal{L}$ such that in C, the variable x is not free.

(1) $\forall x \forall y A \Leftrightarrow \forall y \forall x A$, $\exists x \exists y A \Leftrightarrow \exists y \exists x A$

(2) $(\exists x)(\forall y)A \Rightarrow (\forall y)(\exists x)A$, $\forall x A \Rightarrow \exists x A$

(3) $\forall x(A \wedge B) \Leftrightarrow \forall x A \wedge \forall x B$

(4) $\exists x(A \vee B) \Leftrightarrow \exists x A \vee \exists x B$

(5) $\forall x A \vee \forall x B \Rightarrow \forall x(A \vee B)$

(6) $\exists x(A \wedge B) \Rightarrow \exists x A \wedge \exists x B$

(7) $\neg\forall x A \Leftrightarrow \exists x(\neg A)$, $\neg\exists x A \Leftrightarrow \forall x(\neg A)$ (De Morgan laws)

(8) $C \wedge \forall x A \Leftrightarrow \forall x(C \wedge A)$,
$C \vee \forall x A \Leftrightarrow \forall x(C \vee A)$,
$C \wedge \exists x A \Leftrightarrow \exists x(C \wedge A)$,
$C \vee \exists x A \Leftrightarrow \exists x(C \vee A)$.

(9) $C \to \forall x A \Leftrightarrow \forall x(C \to A)$,
$C \to \exists x A \Leftrightarrow \exists x(C \to A)$,
$\forall x A \to C \Leftrightarrow \exists x(A \to C)$,
$\exists x A \to C \Leftrightarrow \forall x(A \to C)$.

(10) $\forall x \varphi \Rightarrow \varphi_t^x$ and $\varphi_t^x \Rightarrow \exists x \varphi$ (if in the formula φ, the substitution of the free variable x with the term t is permitted).

Exercise 17 a) Prove that in (2), (5) and (6), the inverse implications are not true (by giving counterexamples).

b) Prove (9) using (8) and (7).

Exercise 18 We consider the structure $\mathcal{M} = (\mathbb{N}, S, P)$, where S and P are predicates of 3 variables defined as follows: $S(x, y, z)$ is true if and only if $x + y = z$, while $P(x, y, z)$ is true if and only if $xy = z$.

1. Write down a formula with a free variable x, true if and only if:
   (a) $x = 0$;
   (b) $x = 1$;
   (c) $x = 2$;
   (d) x is an even number;
   (e) x is an odd number;
   (f) x is a prime number.
2. Write down a formula with two free variables x, y, true if and only if:
   (a) $x = y$;
   (b) $x \leq y$;
   (c) $x < y$;
   (d) x divides y;
   (e) x and y are twin prime numbers (their difference is 2).
3. Write down a formula with free three variables x, y, z, true if and only if:
   (a) z is the least common multiple of x and y;
   (b) z is the greatest common divisor of x and y;
4. Write down the proposition (closed formula) which expresses:
   (a) the commutativity of the addition;
   (b) the associativity of the addition;
   (c) the commutativity of the multiplication;
   (d) the associativity of the multiplication;
   (e) the distributivity of the addition with respect to multiplication;
   (f) for any natural number there exists one strictly larger;
   (g) the infinity of the set of prime numbers;
   (h) the infinity of the set of pairs of twin prime numbers;
   (i) that every natural number is sum of 4 perfect squares;
   (j) the existence of the least common multiple and of the greatest common divisor;
   (k) every even number > 2 is the sum of two prime numbers.

Chapter 3

SETS

3.1 Naive and axiomatic set theory

We start with a recapitulation of the knowledge acquired in high school.

3.1.1 By a **set** we mean a well-determined collection of unique things (objects, notions), called its **elements**. The fact that the element a **belongs** to the set A is denoted $a \in A$; the notation $b \notin A$ means: b does not belong to A, that is $\neg(b \in A)$.

These three notions are primary, that is, we do not define them.

A set can be given by enumerating its elements, for example $A = \{1, 2, 3, 4\}$, $B = \{x, y, z\}$, or by a property (predicate) $P(x)$:

$$A = \{x \mid P(x)\},$$

for instance $A = \{x \mid x \in \mathbb{R} \text{ and } 0 \leq x \leq 3\}$.

The sets A and B are **equal**, $A = B$, if they have the same elements.

The empty set, denoted by $\emptyset$, is the *unique* set with no elements.

Definition 3.1.2 a) A set A is a **subset** of the set B, if every element of A is an element of B; notation: $A \subseteq B$. Any nonempty set A has two **trivial** subsets: $\emptyset$ and A.

Note that we have $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$. If $A \subseteq B$ and there exists $x \in B$ such that $x \notin A$, then we say that A is **proper subset** of B. Notation: $A \subset B$.

b) The subsets of a set A form **the power set** of A:

$$\mathcal{P}(A) = \{X \mid X \subseteq A\},$$

that is, $X \in \mathcal{P}(A) \Leftrightarrow X \subseteq A$.

Exercise 19 Prove that the sets $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}, \ldots$ are pairwise distinct. (Hint: use mathematical induction.)

Definition 3.1.3 a) The **intersection** of the sets A and B is the set of common elements, that is

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$

If $A \cap B = \emptyset$, then we say that A and B are **disjoint**.

b) The **union** of the sets A and B is the set

$$A \cup B = \{x \mid x \in A \text{ sau } x \in B\}.$$

c) The **difference** of the sets $A \setminus B$ is the set

$$A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}.$$

If $B \subseteq A$, then $A \setminus B$ is **the complement** of B with respect to A. Notation: $\complement_A(B)$.

d) **The symmetric difference** a of the sets A and B is the set

$$A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

Exercise 20 Let A, B, C be sets included in the universe U. Prove the following basic properties:
a) $A \subseteq A$ (*reflexivity*);
b) if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$ (*transitivity*);
c) if $A \subseteq B$ and $B \subseteq A$, then $A = B$ (*antisymmetry*);
d) $A \cup B = B \cup A$, $A \cap B = B \cap A$ (*commutativity*);
e) $(A \cup B) \cup C = A \cup (B \cup C))$, $(A \cap B) \cap C = A \cap (B \cap C))$ (*associativity*);
f) $A \cap A = A$, $A \cap A = A$ (*idempotence*);
g) $A \cup (A \cap B) = A$; $A \cap (B \cup A) = A$ (*absorption*);
h) $A \cup \emptyset = A$; $A \cap \emptyset = \emptyset$;
i) $A \cup \complement A = U$; $A \cap \complement A = \emptyset$;
j) $\complement\complement A = A$;

Exercise 21 Let A, B, C sets. Prove that:
a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$; $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (*distributivity*);
b) $A \setminus B = A \cap \complement B$;
c) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C) = (A \setminus B) \setminus C$;
d) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$;
e) $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$;
f) $(A \cap B) \setminus C = A \cap (B \setminus C) = (A \setminus C) \cap B$;
g) $\complement(A \cup B) = \complement A \cap \complement B$; $\complement(A \cap B) = \complement A \cup \complement B$ (*De Morgan laws*).

Exercise 22 Prove that for any sets A, B, C we have:
a) $A \triangle B = (A \cap \complement B) \cup (B \cap \complement A)$;
b) $A \triangle B = B \triangle A$;
c) $(A \triangle B) \triangle C = A \triangle (B \triangle C)$;
d) $A \triangle \emptyset = A$; $\complement A = A \triangle U$; $A \triangle A = \emptyset$;
e) $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$;
f) $A \cup B = A \triangle B \triangle (A \cap B)$.

Exercise 23 Prove that for any sets A, B, C, if $A \cap C = B \cap C$ and $A \cup C = B \cup C$ then $A = B$.

Exercise 24 Let A, B, C given sets. Find the set X which satisfies:
a) $A \cap X = B$, $A \cup X = C$;
b) $A \setminus X = B$, $X \setminus A = C$.

Definition 3.1.4 a) Let a and b be two elements. The set $\{\{a\}, \{a, b\}\}$ is denoted by (a, b), and it is called the **ordered pair** with **the first component** a and **the second component** b.
b) For $n > 2$, we define inductively $(a_1, \dots, a_n) := ((a_1, \dots, a_{n-1}), a_n)$.
c) The **cartezian product** of the sets $A_1, A_2, \dots, A_n$ (where $n \geq 1$) is the set

$$A_1 \times A_2 \times \cdots \times A_n = \{(x_1, x_2, \dots, x_n) \mid x_1 \in A_1,\ x_2 \in A_2, \dots, x_n \in A_n\}.$$

If for an index i we have $A_i = \emptyset$, then $A_1 \times A_2 \times \cdots \times A_n = \emptyset$.

Exercise 25 If a, b, c, d are elements, then $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.

Exercise 26 Let A, B, C, D be sets. Prove that:
a) $(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D)$;
b) the statement $(A \cup B) \times (C \cup D) = (A \times C) \cup (B \times D)$ is not true in general;
c) $(A \cup B) \times C = (A \times C) \cup (B \times C)$;
d) $(A \cap B) \times C = (A \times C) \cap (B \times C)$;
e) $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$.

3.1.5 The approach in this paragraph belongs to the so-called **naive set theory**, developed by the German mathematician Georg Cantor after 1870. It was later shown that this theory leads to contradictions, due to the fact that it allows the formation of “large” sets. **Bertrand Russell’s paradox** (1902) is among the best known: we consider the set R of all sets that are not contained as an element, then $R \in R$ means that $R \notin R$, and $R \notin R$ means that $R \in R$; in both cases we have a contradiction that comes from the fact that the theory allows R to be considered a set.

The **axiomatic set theory** was created to eliminate these contradictions. The most used are the axiomatizations developed by Zermelo and Fraenkel (ZF), respectively von Neumann, Bernays and Gödel (NBG).

From the point of view of predicate logic, the axioms of both theories can be given by closed formulas in the first-order language $\mathcal{L}_S$, mentioned in the previous chapter, which uses a single binary predicate $\in$ (“belongs to”). Kurt Gödel proved in 1939 that both axiomatic systems admit a model, so they are non-contradictory.

Chapter 4

RELATIONS AND FUNCTIONS

A binary relation or correspondence between the elements of the sets A and B is a set of pairs in $A \times B$. This concept formalizes and generalizes notions such as *greater than*, *equal to*, *divides*, *belongs to*, *is included in*, *parallel to*, *perpendicular to*, *congruent with*, *adjacent to* etc. The concept of function is a particular case of that of relation.

4.1 Binary relations

Definition 4.1.1 Let $n \in \mathbb{N}^*$ and let $A_1, A_2, \dots, A_n$ be sets.

a) We call **n-ary relation** the system $\rho = (A_1, A_2, \dots, A_n, R)$, where $R \subseteq A_1 \times A_2 \times \cdots \times A_n$.

If $n = 2$, then $\rho = (A_1, A_2, R)$ is a **binary relation** (or **correspondence**) between the elements of the sets A_1 and A_2, where $R \subseteq A_1 \times A_2$. In the following we deal only with binary relations, called relations for short.

b) Let $\rho = (A, B, R), R \subseteq A \times B$ a relation. The set R is called the **graph** of ρ and **we denote**:

$$(a, b) \in R \Leftrightarrow a\rho b,$$

and we read: a is in the relation ρ with b. Otherwise, $(a, b) \notin R \Leftrightarrow a \not\rho b$. We often identify the relation with its graph.

c) We say that ρ is a **hhomogeneouseous relation**, if $A = B$.

d) ρ is the **empty relation**, if $R = \emptyset$; ρ is the **universal relation**, if $R = A \times B$.

e) On the set A we define the **diagonal relation**

$$1_A = (A, A, \Delta_A), \quad \Delta_A = \{(a, a) \mid a \in A\}$$

(where $a 1_A b \Leftrightarrow a = b$).

Example 4.1.2 1) Let $A = \{a, b, c, d\}$, $B = \{1, 2\}$ and $\rho = (A, B, R)$, where $R = \{(a, 1), (a, 2), (b, 2), (c, 1)\}$. Then $a\rho 1, a\rho 2$ and $c \not\rho 2$.

2) The similarity relation on the set of triangles in the plane.

3) The divisibility relation on $\mathbb{Z}$ is the following hhomogeneouseous relation: $\rho = (\mathbb{Z}, \mathbb{Z}, R)$, where

$$R = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid a|b\} = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid \exists c \in \mathbb{Z} : \ b = ac\}.$$

4) If $A = \emptyset$ or $B = \emptyset$, then there exists a unique relation $\rho = (A, B, R)$, the empty relation, with graph $R = \emptyset$.

4.1.1 Operations with relations

Definition 4.1.3 a) We say that $\rho = (A, B, R)$ is a **subrelation** of the relation $\sigma = (A, B, S)$ (notation; $\rho \subseteq \sigma$), if $R \subseteq S$, that is, if for any $(a, b) \in A \times B$ we have $a\rho b \Rightarrow a\sigma b$.

We consider the relations $\rho = (A, B, R)$, $\rho' = (A, B, R')$ and $\sigma = (C, D, S)$.

b) The **intersection** of the relations ρ and ρ' is the relation $\rho \cap \rho' = (A, B, R \cap R')$, hence $a(\rho \cap \rho')b \Leftrightarrow a\rho b \wedge a\rho' b$.

c) The **union** of the relations ρ and ρ' is the relation $\rho \cup \rho' = (A, B, R \cup R')$, hence $a(\rho \cup \rho')b \Leftrightarrow a\rho b \vee a\rho' b$.

d) The **complement** of the relation ρ is the relation $\complement\rho = (A, B, \complement R)$, where $\complement R$ is with respect to $A \times B$. Hence $a\complement\rho b \Leftrightarrow a \not\rho b$.

e) The **inverse** of the relation ρ is the relation $\rho^{-1} = (B, A, R^{-1})$, where

$$R^{-1} = \{(b, a) \in B \times A \mid (a, b) \in R\}.$$

Hence $b\rho^{-1}a \Leftrightarrow a\rho b$.

f) The **composition** of the relations ρ and σ is the relation $\sigma \circ \rho = (A, D, S \circ R)$, where

$$S \circ R = \{(a, d) \in A \times D \mid \exists x \in B \cap C \mid (a, x) \in R,\ (x, d) \in S\},$$

that is $a(\sigma \circ \rho)b \Leftrightarrow \exists x \in B \cap C : a\rho x$ and $x\rho d$. We denote: $\rho \circ \rho = \rho^2$.

Example 4.1.4 1) On the set $\mathbb{Z}$, the relation of equality "=" is a subrelation of the relation $\leq$. The relation of divisibility $|$ is not a subrelation of $\leq$, because, for instance, $2|-6$ and $2 \not\leq -6$.

2) On $\mathbb{R}$, the intersection of $\leq$ and $\geq$ is the relation of equality $=$; the union of the relations "=" and "<" is the relation "$\leq$"; the complement of $<$ is $\geq$; the inverse of $<$ is the relation $>$.

3) Composition of relations it is not commutative, that is, in general, $\sigma \circ \rho \neq \rho \circ \sigma$. Indeed, consider the relations "<" respectively ">" on $\mathbb{N}$. Then $a(< \circ >)b \Leftrightarrow \exists c \in \mathbb{N} : a > c$ and $c < b \Leftrightarrow a, b \in \mathbb{N}^* \times \mathbb{N}^*$, that is, the graph of $< \circ >$ is the set $\mathbb{N}^* \times \mathbb{N}^*$; on the other hand, $a(> \circ <)b \Leftrightarrow \exists c \in \mathbb{N} : a < c$ and $c > b \Leftrightarrow a, b \in \mathbb{N} \times \mathbb{N}$, that is, $> \circ <$ have the graph $\mathbb{N} \times \mathbb{N}$.

Theorem 4.1.5 *Let* $\rho = (A, B, R)$*,* $\sigma = (C, D, S)$ *and* $\tau = (E, F, T)$ *be relations. Then:*

1) $(\tau \circ \sigma) \circ \rho = \tau \circ (\sigma \circ \rho)$ (composition of relations is associative)*,*

2) $\rho \circ 1_A = 1_B \circ \rho = \rho$ (the relation of equality is neutral element with respect to composition)*.*

Proof. 1) Aratam associativity of composition. We have $\tau \circ \sigma = (C, F, T \circ S)$, $(\tau \circ \sigma) \circ \rho = (A, F, (T \circ S) \circ R)$, $\sigma \circ \rho = (A, D, S \circ R)$ and $\tau \circ (\sigma \circ \rho) = (A, F, T \circ (S \circ R)$. Furthermore, for any $(a, f) \in A \times F$ we have:

$$\begin{aligned}
(a, f) \in (T \circ S) \circ R &\Leftrightarrow a(\tau \circ \sigma) \circ \rho f && \text{(notation)}\\
&\Leftrightarrow (\exists x)\ x \in B \cap C \text{ and } (a\rho x \text{ and } x(\tau \circ \sigma)f) && \text{(Definition 4.1.3. f))}\\
&\Leftrightarrow (\exists x)\ x \in B \cap C \text{ and } (a\rho x \text{ and } (\exists y)\ y \in E \cap D \text{ and } (x\sigma y \text{ and } y\tau f)) && \text{(Definition 4.1.3. f))}\\
&\Leftrightarrow (\exists x)\ (\exists y)\ x \in B \cap C \text{ and } y \in E \cap D \text{ and } (a\rho x \text{ and } x\sigma y \text{ and } y\tau f) && \text{(tautology 2.3.9 (8))}\\
&\Leftrightarrow (\exists y)\ y \in E \cap D \text{ and } (\ (\exists x)\ x \in B \cap C \text{ and } a\rho x \text{ and } x\sigma y) \text{ and } y\tau f && \text{(tautologies 2.3.9 (1), (8))}\\
&\Leftrightarrow (\exists y)\ y \in E \cap D \text{ and } (a(\sigma \circ \rho)y \text{ and } y\tau f) && \text{(Definition 4.1.3. f))}\\
&\Leftrightarrow a\tau \circ (\sigma \circ \rho)f && \text{(Definition 4.1.3. f))}\\
&\Leftrightarrow (a, f) \in T \circ (S \circ R) && \text{(notation).}
\end{aligned}$$

We have shown in this way that $(T \circ S) \circ R = T \circ (S \circ R)$. ■

Exercise 27 Consider the sets $A = \{1, 2\}$, $B = \{1, 2, 3\}$, $C = \{1, 2, 3, 4\}$, $R_1 = \{(1, 2), (1, 3), (2, 3)\} \subseteq A \times B$, $R_2 = \{(1, 4), (3, 1), (3, 4)\} \subseteq B \times C$, $\rho_1 = (A, B, R_1)$, $\rho_2 = (B, C, R_2)$. Find the relations: $\rho_2 \circ \rho_1$, $\rho_1 \circ \rho_2$, ρ_1^{-1}, ρ_1^{-1}, $(\rho_1 \circ \rho_2)^{-1}$, $\rho_2^{-1} \circ \rho_1^{-1}$.

Exercise 28 Let $\rho = (\mathbb{N}, \mathbb{N}, <)$. Find the relations $<^2$, $<^3$, $< \circ >$ and $> \circ <$.

Exercise 29 Let $A = \{1, 2, 3, 4\}$ and R, S, $S' \subseteq A \times A$, where $R = \{(1, 2),\ (1, 4), (2, 3), (4, 1), (4, 3)\}$, $S = \{(1, 1), (2, 4), (3, 4)\}$, $S' = \{(1, 4), (4, 4)\}$.

Find the relations $(S \cap S') \circ R$, $(S \circ R) \cap (S' \circ R)$, $R \circ (S \cap S')$ and $(R \circ S) \cap (R \circ S')$.

Exercise 30 We consider the relations $\rho = (A, B, R)$, $\rho' = (A, B, R')$, $\sigma = (C, D, S)$ and $\sigma' = (C, D, S')$. Prove (specifying the tautologies in the list 2.3.9 that were used):

a) $(\rho^{-1})^{-1} = \rho$; $(\complement\rho)^{-1} = \complement\rho^{-1}$;

b) $(\sigma \circ \rho)^{-1} = \rho^{-1} \circ \sigma^{-1}$;

c) $(\rho \cap \rho')^{-1} = \rho^{-1} \cap \rho'^{-1}$; $(\rho \cup \rho')^{-1} = \rho^{-1} \cup \rho'^{-1}$;

d) $\sigma \circ (\rho \cup \rho') = (\sigma \circ \rho) \cup (\sigma \circ \rho')$; $(\sigma \cup \sigma') \circ \rho = (\sigma \circ \rho) \cup (\sigma' \circ \rho)$;

e) $\sigma \circ (\rho \cap \rho') \subseteq (\sigma \circ \rho) \cap (\sigma \circ \rho')$; $(\sigma \cap \sigma') \circ \rho \subseteq (\sigma \circ \rho) \cap (\sigma' \circ \rho)$;

f) if $\sigma \subseteq \sigma'$, $\rho \subseteq \rho'$ then $\sigma \circ \rho \subseteq \sigma' \circ \rho'$.

4.1.2 The section of a relation with respect to a subset

Definition 4.1.6 Let $\rho = (A, B, R)$ be a relation, and let $X \subseteq A$. The set

$$\rho(X) = \{b \in B \mid \exists x \in X \mid x\rho b\} \subseteq B$$

is called **the section of the relation** ρ **with respect to the subset** X. If the subset $X = \{x\}$ has a single element, then we denote:

$$\rho\langle x\rangle = \rho(\{x\}) = \{b \in B \mid x\rho b\}.$$

Example 4.1.7 In Example 4.1.4 1) we have $\rho(\{a,b\}) = \{1,2\}$, $\rho(\{c,d\}) = \{1\}$, $\rho\langle a\rangle = \{1,2\}$, $\rho\langle d\rangle = \emptyset$, $\rho(A) = \{1,2\}$, $\rho^{-1}(B) = \{a,b,c\}$.

Theorem 4.1.8 *Let* $\rho = (A,B,R)$ *and* $\sigma = (C,D,S)$ *be relations, and let* $X \subseteq A$. *Then we have:*

$$(\sigma \circ \rho)(X) = \sigma(\rho(X) \cap C);$$

if in addition $B = C$*, then* $(\sigma \circ \rho)(X) = \sigma(\rho(X))$.

Proof. For any $d \in D$ we have:

$$\begin{aligned}
d \in (\sigma\circ\rho)(X) &\Leftrightarrow (\exists x)\ x \in X \text{ and } x(\sigma\circ\rho)d \Leftrightarrow && \text{(Definition 4.1.6)}\\
&\Leftrightarrow (\exists x)\ x \in X \text{ and } ((\exists z)\ z \in B \cap C \text{ and } x\rho z \text{ and } z\sigma d) \Leftrightarrow && \text{(Definition 4.1.3. f))}\\
&\Leftrightarrow (\exists x)\ (\exists z)\ (x \in X \text{ and } z \in B \cap C \text{ and } x\rho z \text{ and } z\sigma d \Leftrightarrow && \text{(tautology 2.3.9 (8))}\\
&\Leftrightarrow (\exists z)\ z \in B \cap C \text{ and } ((\exists x)\ x \in X \text{ and } x\rho z) \text{ and } z\sigma d \Leftrightarrow && \text{(tautologies 2.3.9 (1), (8))}\\
&\Leftrightarrow (\exists z)\ z \in B \cap C \text{ and } (z \in \rho(X) \text{ and } z\sigma d) \Leftrightarrow && \text{(Definition 4.1.6)}\\
&\Leftrightarrow (\exists z)\ z \in \rho(X) \cap C \text{ and } z\sigma y \Leftrightarrow && \text{(because } \rho(X) \subseteq B \text{)}\\
&\Leftrightarrow d \in \sigma(\rho(X) \cap C)), && \text{(Definition 4.1.6)}
\end{aligned}$$

hence the statement e proved. ■

Exercise 31 Let $A = \{a_1, a_2, a_3, a_4\}$, $B = \{b_1, b_2, b_3, b_4, b_5\}$, $X = \{a_2, a_4\}$, $Y = \{b_1, b_2, b_4, b_5\}$ and consider the relation $R = \{(a_1, b_2), (a_3, b_5), (a_1, b_3), (a_2, b_4)\} \subseteq A \times B$. Find the sets $R(X)$, $R\langle a_2\rangle$, $R^{-1}(Y)$, $R^{-1}\langle b_5\rangle$, $R^{-1}(B)$ and $R(A)$.

Exercise 32 Let $\delta = (\mathbb{N}, \mathbb{N}, |)$ be the relation of divisibility. Find the sets $\delta\langle 1\rangle$, $\delta^{-1}(\{4,9\})$, $\delta^{-1}(\mathbb{N})$ and $\delta(\mathbb{N})$.

Exercise 33 Let $\rho = (A,B,R)$ and $\rho' = (A,B,R')$ be relations. Prove that the following statements are equivalent:
(i) $\rho \subseteq \rho'$;
(ii) $(\forall\ x \in A)\ \rho\langle x\rangle \subseteq \rho'\langle x\rangle$;
(iii) $(\forall\ X \subseteq A)\ \rho(X) \subseteq \rho'(X)$.

Exercise 34 Let $\rho = (A,B,R)$ and $\rho' = (A,B,R')$ relations and let $X, X' \subseteq A$. Prove (specifying the tautologies in the list 2.3.9 that were used):
a) if $X \subseteq X'$ and $\rho \subseteq \rho'$, then $\rho(X) \subseteq \rho'(X')$;
b) $\rho(X \cup X') = \rho(X) \cup \rho(X')$; $(\rho \cup \rho')(X) = \rho(X) \cup \rho'(X)$;
c) $\rho(X \cap X') \subseteq \rho(X) \cap \rho(X')$; $(\rho \cap \rho')(X) \subseteq \rho(X) \cap \rho'(X)$;

Remark 4.1.9 In c) the equality does not hold in general. Let, for instance, $\rho = (A,A,R)$, where $A = \{1,2,3\}, R = \{(1,1), (1,3), (2,2), (3,1)(3,3)\}$, and let $X = \{1,2\}, X' = \{2,3\}$. Then $\rho(X) \cap \rho(X') = \{1,2,3\}$ and $\rho(X \cap X') = \rho\langle 2\rangle = \{2\}$.

4.2 Functions

Definition 4.2.1 a) The relation $f = (A,B,F)$, where $F \subseteq A \times B$, is called a **function (functional relation)**, if for any $a \in A$, the section $f\langle a\rangle$ has exactly one element.

b) If $f = (A,B,F)$ is a function, then A is called the **domain** of f, notation $A = \operatorname{dom} f$.

c) The set B is the **codomain** lui f, notation $B = \operatorname{codom} f$, while the section $f(A)$ is the **image** lui f; notation: $f(A) = \operatorname{Im} f$.

d) The set $F \subseteq A \times B$ is **the graph** of the function f.

If $f = (A,B,F)$ is a function, then we use the following notation:

$$f : A \to B, \quad A \xrightarrow{f} B.$$

If $a \in A$, then the element $b \in B$ determined by the equality $f\langle a\rangle = \{b\}$ is denoted $b = f(a)$ or $a \mapsto b = f(a)$.

Remark 4.2.2 a) The functions $f : A \to B$ and $f' : A' \to B'$ are equal $(f = f')$ if and only if $A = A', B = B'$ and $f(a) = f(a')$ for any $a \in A$.

b) If $A = \emptyset$, then the unique relation $\rho = (A,B,R)$ is the empty relation $(R = \emptyset)$; this is a function for any set B.

If $A \neq \emptyset$ and $B = \emptyset$, then the empty relation $\rho = (A, \emptyset, \emptyset)$ is not a function.

c) If $f : A \to B$ is a function, and $X \subseteq A$, $Y \subseteq B$, $y \in Y$, then

$$f(X) = \{b \in B \mid \exists x \in X : f(x) = b\} = \{f(x) \mid x \in X\},$$
$$f^{-1}(Y) = \{a \in A \mid \exists y \in Y : af^{-1}y\} = \{a \in A \mid \exists y \in Y : f(a) = y\} = \{a \in A \mid f(a) \in Y\}$$

and

$$f^{-1}\langle y\rangle = f^{-1}(y) = \{a \in A \mid f(a) = y\},$$

and the graph is $F = \{(a, f(a)) \mid a \in A\}$.

Example 4.2.3 1) In Example 4.1.1. 1), the relation ρ it is not a function, because, for instance, $\rho\langle a\rangle = \{1, 2\}$. The relation $\rho' = (A, B, R')$, where $A = \{a, b, c, d\}$, $B = \{1, 2\}$, and $R' = \{(a, 1), (b, 1), (c, 2), (d, 2)\}$ is a function.

Theorem 4.2.4 *1) Let* $f = (A, B, F)$ *and* $g = (C, D, G)$ *functions. The composed relation* $g \circ f = (A, D, G \circ F)$ *is a function if and only if* $f(A) \subseteq C$*, that is* $\operatorname{Im} f \subseteq \operatorname{Dom} g$*, and in this case,* $(g \circ f)(a) = g(f(a))$ *for any* $a \in A$*.*

2) If $f : A \to B, g : B \to C, h : C \to D$ *are functions, then* $f \circ 1_A = 1_B \circ f = f$ *and* $(h \circ g) \circ f = h \circ (g \circ f)$*.*

Proof. 1) "$\Rightarrow$" We assume that $g \circ f$ is a function, and let $b \in f(A)$. We only have to show that $b \in C$. Indeed, because $b \in f(A)$, there exists $a \in A$ such that $b = f(a)$. Let $d = (g \circ f)(a)$ (where $g \circ f$ is a function), that is, $a(g \circ f)d$, hence there exists $c \in B \cap C$ such that afc and cgd. From this, afc and afb; because f is a function, we get $b = c \in B \cap C$.

"$\Leftarrow$" We now assume that $f(A) \subseteq C$, and let $a \in A$. Since f is a function, there exists $b \in f(A)$ such that $f(a) = b$ (that is, afb). Here $b \in f(A) \subseteq C$, and because g is function, there exists $d \in D$ such that $g(b) = d$ (that is, $b\,g\,d$). From this, $a\,(g \circ f)\,d$ and we have

$$(g \circ f)\langle a\rangle = g\langle f\langle a\rangle\rangle = g\langle f(a)\rangle = \{g(f(a))\},$$

hence $g \circ f$ is a function, and $(g \circ f)(a) = d = g(b) = g(f(a))$.

2) Follows from the property concerning relations, or may be easily proved directly. ■

Exercise 35 Let $\rho = (A, B, R)$ be a relation. Prove that ρ is a function if and only if

$$1_A \subseteq \rho^{-1} \circ \rho \quad \text{and} \quad \rho \circ \rho^{-1} \subseteq 1_B.$$

Exercise 36 Let $f : A \to B$ be a function. Prove that:

a) For any $X \subseteq A$ we have $X \subseteq f^{-1}(f(X))$;

b) For any $Y \subseteq B$ we have $Y \supseteq f(f^{-1}(Y))$;

c) $f \circ f^{-1} \circ f = f$.

4.2.1 Commutative diagrams

We consider the functions $f : A \to B$, $g : B \to C$ and $h : A \to C$, represented by the following diagram:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "g"'] & B \\
C \arrow[ur, "h"'] &
\end{tikzcd}$$

We say that this is a **commutative diagram** if $f = h \circ g$. We also have other situations, for instance:

$$\begin{tikzcd}
A \arrow[r, "k"] \arrow[d, "f"'] & D \arrow[d, "h"] \\
B \arrow[r, "g"] & C
\end{tikzcd}
\qquad
\begin{tikzcd}
A \arrow[r, "k"] \arrow[d, "f"'] & D \\
B \arrow[r, "g"] & C \arrow[u, "h"']
\end{tikzcd}$$

These are **commutative diagrams** if $h \circ k = g \circ f$, respectively $h \circ g \circ f = k$.

4.2.2 Family of elements and family of sets

Definition 4.2.5 a) Let $f : I \to A$ be a function, and let $F = \{(i, f(i)) \mid i \in I\}$ be the graph of f. We often identify the function f with F, and we denote $(a_i)_{i\in I}$, where $a_i = f(i)$; we say that $(a_i)_{i\in I}$ is a **family of elements**, while I is **the set of indices**.

Analogously, if $f : I \to \mathcal{P}(U)$ a function, then we say that $(A_i)_{i\in I}$ is a **family of sets**, where $A_i = f(i) \subseteq U$.

b) **The union of the family of sets** $(A_i)_{i\in I}$ is the set

$$\bigcup_{i\in I} A_i = \{a \in U \mid \exists i \in I : a \in A_i\}.$$

c) **The intersection of the family of sets** $(A_i)_{i\in I}$ is the set

$$\bigcap_{i\in I} A_i = \{a \in U \mid \forall i \in I : a \in A_i\}.$$

Remark 4.2.6 Assume that if $I = \emptyset$. Then $\bigcup_{i\in I} A_i = \emptyset$, because then for any $a \in A$ it is not true that $\exists i \in I :\ a \in A_i$. On the other hand, $\bigcap_{i\in I} A_i = A$, because the statement $\exists i \in I :\ a \notin A_i$ is false for any $a \in A$, hence its negation $\forall i \in I : a \in A_i$ is true for any $a \in A$.

Exercise 37 Prove the following identities (sspecifying the tautologies in the list 2.3.9 that were used), where $A_{ij}, A_i, B_j, A \in \mathcal{P}(U)$ for any $i \in I$, $j \in J$:

a) $\bigcup_{i\in I}\bigcup_{j\in J} A_{ij} = \bigcup_{j\in J}\bigcup_{i\in I} A_{ij}$;

b) $\bigcap_{i\in I}\bigcap_{j\in J} A_{ij} = \bigcap_{j\in J}\bigcap_{i\in I} A_{ij}$;

c) $\complement(\bigcup_{i\in I} A_i) = \bigcap_{i\in I} \complement(A_i)$;

d) $\complement(\bigcap_{i\in I} A_i) = \bigcup_{i\in I} \complement(A_i)$;

e) $(\bigcup_{i\in I} A_i) \cup (\bigcup_{i\in I} B_i) = \bigcup_{i\in I}(A_i \cup B_i)$;

f) $\bigcup_{j\in J}(A \cap B_j) = A \cap (\bigcup_{j\in J} B_j)$;

g) $\bigcap_{j\in J}(A \cup B_j) = A \cup (\bigcap_{j\in J} B_j)$;

h) $\bigcup_{i\in I}(\bigcap_{j\in J} A_{ij}) \subseteq \bigcap_{j\in J}(\bigcup_{i\in I} A_{ij})$;

i) $(\bigcup_{i\in I} A_i) \cap (\bigcup_{j\in J} B_j) = \bigcup_{(i,j)\in I\times J}(A_i \cap B_j)$;

j) $(\bigcap_{i\in I} A_i) \cup (\bigcap_{j\in J} B_j) = \bigcap_{(i,j)\in I\times J}(A_i \cup B_j)$.

Exercise 38 Prove that (specifying the tautologies in the list 2.3.9 that were used):

a) $(\bigcap_{i\in I} X_i) \times (\bigcap_{i\in I} Y_i) = \bigcap_{i\in I}(X_i \times Y_i)$;

b) $(\bigcup_{i\in I} X_i) \times (\bigcup_{j\in J} Y_j) = \bigcup_{(i,j)\in I\times J}(X_i \times Y_j)$.

Exercise 39 Let $f : A \to B$ a function and $X_i \subseteq A$, $Y_i \subseteq B$ $\forall$ $i \in I$. Prove that (specifying the tautologies in the list 2.3.9 that were used):

a) $f(\bigcup_{i\in I} X_i) = \bigcup_{i\in I} f(X_i)$;

b) $f(\bigcap_{i\in I} X_i) \subseteq \bigcap_{i\in I} f(X_i)$. Give an example in which the inclusion is strict;

c) $f^{-1}(\bigcup_{i\in I} Y_i) = \bigcup_{i\in I} f^{-1}(Y_i)$;

d) $f^{-1}(\bigcap_{i\in I} Y_i) = \bigcap_{i\in I} f^{-1}(Y_i)$.

Exercise 40 Prove that $\mathcal{P}(\bigcap_{i\in I} A_i) = \bigcap_{i\in I} \mathcal{P}(A_i)$.

Exercise 41 Consider the relations $\rho_i = (A, B, R_i)$, $i \in I$ and $\sigma = (C, D, S)$. Prove that (specifying the tautologies in the list 2.3.9 that were used):

a) $\sigma \circ (\bigcup_{i\in I} \rho_i) = \bigcup_{i\in I}(\sigma \circ \rho_i)$;

b) $(\bigcup_{i\in I} \rho_i) \circ \sigma = \bigcup_{i\in I}(\rho_i \circ \sigma)$;

c) $\sigma \circ (\bigcap_{i\in I} \rho_i) \subseteq \bigcap_{i\in I}(\sigma \circ \rho_i)$;

d) $(\bigcap_{i\in I} \rho_i) \circ \sigma \subseteq \bigcap_{i\in I}(\rho_i \circ \sigma)$.

4.3 Injective, surjective and bijective functions

Definition 4.3.1 Let $f : A \to B$ be a function. We say that

a) f is **injective**, if $\forall x_1, x_2 \in A, x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$, or equivalently, $\forall x_1, x_2 \in A, f(x_1) = f(x_2) \Rightarrow x_1 = x_2$;

b) f is **surjective**, if $\forall y \in B\ \exists x \in A : f(x) = y$, or equivalently, $f(A) = B$;

c) f is **bijective**, if is injective and surjective, or equivalently, if $\forall y \in B$, there exists a unique $x \in A$ such that $f(x) = y$.

Example 4.3.2 1) The function $f : \mathbb{R} \to \mathbb{R}, \quad f(x) = x^2$ is not injective, because for instance $-1 \neq 1$ and $f(-1) = f(1) = 1$; it is not surjective, because for instance $y = -1 \in \mathbb{R}$, but it does not exist $x \in \mathbb{R}$ such that $f(x) = x^2 = -1$.

The function $g : [0, \infty) \to \mathbb{R}, \ g(x) = x^2$ is injective and it is not surjective, the function $h : [0, \infty) \to [0, \infty), \quad h(x) = x^2$ is injective and surjective, hence bijective.

2) For any set A, the diagonal relation $1_A = (A, A, \Delta_A)$ is a bijective function.

Theorem 4.3.3 (the characterization of injective functions) *Let $f : A \to B$ be a function. The following statements are equivalent:*

(i) f *is injective;*

(ii) *for any set A' and for any functions $\alpha, \beta : A' \to A$, if $f \circ \alpha = f \circ \beta$, then $\alpha = \beta$* (we say that f is **left cancellable**)*;*

(iii) (*assuming that $A \neq \emptyset$*) f *has a* **left inverse (retraction)***, that is there exists a function $r : B \to A$ such that $r \circ f = 1_A$.*

Proof. (i) $\Rightarrow$ (ii) If $f \circ \alpha = f \circ \beta$, then for any $a \in A'$ we have $f(\alpha(a)) = f(\beta(a))$; from the injectivity of f it follows that $\alpha(a) = \beta(a)$; consequently, $\alpha = \beta$.

(ii) $\Rightarrow$ (i) We assume that the statement (ii) is true, and that f it is not injective, that is, there exist $x_1, x_2 \in A$, such that $x_1 \neq x_2$ and $f(x_1) = f(x_2)$.

Let $A' = \{x_1, x_2\}$, and let $\alpha, \beta : A' \to A$, $\alpha(x_1) = x_1$, $\alpha(x_2) = x_2$, $\beta(x_1) = x_1$, $\beta(x_2) = x_1$. Then $\alpha \neq \beta$, but $f \circ \alpha = f \circ \beta$, because

$$(f \circ \alpha)(x_1) = f(\alpha(x_1)) = f(x_1) = f(\beta(x_1)) = (f \circ \beta)(x_1),$$
$$(f \circ \alpha)(x_2) = f(\alpha(x_2)) = f(x_2) = f(x_1) = f(\beta(x_2)) = (f \circ \beta)(x_2),$$

hence we have a contradiction.

(i) $\Rightarrow$ (iii) We assume that f injective, and let $a_0 \in A$. We consider the function

$$r : B \to A, \qquad r(b) = \begin{cases} a, \ \text{if } b = f(a) \in f(A) \\ a_0, \ \text{if } b \in B \setminus f(A) \end{cases},$$

which is well defined, because by the injectivity of f, for any $b \in f(A)$, there exists a unique $a \in A$ such that $f(a) = b$. Hence $(r \circ f)(a) = r(f(a)) = r(b) = a = 1_A(a)$ for any $a \in A$, that is, $r \circ f = 1_A$.

(iii) $\Rightarrow$ (i) Assume that there exists a function $r : B \to A$ such that $r \circ f = 1_A$. If $x_1, x_2 \in A$ and $f(x_1) = f(x_2)$, then $r(f(x_1)) = r(f(x_2))$, that is, $(r \circ f)(x_1)) = (r \circ f)(x_2))$, hence $x_1 = x_2$. Consequently, f is injective.

Theorem 4.3.4 (the characterization of surjective functions) *Let $f : A \to B$ be a function. The following statements are equivalent:*

(i) f *is surjective;*

(ii) *for any set B' and for any functions $\alpha, \beta : B \to B'$, if $\alpha \circ f = \beta \circ f$, then $\alpha = \beta$* (we say that f is **right cancellable**)*;*

(iii) f *has a* **right inverse (section)***, that is, there exists a function $s : B \to A$ such that $f \circ s = 1_B$.*

Proof. (i) $\Rightarrow$ (ii) Assume that f is surjective, and that $\alpha \circ f = \beta \circ f$, that is $\alpha(f(a)) = \beta(f(a))$ for any $a \in A$. By the surjectivity of f we have that for any $b \in B$, there exists $a \in A$ such that $b = f(a)$; we get $\alpha(b) = \beta(b)$, that is $\alpha = \beta$.

(ii) $\Rightarrow$ (i) We assume that the statement (ii) is true, but f it is not surjective, that is there exists $b_0 \in B \setminus f(A)$. In the case $A \neq \emptyset$, let $B' = B$, and consider the functions α, $\beta : B \to B$, where $\alpha = 1_B$ and

$$\beta(b) = \begin{cases} b, & \text{if } b \neq b_0, \\ b_0', & \text{if } b = b_0, \end{cases}$$

where $b_0' \in f(A)$. Then $\alpha \neq \beta$, because $\beta(b_0) = b_0' \neq b_0$ (since $b_0 \notin f(A)$, $b_0' \in f(A)$), but $\alpha \circ f = \beta \circ f$, because $(\alpha \circ f)(a) = \alpha(f(a)) = f(a) = \beta(f(a)) = (\beta \circ f)(a)$ for any $a \in A$, which is a contradiction.

In the case $A = \emptyset$, let $B' = \{0, 1\}$, $\alpha, \beta : B \to B'$, $\alpha(b) = 0$, $\beta(b) = 1$ for any $b \in B$. Then $\alpha \neq \beta$ and $\alpha \circ f = \beta \circ f = \emptyset$.

(i) $\Rightarrow$ (iii) We assume that the function f is surjective. Then for any $b \in B$, $f^{-1}(b) = \{a \in A \mid f(a) = b\} \neq \emptyset$. For any b we choose an element $a \in f^{-1}(b)$; thus we get the function $s : B \to A$, $s(b) = a$, and we have

$$(f \circ s)(b) = f(s(b)) = f(a) = b = 1_B(b),$$

that is $f \circ s = 1_B$.

(iii) $\Rightarrow$ (i) Let $s : B \to A$ a function such that $f \circ s = 1_B$. Then for any $b \in B$ we have $b = 1_B(b) = f(s(b))$; denoting $a = s(b) \in A$, we have $f(a) = b$; consequently, f is surjective.

Theorem 4.3.5 (the characterization of bijective functions) *Let* $f : A \to B$ *be a function. The following statements are equivalent:*

(i) f *is bijective;*

(ii) *The inverse relation* f^{-1} *is a function; in this case we have* $f^{-1} \circ f = 1_A, \quad f \circ f^{-1} = 1_B$*;*

(iii) f *is* **invertible**, *that is there exists a function* $g : B \to A$*, such that*

$$g \circ f = 1_A, \qquad f \circ g = 1_B.$$

Proof. (i) $\Leftrightarrow$ (ii) f is bijective $\Leftrightarrow$ for any $b \in B$, the set $f^{-1}(b) = \{a \in A \mid f(a) = b\}$ has exactly one element $\Leftrightarrow$ the relation f^{-1} is a function.

Furthermore, for all $a, a' \in A$ we have $a(f^{-1} \circ f)a' \Leftrightarrow \exists b \in B : afb$ and $bf^{-1}a' \Leftrightarrow \exists b \in B : f(a) = b$ and $f(a') = b \Leftrightarrow a = a'$ (because f is injective) $\Leftrightarrow a1_Aa'$, that is $f^{-1} \circ f = 1_A$.

Similarly, for all $b, b' \in B$ we have $b(f \circ f^{-1})b' \Leftrightarrow \exists a \in A : bf^{-1}a$ and $afb' \Leftrightarrow \exists a \in A : f(a) = b$ and $f(a) = b' \Leftrightarrow b = b'$ (because f is surjective) $\Leftrightarrow b1_Bb'$, that is $f \circ f^{-1} = 1_B$.

(i) $\Rightarrow$ (iii) If f is bijective, then let $g = f^{-1}$, about which we have just shown that it satisfies condition (iii).

(iii) $\Rightarrow$ (i) Follows from the implications (iii) $\Rightarrow$ (i) of the previous two theorems.

Remark 4.3.6 If f is a bijective function, then the function f^{-1} is also bijective, since $(f^{-1})^{-1} = f$.

Exercise 42 Let $f : A \to B$ and $g : B \to C$ be functions. Prove that:

a) If f and g is injective (surjective), then $g \circ f$ is injective (surjective);

b) If $g \circ f$ is injective (surjective), then f is injectiv (g is surjective);

c) If $g \circ f$ is injective and f is surjective, then g is injective;

d) If $g \circ f$ is surjective and g is injective, then f is surjective.

Exercise 43 Let $f : A \to B$ be a function, $X_1, X_2 \subseteq A$, $(X_i)_{i \in I}, X_i \subseteq A$, and $Y_1, Y_2 \subseteq B$. Prove that:

a) $f^{-1}(Y_1 \setminus Y_2) = f^{-1}(Y_1) \setminus f^{-1}(Y_2)$;

b) if f is injective, then

(1) $f(X_1 \setminus X_2) = f(X_1) \setminus f(X_2)$,

(2) $f(\bigcap_{i \in I} X_i) = \bigcap_{i \in I} f(X_i)$.

Exercise 44 Let $f : A \to B$ be a function.

a) Prove that the following statements are equivalent:

(i) f is injective;

(ii) $f^{-1} \circ f = 1_A$;

(iii) for all $X \subseteq A$ we have $f^{-1}(f(X)) = X$;

(iv) for all $X \subseteq A$ we have $f(\complement(X)) \subseteq \complement f(X)$;

(v) for all $X_1, X_2 \subseteq A$ we have $f(X_1 \cap X_2) = f(X_1) \cap f(X_2)$.

b) Prove that the following statements are equivalent:

(i) f is surjective;

(ii) $f \circ f^{-1} = 1_B$;

(iii) for all $Y \subseteq B$ we have $f(f^{-1}(Y)) = Y$;

(iv) for all $X \subseteq A$ we have $\complement f(X) \subseteq f(\complement(X))$.

Exercise 45 Let $f : A \to B$ be a function.

a) Assume that f is surjective. Prove that f is injective $\Leftrightarrow$ f has exactly one right inverse.

b) Assume that $A \neq \emptyset$ and that f is injective. If f is surjective, then show that f has exactly one left inverse; show also that the converse statement is not true.

Exercise 46 Let $A \neq \emptyset$ and let $f : A \to B$ be a function. Prove that there exists a function $g : B \to A$ such that $f \circ g \circ f = f$.

4.3.1 The direct product of a family of sets and of a family of functions

Let $(A_i)_{i\in I}$ a family of sets. By definition,

$$\prod_{i\in I} A_i = \{f : I \to \bigcup_{i\in I} A_i \mid \forall\, i \in I:\ f(i) \in A_i\} =$$
$$= \{(a_i)_{i\in I} \mid \forall\, i \in I:\ a_i \in A_i\}$$

is **the generalized cartesian product** of the family $(A_i)_{i\in I}$. The function

$$p_j : \prod_{i\in I} A_i \to A_j, \quad p_j((a_i)_{i\in I}) = a_j$$

is called a **canonical projection**, and the pair $(\prod_{i\in I} A_i, (p_i)_{i\in I})$ is the **direct product** of the family $(A_i)_{i\in I}$.

Furthermore, if $(f_i : A_i \to A'_i)_{i\in I}$ is a family of functions, then

$$\prod_{i\in I} f_i : \prod_{i\in I} A_i \to \prod_{i\in I} A'_i, \quad (\prod_{i\in I} f_i)((a_i)_{i\in I}) = (f_i(a_i))_{i\in I}$$

is the **direct product** of the family $(f_i)_{i\in I}$.

Note that $\prod_{i\in I} A_i$ is nonempty if and only if $I \neq \emptyset$ and $A_i \neq \emptyset$ for any $i \in I$. If $I = \{1\}$, then $\prod_{i\in I} A_i = A_1$; if $I = \{1, 2\}$, then $\prod_{i\in I} A_i$ is identified with the cartesian product $A_1 \times A_2$. In this case, if $f_i : A_i \to A'_i$, $i = 1, 2$, then

$$f_1 \times f_2 : A_1 \times A_2 \to A'_1 \times A'_2, \qquad (f_1 \times f_2)(a_1, a_2) = (f_1(a_1), f_2(a_2)).$$

Exercise 47 Let $f : A \to A'$, $g : B \to B'$ be functions, and let $(f_i : A_i \to A'_i)_{i\in I}$ be a family of functions. Prove that:

a) f and g are injective (surjective) $\Leftrightarrow$ $f \times g$ is injective (surjective);

b) If f_i are injective (respectively surjective) for any $i \in I$, then $\prod_{i\in I} f_i$ injective (respectively surjective).

4.3.2 The direct sum of a family of sets and of a family of functions

Let $(A_i)_{i\in I}$ be a family of sets. By definition,

$$\coprod_{i\in I} A_i = \bigcup_{i\in I} A_i \times \{i\} = \{(a_i, i) \mid\ i \in I,\ a_i \in A_i\}$$

is the **disjoint union**of the family $(A_i)_{i\in I}$. The function

$$q_j : A_j \to \coprod_{i\in I} A_i, \qquad q_j(a_j) = (a_j, j)$$

is called a **canonical injection**, while the pair $(\coprod_{i\in I} A_i, (q_i)_{i\in I})$ is the **direct sum** of the family $(A_i)_{i\in I}$.

Furthermore, if $(f_i : A_i \to A'_i)_{i\in I}$ is a family of functions, then

$$\coprod_{i\in I} f_i : \coprod_{i\in I} A_i \to \coprod_{i\in I} A'_i, \qquad (\coprod_{i\in I} f_i)(a_i, i) = (f_i(a_i), i)$$

is the **direct sum** of the family $(f_i)_{i\in I}$.

Note that $\coprod_{i\in I} A_i$ is the empty set if and only if $I = \emptyset$ or $A_i = \emptyset$ for any $i \in I$.

Exercise 48 Let $f : A \to A'$, $g : B \to B'$ be functions and let $(f_i : A_i \to A'_i)_{i\in I}$ be a family of functions. Prove that:

a) If f and g are injective (surjective), then $f \amalg g$ is injective (surjective);

b) If f_i injective (respectively surjective) for any $i \in I$, then $\coprod_{i\in I} f_i$ is injective (respectively surjective).

4.3.3 The set $\mathrm{Hom}(A, B)$ and the function $\mathrm{Hom}(f, g)$

Let A and B be sets. We denote by $\mathrm{Hom}(A, B)$ the set of all functions $f : A \to B$:

$$\mathrm{Hom}(A, B) = \{f \mid f : A \to B\}.$$

Note that if $A = \emptyset$, then $\mathrm{Hom}(\emptyset, B) = \{\emptyset\}$, while if $A \neq \emptyset$, $B = \emptyset$, then $\mathrm{Hom}(A, \emptyset) = \emptyset$.

Let $f : A' \to A$ and $g : B \to B'$ be functions; we define the following function:

$$\mathrm{Hom}(f,g) : \mathrm{Hom}(A,B) \to \mathrm{Hom}(A',B'), \qquad \mathrm{Hom}(f,g)(\alpha) = g \circ \alpha \circ f,$$

so the following diagram is commutative:

$$\begin{tikzcd}
A' \arrow[r, "f"] \arrow[d, "(g^f)(\alpha)"'] & A \arrow[d, "\alpha"] \\
B' & B \arrow[l, "g"]
\end{tikzcd}$$

We also use the notations $\mathrm{Hom}(A,B) = B^A$ and $\mathrm{Hom}(f,g) = g^f$.

Exercise 49 Let $A' = B' = \{1,2,3\}$, $A = B = \{1,2\}$, $f(1) = f(2) = 1$, $f(3) = 2$, $g(1) = 2$ and $g(2) = 3$. Find the function $\mathrm{Hom}(f,g)$.

Exercise 50 Let $f : A' \to A$ and $g : B \to B'$ be two functions. Prove that:
a) If f is surjective and g is injective, then $\mathrm{Hom}(f,g)$ is injective;
b) If $A' \neq \emptyset$, g is surjective and f is injective, then $\mathrm{Hom}(f,g)$ is surjective;
c) g is injective if and only if for any set A, $\mathrm{Hom}(1_A, g) : \mathrm{Hom}(A,B) \to \mathrm{Hom}(A,B')$ is injective;
d) f is surjective if and only if for any set B, $\mathrm{Hom}(f, 1_B) : \mathrm{Hom}(A,B) \to \mathrm{Hom}(A',B)$ is injective.

4.3.4 The power set, and the characteristic function of a subset

Recall that the power set of a set A is the set $\mathcal{P}(A) = \{X \mid X \subseteq A\}$, that is, we have $X \in \mathcal{P}(A) \iff X \subseteq A$. O function $f : A \to B$ induces the functions

$$f_* : \mathcal{P}(A) \to \mathcal{P}(B), \qquad f_*(X) = f(X),$$

$$f^* : \mathcal{P}(B) \to \mathcal{P}(A), \qquad f^*(Y) = f^{-1}(Y).$$

Exercise 51 Let $f : A \to B$ and $g : B \to C$ be functions. Prove that:
a) $1_{A*} = 1_A{}^* = 1_{\mathcal{P}(A)}$;
b) $(g \circ f)_* = g_* \circ f_*$; $(g \circ f)^* = f^* \circ g^*$;
c) $f^* \circ f_* \circ f^* = f^*$;
d) If $\varphi = f^* \circ f_*$ and $\psi = f_* \circ f^*$, then $\varphi \circ \varphi = \varphi$ and $\psi \circ \psi = \psi$.

Exercise 52 Let $f : A \to B$ be a function.
a) Prove that the following statements are equivalent:
(i) f is injective; (ii) f_* is injective; (iii) $f^* \circ f_* = 1_{\mathcal{P}(A)}$; (iv) f^* is surjective.
b) Prove that the following statements are equivalent:
(i) f is surjective; (ii) f_* is surjective; (iii) $f_* \circ f^* = 1_{\mathcal{P}(B)}$; (iv) f^* is injective.

Definition 4.3.7 Let A a set and $X \subseteq A$. The function

$$\chi_X : A \to \{0,1\}, \qquad \chi_X(x) = \begin{cases} 1, & \text{if } x \in X, \\ 0, & \text{if } x \notin X \end{cases}$$

is called the **characteristic function** of the subset X. Thus we have the function

$$\varphi_A : \mathcal{P}(A) \to \mathrm{Hom}(A,\{0,1\}), \qquad \varphi_A(X) = \chi_X$$

Exercise 53 Let A a set. Prove that the function φ_A is bijective, and we have $\varphi_A^{-1}(\chi) = \chi^{-1}(1)$, for any function $\chi : A \to \{0,1\}$;

Exercise 54 If $X, Y \subseteq A$, then:
(1) $X \subseteq Y \Leftrightarrow \chi_X(x) \le \chi_Y(x), \quad \forall x \in A$,
(2) $\chi_{\bar{X}}(x) = 1 - \chi_X(x), \quad \forall x \in A$,
(3) $\chi_{X \cap Y}(x) = \chi_X(x)\chi_Y(x), \quad \forall x \in A$,
(4) $\chi_{X \cup Y}(x) = \chi_X(x) + \chi_Y(x) - \chi_X(x)\chi_Y(x), \quad \forall x \in A$,
(5) $\chi_{X \setminus Y}(x) = \chi_X(x)(1 - \chi_Y(x)), \quad \forall x \in A$,
(6) $\chi_{X \Delta Y}(x) = \chi_X(x) + \chi_Y(x) - 2\chi_X(x)\chi_Y(x), \quad \forall x \in A$.

Remark 4.3.8 The above properties of the characteristic function can be used to prove equalities between sets. For example, let us show that $(X\Delta Y)\Delta Z = X\Delta(Y\Delta Z)$:

$$\begin{aligned}\chi_{(X\Delta Y)\Delta Z} &= \chi_{X\Delta Y} + \chi_Z - 2\chi_{X\Delta Y}\chi_Z = \\ &= \chi_X + \chi_Y - 2\chi_X\chi_Y - 2(\chi_X + \chi_Y - 2\chi_X\chi_Y)\chi_Z = \\ &= \chi_X + \chi_Y + \chi_Z - 2(\chi_X\chi_Y + \chi_X\chi_Z + \chi_Y\chi_Z) + 4\chi_X\chi_Y\chi_Z. \qquad (*)\end{aligned}$$

Similarly, if we compute $\chi_{X\Delta(Y\Delta Z)}$, we get the same term (*). Thus, by the commutativity of Δ, and from (*) we deduce

$$\begin{aligned}\chi_{X\Delta(Y\Delta Z)} &= \chi_{(Y\Delta Z)\Delta X} = \\ &= \chi_Y + \chi_Z + \chi_X - 2(\chi_Y\chi_Z + \chi_Y\chi_X + \chi_Z\chi_X) + 4\chi_Y\chi_Y\chi_X = \\ &= \chi_X + \chi_Y + \chi_Z - 2(\chi_X\chi_Y + \chi_X\chi_Z + \chi_Y\chi_Z) + 4\chi_X\chi_Y\chi_Z.\end{aligned}$$

4.4 Equivalence relations

4.4.1 Important classes of homogeneous relations

Definition 4.4.1 Let $\rho = (A, A, R)$ be a homogeneous relation. We say that:

a) ρ is **reflexive**, if for any $x \in A$, $x\rho x$, that is

$$(\forall\, x \in A)(x\rho x);$$

ρ is **irreflexive** if for any $a \in A$ we have $a \not\rho\, a$, that is, we have $\neg(a\rho a)$;

b) ρ is **transitive**, if for any $x, y, z \in A, x\rho y$ and $y\rho z$ implies $x\rho z$, that is,

$$(\forall\, x, y, z \in A)(x\rho y \wedge y\rho z \to x\rho z);$$

c) ρ is **symmetric**, if for any $x, y \in A, x\rho y$ implies $y\rho x$, that is,

$$(\forall\, x, y \in A)(x\rho y \to y\rho x);$$

d) ρ is **antisymmetric**, if for any $x, y \in A, x\rho y$ and $y\rho x$ implies $x = y$, that is,

$$(\forall\, x, y \in A)(x\rho y \wedge y\rho x \Rightarrow x = y);$$

ρ is **asymmetric**, if for any $x, y \in A$, $x\rho y$ implica $y \not\rho x$;

e) ρ is **preorder**, if ρ is reflexive and transitive. In this case we say that (A, ρ) is a **preordered set**;

f) ρ is an **equivalence relation**, if ρ is reflexive, transitive and symmetric. We denote by $\mathcal{E}(A)$ the set of the relations of equivalence defined on A;

g) ρ is an **order relation** (also called **partial order**), if ρ is reflexive, transitive and antisymmetric. In this case, we say that (A, ρ) is an **ordered set** (also called **poset** in the literature);

h) ρ is a **strict order**, if ρ is irreflexive and transitive.

i) ρ is a **tolerance relation**, if ρ is reflexive and symmetric.

Remark 4.4.2 The following statements are easy to prove:

1) ρ is reflexive $\Leftrightarrow 1_A \subseteq \rho$;
2) ρ is transitive $\Leftrightarrow \rho^2 \subseteq \rho$;
3) ρ is symmetric $\Leftrightarrow \rho = \rho^{-1}$;
4) ρ is antisymmetric $\Leftrightarrow \rho \cap \rho^{-1} \subseteq 1_A$;
5) ρ is reflexive and antisymmetric $\Rightarrow \rho \cap \rho^{-1} = 1_A$;
6) ρ is irreflexive $\Leftrightarrow \rho \cap 1_A = \emptyset$;
6) ρ is asymmetric $\Leftrightarrow \rho \cap \rho^{-1} = \emptyset$;
7) ρ is a preorder $\Rightarrow \rho^2 = \rho$;
8) ρ is an equivalence $\Leftrightarrow 1_A \subseteq \rho$ and $\rho = \rho^2 = \rho^{-1}$;
9) ρ is an equivalence and an order $\Leftrightarrow \rho = 1_A$.

Example 4.4.3 1) On the set $\mathbb{Z}$ of integers, the divisibility relation is a preorder; it is not symmetric (because $3 \mid 6$ but $6 \nmid 3$), and it is not antisymmetric (because for instance $3 | -3$ and $-3 | 3$, but $-3 \neq 3$).

2) On the set $\mathbb{N}$ of natural numbers, the divisibility relation is an order, hence $(\mathbb{N}, |)$ is an ordered set.

3) On the set $\mathbb{Z}$ of integers, the relation of congruence modulo n, defined by $a \equiv b \pmod{n} \Leftrightarrow n | a - b$, is an equivalence.

4) The universal relation $(A, A, A \times A)$ is an equivalence.

5) The restriction to a subset of an equivalence is also an equivalence. More precisely, if $\rho = (A, A, R)$ is equivalence on A, and $B \subseteq A$, then $(B, B, R \cap (B \times B))$ is an equivalence on B.

4.4.2 Equivalences and partitions

Definition 4.4.4 If ρ is an equivalence relation on the set A, then the section

$$\rho\langle x\rangle = \{y \in A \mid x\rho y\}$$

with respect to the element $x \in A$ is called an **equivalence class**. The set of all these classes is called **the quotient set (or factor set)** modulo ρ:

$$A/\rho = \{\rho\langle x\rangle \mid x \in A\}.$$

Example 4.4.5 1) On the set $\mathbb{Z}$ of integers, the congruence relation $a \equiv b \pmod{n}$ (where $n \neq 0$) corresponds to the quotient set

$$\mathbb{Z}/\equiv \pmod{n} = \{\widehat{0}, \widehat{1}, \widehat{2}, \dots, \widehat{n-1}\},$$

where

$$\begin{aligned}\widehat{k} = \equiv \pmod{n}\langle k\rangle &= \{x \in \mathbb{Z} \mid x \equiv k \pmod{n}\}\\ &= \{x \in \mathbb{Z} \mid n|x-k\}\\ &= \{x \in \mathbb{Z} \mid \exists j \in \mathbb{Z} :\ x = jn + k\}\\ &= n\mathbb{Z} + k\end{aligned}$$

2) $A/1_A = \{\{x\} : x \in A\}$ and $A/(A \times A) = \{A\}$.

Lemma 4.4.6 *If ρ is an equivalence on the set A, and $x, y \in A$, then the following statements are equivalent:*
(i) $x\rho y$; (ii) $y \in \rho\langle x\rangle$; (iii) $\rho\langle x\rangle = \rho\langle y\rangle$.

Proof. (i) ⇔ (ii) is clear by the definition.

(i) ⇒ (iii) Let $x\rho y$, and let $z \in \rho\langle x\rangle$. Then $x\rho y$ and $x\rho z \Rightarrow z\rho x$ and $x\rho y$ (because ρ is symmetric), hence $z\rho y$ (because ρ is transitive) $\Rightarrow z \in \rho\langle y\rangle$, hence $\rho\langle x\rangle \subseteq \rho\langle y\rangle$. Analogously, $\rho\langle y\rangle \subseteq \rho\langle x\rangle$, hence (iii) holds.

(iii) ⇒ (i) If $\rho\langle x\rangle = \rho\langle y\rangle$, then $y \in \rho\langle y\rangle = \rho\langle x\rangle \Rightarrow y\rho x \Rightarrow x\rho y$.

Definition 4.4.7 Let A be a nonempty set, and let $\pi \subseteq \mathcal{P}(A) \setminus \{\emptyset\}$. We say that π is a **partition** of A if the following two conditions hold:

(1) $A = \bigcup_{B\in\pi} B$,

(2) For all $B_1, B_2 \in \pi$, $B_1 \neq B_2 \Rightarrow B_1 \cap B_2 = \emptyset$, that is, every two distinct classes from π are disjoint.

If $B \in \pi$ and $b \in B$, then we say that b is a **representative** of B. We denote by $P(A)$ the set of partitions of A.

Equivalence relations and partitions determine each other.

Theorem 4.4.8 *Let A be a nonempty set.*

1) *If ρ is an equivalence on A, then the quotient set*

$$A/\rho = \{\rho\langle x\rangle \mid x \in A\}$$

is a partition of A.

2) *Let $\pi \subseteq \mathcal{P}(A) \setminus \{\emptyset\}$ be a partition of A, and define the relation*

$$\rho_\pi = (A, A, R_\pi), \quad R_\pi = \bigcup_{B\in\pi} (B \times B),$$

that is, $x\rho_\pi y \Leftrightarrow \exists B \in \pi : x, y \in B$.

Then ρ_π is an equivalence relation on A.

3) *Consider the functions*

$$\phi : \mathcal{E}(A) \to \mathcal{P}(A), \quad \phi(\rho) = A/\rho,$$

$$\psi : \mathcal{P}(A) \to \mathcal{E}(A), \quad \psi(\pi) = \rho_\pi.$$

Then $\psi \circ \phi = \mathbf{1}_{\mathcal{E}(A)}$ and $\phi \circ \psi = \mathbf{1}_{\mathcal{P}(A)}$.

Proof. 1) We show that $A=\bigcup_{x\in A}\rho\langle x\rangle$. The inclusion "$\supseteq$" is obvious, because $\rho\langle x\rangle\subseteq A$ for any $x\in A$. Furthermore, for any $y\in A$, $y\in\rho\langle y\rangle$ (because ρ is reflexive), hence $y\in\bigcup_{x\in A}\rho\langle x\rangle$; from this we get the inclusion "$\subseteq$".

Now assume that $\rho\langle x\rangle\cap\rho\langle y\rangle\neq\emptyset$, where $x,y\in A$. We show that the classes $\rho\langle x\rangle$ and $\rho\langle y\rangle$ are equal. Indeed, by hypothesis, $\exists u\in\rho\langle x\rangle\cap\rho\langle y\rangle\Rightarrow x\rho u$ and $y\rho u\Rightarrow x\rho u$ and $u\rho y$ (because ρ is symmetric) $\Rightarrow x\rho y$ (ρ is transitive) $\Rightarrow\rho\langle x\rangle=\rho\langle y\rangle$ by Lemma 4.4.6.

2) ρ_π is reflexive, because $\forall x\in A=\bigcup_{B\in\pi}B\Rightarrow\exists B\in\pi:x\in B\Rightarrow(x,x)\in B\times B\Rightarrow x\rho_\pi x$.

ρ_π is transitive, because $\forall x,y,z\in A:x\rho_\pi y$ and $y\rho_\pi z\Rightarrow\exists B,C\in\pi:x,y\in B$ and $y,z\in C$. Hence $y\in B\cap C$, and we get $B=C$ (because $B\neq C$, so by definition $B\cap C=\emptyset$, contradiction). Hence $x,z\in B=C\Rightarrow x\rho_\pi z$.

Furthermore, ρ_π is symmetric, because $\forall x,y\in A:x\rho_\pi y\Rightarrow\exists B\in\pi:x,y\in B\Rightarrow y\rho_\pi x$.

Hence ρ is an equivalence.

3) For any $\rho\in\mathcal{E}(A)$, $(\psi\circ\phi)(\rho)=\psi(\phi(\rho))=\rho_{\phi(\rho)}=\rho_{A/\rho}$. We show that $\rho_{A/\rho}=\rho$. Indeed,

$$\begin{aligned}x\rho_{A/\rho}y&\Leftrightarrow\exists B\in A/\rho:x,y\in B\Leftrightarrow\\&\Leftrightarrow\exists z\in A:B=\rho\langle z\rangle\in A/\rho\text{ and }\quad x,y\in B=\rho\langle z\rangle\Leftrightarrow\\&\Leftrightarrow\exists z\in A:x\rho z\text{ and }y\rho z\Leftrightarrow\\&\Leftrightarrow\exists z\in A:x\rho z\quad\text{and}\quad z\rho y\quad(\rho\quad\text{is symmetric})\Leftrightarrow\\&\Leftrightarrow x(\rho\circ\rho)y\Leftrightarrow x\rho y\qquad(\text{because}\ \ \rho^2=\rho).\end{aligned}$$

Hence $(\psi\circ\phi)(\rho)=\rho$, that is, $\psi\circ\phi=1_{\mathcal{E}(A)}$.

For any $\pi\in P(A)$, $(\phi\circ\psi)(\pi)=\phi(\psi(\pi))=A/\psi(\pi)=A/\rho_\pi$. We show that $A/\rho_\pi=\pi$. Indeed, $B\in A/\rho_\pi\Leftrightarrow\exists z\in A:B=\rho_\pi\langle z\rangle$. Here $z\in A=\bigcup_{C\in\pi}C$, hence there exists $C\in\pi$ such that $z\in C$, and

$$B=\{x\in A\mid x\rho_\pi z\}=\{x\in A\mid x\in C\}=C\in\pi,$$

hence $A/\rho_\pi\subseteq\pi$.

Conversely, for any $C\in\pi$, there exists $z\in C$ such that

$$C=\{x\in A\mid x\rho_\pi z\}=\rho\langle z\rangle\in A/\rho_\pi,$$

from where we get $\pi\subseteq A/\rho_\pi$. Hence $(\phi\circ\psi)(\pi)=\pi$, that is, $\phi\circ\psi=1_{P(A)}$. ■

Exercise 55 Let $\rho=(A,B,R)$ a relation. Prove:

a) If ρ is reflexive, symmetric and antisymmetric, then $\rho=1_A$;

b) If ρ is reflexive and transitive, then $\rho^2=\rho$.

Exercise 56 Let $A=\{1,2,3,4\}$.

a) If $\rho=\{(1,1),\dots,(4,4),(1,2),(2,1),(3,2),(2,3),(1,3),(3,1)\}$, find the partition corresponding to ρ.

b) If $\pi=\{\{1,2\},\{3\},\{4\}\}$, find the equivalence relation corresponding to π.

Exercise 57 Find all the equivalence relations on a set with $1,2,3$, respectively 4 elements.

Exercise 58 Prove that:

a) $(\mathbb{Z},|)$ is a preordered set, "|" it is not symmetric and it is not antisymmetric;

b) $(\mathbb{N},|)$ is an ordered set;

Exercise 59 On the set $\mathbb{C}$ of complex numbers consider the relations ρ_1 and ρ_2, where $z\rho_1w\Leftrightarrow|z|=|w|$ and $z\rho_2w\Leftrightarrow z=w=0$ or $\arg z=\arg w$. Prove that ρ_1 and ρ_2 are equivalence relations, and represent graphically the classes from $\mathbb{C}/\rho_1$ and $\mathbb{C}/\rho_2$.

Exercise 60 Let ρ_1 and ρ_2 be equivalence relations on the set A. Prove that:

a) ρ_1^{-1} and $\rho_1\cap\rho_2$ are equivalence relations. (More generally, if $(\rho_i)_{i\in I}$ are equivalence relations on A, then $\bigcap_{i\in I}\rho_i$ is an equivalence on the set A.)

b) $\complement\rho_1$ and $\rho_1\cup\rho_2$ are not, in general, equivalence relations;

c) $\rho_1\circ\rho_2$ is an equivalence if and only if $\rho_1\circ\rho_2=\rho_2\circ\rho_1$. In this case, show that $\rho_1\circ\rho_2$ is the smallest equivalence relation which contains ρ_1 and ρ_2.

Exercise 61 Let ρ_1 and ρ_2 two relations on the set A.

a) Prove that $(\rho_1\cup\rho_2)^2=\rho_1^2\cup\rho_2^2\cup(\rho_1\circ\rho_2)\cup(\rho_2\circ\rho_1)$.

b) Assume that ρ_1 and ρ_2 are equivalence relations. Prove that $\rho_1\cup\rho_2$ an equivalence if and only if $\rho_1\circ\rho_2$ and $\rho_2\circ\rho_1$ are subrelations of $\rho_1\cup\rho_2$.

4.5 Factorization theorems for functions

Definition 4.5.1 a) Let $f : A \to B$ be a function. The relation $\ker f$ on the set A, defined by:

$$a_1 \rho a_2 \Leftrightarrow f(a_1) = f(a_2)$$

is called the **kernel** of f.

b) Let ρ be a relation of equivalence on the set A. The function

$$p_\rho : A \to A/\rho, \quad p_\rho(x) = \rho\langle x\rangle$$

is called the **canonical projection** of A onto the quotient set A/ρ.

It is easy to show that the relation $\ker f$ is an equivalence on A, while the canonical projection $p_\rho : A \to A/\rho$ is surjective, and we have $\ker p_\rho = \rho$.

Exercise 62 Let $f : A \to B$ a function. Prove that:
1) $\ker f$ is an equivalence relation on A, and $\ker f = f^{-1} \circ f$,
2) $A/\ker f = \{f^{-1}(b) \mid b \in \operatorname{Im} f\}$,
3) f is injective $\Leftrightarrow \ker f = 1_A$,
4) f is surjective $\Leftrightarrow \operatorname{Im} f = B$.
5) The graph of the relation $f \circ f^{-1}$ is $\Delta_{\operatorname{Im} f}$, where $\Delta_{\operatorname{Im} f} = \{(b, b) \mid b \in \operatorname{Im} f\}$.

Exercise 63 If ρ is an equivalence on A, then the canonical projection $p_\rho : A \to A/\rho$ is surjective, and we have $\ker p_\rho = \rho$.

Theorem 4.5.2 (The 1st Factorization Theorem) *If $f : A \to B$ is a function, then there exists a the unique bijective function $\bar{f} : A/\ker f \to \operatorname{Im} f$ such that the following diagram is commutative, that is, $f = \iota \circ \bar{f} \circ p_{\ker f}$, where $\iota : \operatorname{Im} f \to B$, $i(y) = y$ is the canonical inclusion. Moreover, for any $x \in A$, we have $\bar{f}(\ker f\langle x\rangle) = f(x)$.*

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "p_{\ker f}"'] & B \\
A/\ker f \arrow[r, "\bar{f}"'] & \operatorname{Im} f \arrow[u, "\iota"']
\end{tikzcd}$$

Proof. (*Uniqueness*) We assume that $\bar{f}$ exists, and we prove that it is unique. Indeed, we have:

$$f(x) = (\iota \circ \bar{f} \circ p_{\ker f})(x) = \iota(\bar{f}(p_{\ker f}(x))) = \bar{f}((\ker f)\langle x\rangle).$$

Hence $\bar{f}((\ker f)\langle x\rangle) = f(x)$ is uniquely defined for all $x \in A$.

(*Existence*) Let $\bar{f} : A/\ker f \to \operatorname{Im} f$, $\bar{f}((\ker f)\langle x\rangle) = f(x) \in \operatorname{Im} f$.

• First, we have to show that the definition of $\bar{f}$ does not depend on the choice of the representative $x \in (\ker f)\langle x\rangle$. Indeed, let $y \in (\ker f)\langle x\rangle$ be another representative, that is, $x \ker f y$. Then we have

$$\bar{f}((\ker f)\langle x\rangle) = f(x) = f(y) = \bar{f}((\ker f)\langle x\rangle),$$

hence $\bar{f}$ is indeed well-defined.

• We show that $\bar{f}$ is injective. Indeed, let $x, y \in A$ such that $\bar{f}((\ker f)\langle x\rangle) = \bar{f}((\ker f)\langle y\rangle)$. In follows that $f(x) = f(y)$, hence $x \ker f y$, that is, $(\ker f)\langle x\rangle = (\ker f)\langle y\rangle$.

• We show that $\bar{f}$ is surjective. Indeed, let $y \operatorname{Im} f$, and let $x \in A$ such that $f(x) = y$. It follows that $y = \bar{f}((\ker f)\langle x\rangle)$, hence $\bar{f}$ is surjective.

• Finally, we show that the diagram is commutative. Indeed, for any $x \in A$ we have

$$(\iota \circ \bar{f} \circ p_{\ker f})(x) = \iota(\bar{f}(p_{\ker f}(x))) = \bar{f}((\ker f)\langle x\rangle) = f(x),$$

hence $\iota \circ \bar{f} \circ p_{\ker f} = f$. ■

Exercise 64 Apply the 1st Factorization Theorem to each of the following four functions:
a) $f, g : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, $g(x) = x^4$;
b) $f, g : \mathbb{C} \to \mathbb{C}$, $f(z) = z^2$, $g(z) = z^4$.

Chapter 5

ORDERED SETS

The concept of ordered set formalizes and generalizes the intuitive idea of ordering or arranging the objects of a collection.

5.1 Order relations

Let $\rho = (A, A, R)$ be a homogeneous relation. Recall that ρ is an **order relation**, respectively (A, ρ) is an **ordered set (poset)** if ρ is reflexive, transitive and antisymmetric. If ρ is an order relation, then instead of $x\rho y$ we often denote $x \leq y$. We also denote $\mathcal{O}(A) = \{\rho = (A, A, R) \mid \rho \text{ relation of order }\}$ the set of the relations of order on A.

Recall that ρ is a **strict order** if ρ is irreflexive and transitive. Notations: $x < y$, if $x \leq y$ and $x \neq y$; $x > y$, if $y < x$ etc.

Definition 5.1.1 We say that (A, ρ) is a **totally ordered set** (or **chain**) if :

$$\text{for any } \ x, y \in A \ \text{ are loc } \ x\rho y \ \text{ or } \ y\rho x$$

(this means that $\rho \cup \rho^{-1} = A \times A$ is the universal relation, that is, any two elements of A are **comparable** with respect to the relation ρ).

Example 5.1.2 1) $(\mathbb{N}, \leq), (\mathbb{Z}, \leq), (\mathbb{Q}, \leq), (\mathbb{R}, \leq)$ are totally ordered sets.

2) $(\mathbb{N}, |)$, (where "$|$" is the relation of divisibility) is an ordered set which it is not totally ordered, because, for instance, 2 and 3 nu are comparable.

3) If A is a set, then $(\mathcal{P}(A), \subseteq)$ is an ordered set. If A has more than one element, then $(\mathcal{P}(A), \subseteq)$ it is not totally ordered.

4) If (A, ρ) is an ordered (totally ordered) set, and $B \subseteq A$, then $(B, \rho \cap (B \times B))$ is ordered (totally ordered).

Exercise 65 Let $A \neq \emptyset$ and let $\rho, \rho' \in \mathcal{O}(A)$. Prove that:
a) $\rho \cap \rho', \rho^{-1} \in \mathcal{O}(A)$.
b) $\complement\rho \notin \mathcal{O}(A)$.
c) In general $\rho \cup \rho' \notin \mathcal{O}(A)$.
d) If σ is a strict order on A, then σ is *asymmetric*, while $\sigma \cup 1_A \in \mathcal{O}(A)$.
e) $\sigma := \rho \setminus 1_A$ is a strict order relation on A.
f) The order relation ρ is total $\Longleftrightarrow$ ρ satisfies the **trichotomy** property, that is, for any $x, y \in A$, exactly one of the following three statements is true: (1) $a\sigma b$; (2) $a = b$; (3) $a\sigma^{-1}b$.

A finite ordered set may be represented graphically by a **Hasse diagram**, according to the following rule: if $x < y$ and if it does not exist $z \in A$ such that $x < z < y$, then we place the dot y above the dot x and we join them with a segment.

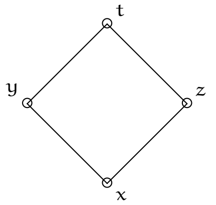
Example 5.1.3 Let $A = \{x, y, z, t\}$, and consider the order relations on the set A with graphs

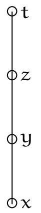
$$R = \{(x, x), (y, y), (z, z), (t, t), (x, y), (x, z), (x, t), \ (y, t), (z, t)\},$$

respectively

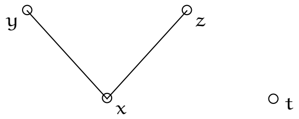
$$R = \{(x, x), (y, y), (z, z), (t, t), (x, y), (x, z), (x, t), (y, z), \ (y, t), (z, t)\}.$$

Then the Hasse diagrams are:





In the following diagram $x < y, x < z$, y and z are incomparable, and t is incomparable with x, y, z.



Exercise 66 Draw the Hasse diagrams of the following ordered sets:
a) $(\mathcal{P}(\{a, b, c, d\}), \subseteq)$;
b) The set of divisors of 60, ordered by divisibility;

Definition 5.1.4 Let $(A, \leq)$ and $(B, \leq)$ be two ordered sets, and let $f : A \to B$ be a function.
a) We say that f is **increasing (order-preserving)**, respectively **decreasing (order-reversing)**, if for any $x, y \in A$,

$$x \leq y \Rightarrow f(x) \leq f(y) \qquad (f(y \leq f(x));$$

further, f is **order isomorphism**, if f is increasing, bijective and f^{-1} is increasing;

Example 5.1.5 1) The sets ordered $\mathbb{N} = \{1, 2, 3, \dots\}$ and $2\mathbb{N} = \{2, 4, 6, \dots\}$ are isomorphic, because $f : \mathbb{N} \to 2\mathbb{N}$, $f(n) = 2n$ is an order isomorphism.
2) The sets ordered $(\mathbb{N}, |)$ and $(\mathbb{N}, \leq)$ are not isomorphic. Indeed, if we would have an isomorphism $f : (\mathbb{N}, |) \to (\mathbb{N}, \leq)$, then let $f(2) = n$ and $f(3) = m$, where $n \neq m$. If $n < m$, then $f^{-1}(n) = 2|3 = f^{-1}(m)$, while if $m < n$, then $f^{-1}(m) = 3|2 = f^{-1}(n)$, hence we have contradiction in both cases.

Exercise 67 Find all the order relations on the set $A = \{a, b, c\}$ (using Hasse diagrams). Classify these orderings into isomorphism classes.

Exercise 68 Let $(A, \leq)$, $(B, \leq)$ and $(C, \leq)$ be ordered sets, and let $f : A \to B$ and $g : B \to C$ be two functions. Prove that:
a) If f and g are increasing (decreasing), then $g \circ f$ is increasing.
b) If f is increasing (decreasing) and g is decreasing (increasing), then $g \circ f$ is decreasing.

Exercise 69 Let $(A, \leq)$ and $(B, \leq)$ be ordered sets, and let $f : A \to B$ be a bijective and increasing function. Prove that:
a) If A is totally ordered, then f^{-1} is increasing, and B is also totally ordered.
b) The function $1_{\mathbb{N}^*} : (\mathbb{N}^*, |) \to (\mathbb{N}^*, \leq)$ is bijective and increasing, but it is not an order isomorphism.

Definition 5.1.6 Let $(A, \leq)$ be an ordered set, and let $x \in A$. We say that x is **a least element** or **minimum** (respectively **a greatest element** or **maximum**) of A, if for any $a \in A$, $x \leq a$ (respectively for any $a \in A$, $a \leq x$).
Notation: $x = \min A$ (respectively $x = \max A$).

Remark 5.1.7 If there exists a least element (a greatest element), then it is unique. Indeed, if for instance x and x' are both least elements, then $x \leq x'$ (because x is a least element) and $x' \leq x$ (because x' is a least element), hence by antisymmetry we get $x = x'$.

Example 5.1.8 1) In $(\mathbb{N}, \leq)$, $x = 0$ is the least element and there is no greatest element.
2) In $(\mathbb{N}, |)$, 1 is the least element, and 0 is the greatest element, because $1|a$ and $a|0$ for any $a \in \mathbb{N}$.
3) In $(\mathbb{N} \setminus \{0, 1\}, |)$, there is no least element, and no greatest element.
4) In $(\mathcal{P}(A), \subseteq)$, $\min \mathcal{P}(A) = \emptyset$, and $\max \mathcal{P}(A) = A$.

Definition 5.1.9 In the ordered set $(A, \leq)$, the element x is **minimal** (respectively **maximal**), if $\forall a \in A : a \leq x \Rightarrow a = x$ (respectively $\forall a \in A : x \leq a \Rightarrow a = x$). In opther words, $x \in A$ is a minimal element (respectively maximal element), if no element a of A satisfies $a < x$ (respectively $a > x$).

Example 5.1.10 1) In $(\mathbb{N}, \leq)$, the number 0 is a minimal element, and there are no maximal elements.

2) In $(\mathbb{N}, |)$, the number 1 is a minimal element and 0 is a maximal element.

3) In $(\mathbb{N} \setminus \{0, 1\}, |)$, the prime numbers are minimal elements, and there are no maximale elements.

4) It is clear by definitions that if there exists the least (respectively greatest) element, then this is the unique minimal (respectively maximal) element. The converse statement is not true: for instance, if $A = \{2^k \mid k \in \mathbb{N}\} \cup \{3, 9\}$, then in the ordered set $(A, |)$, $a = 9$ is the unique maximal element and there is no greatest element.

5) If $(A, \leq)$ is a totally ordered set, then the notions of minimal (respectively maximal) element and the least element (respectively greatest element) are equivalent.

Exercise 70 Let $(A, \leq)$ be an ordered set. Prove that if there exists $a = \min A$, then a is the unique minimal element of A, while the converse statement is not true.

5.2 Lattices

Definition 5.2.1 Let $(A, \leq)$ a ordered set, $x \in A$ and let $B \subseteq A$.

a) We say that x is **minorant (lower bound)** (respectively **majorant (upper bound)**) of B, if for any $b \in B$ we have $x \leq b$ (respectively for any $b \in B$ we have $b \leq x$).

b) We say that x is an **infimum** or (respectively a **supremum**) of B, if x is *greatest minorant* of B (respectively x is *the least majorant* of B). Notation: $x = \inf B$ or $x = \inf_A B$ (respectively $x = \sup B$ or $x = \sup_A B$).

Example 5.2.2 1) In $(\mathbb{N} \setminus \{0, 1\}, |)$, the subset $B = \{2k + 1 \mid k \in \mathbb{N}\}$ has a unique minorant $x = 1$, hence $\inf B = 1$; further, B has no majorants, and no supremum.

2) In $(\mathbb{R}, \leq)$ the interval $B = (1, 3]$ has as minorant every $x \leq 1$, and as majorant every $x \geq 3$; further, $\inf B = 1 \notin B$, $\sup B = 3 \in B$.

3) If $(A, \leq)$ is an ordered set, then every element of A is a minorant and a majorant of $B = \emptyset$. Furthermore, $\exists \inf \emptyset \Leftrightarrow \exists \max A \Leftrightarrow \exists \sup A$, $\exists \sup \emptyset \Leftrightarrow \exists \min A \Leftrightarrow \exists \inf A$ and then $\inf \emptyset = \max A = \sup A$, $\sup \emptyset = \min A = \inf A$.

4) Any subset $B \subseteq A$ has at most one infimum, and at most one supremum; further, if B has minorant x (respectively majorant y) which belongs to B, then $x = \inf B$ (respectively $y = \sup B$).

Exercise 71 Let $(A, \leq)$ be an ordered set, and let $X \subseteq B \subseteq A$.

a) If there exists $\inf_B X$ and $\inf_A X$, then $\inf_B X \geq \inf_A X$.

b) If there exists $\sup_B X$ and $\sup_A X$, then $\sup_B X \leq \sup_A X$.

Definition 5.2.3 a) The ordered set $(A, \leq)$ is called a **lattice**, if every subset with two elements of A has infimum and supremum (that is, for any $a, b \in A$, $a \neq b$, $\exists \inf\{a, b\}$ and $\exists \sup\{a, b\}$).

b) $(A, \leq)$ is a **complete lattice**, if every subset of A has infimum and supremum (that is, for any $B \subseteq A$, $\exists \inf B$ and $\exists \sup B$).

Example 5.2.4 1) $(\mathbb{N}, |)$ is lattice. Indeed, for any $a, b \in \mathbb{N}$, $\inf\{a, b\} = \mathrm{cmmdc}(a, b)$ while $\sup\{a, b\} = \mathrm{cmmmc}(a, b)$.

2) If $(A, \leq)$ is totally ordered, then $(A, \leq)$ is a lattice: $\forall a, b \in A : \inf\{a, b\} = \min\{a, b\}$, and $\sup\{a, b\} = \max\{a, b\}$.

3) $(\mathbb{R}, \leq)$ is not a complete lattice, because for instance, the interval $B = (-\infty, 0)$ has no minorant, hence it does not have supremum in $(\mathbb{R}, \leq)$.

4) $(\mathcal{P}(A), \subseteq)$ is a complete lattice. If $X = \{X_i \mid i \in I\} \subseteq \mathcal{P}(A)$, then $\inf X = \bigcap_{i \in I} X_i$, and $\sup X = \bigcup_{i \in I} X_i$.

Exercise 72 Find all the latticesle with 1, 2, 3, 4, 5 and respectively 6 elements (up to isomorphism, using diagrams Hasse).

Theorem 5.2.5 (the characterization of complete lattices) *Let $(A, \leq)$ be an ordered set. The following statements are equivalent:*

(i) $(A, \leq)$ *is a complete lattice;*

(ii) *every subset of* A *has infimum;*

(iii) *every subset of* A *has supremum.*

Proof. (i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii) are clear by definition.

(ii) $\Rightarrow$ (i) We must show that every subset B of A has supremum. We denote by C the set of majorants lui B. We have $C \neq \emptyset$, because by (ii), there exists $\inf \emptyset = \max A \in C$. Let $x = \inf C$, which exists by the hypothesis. We show that $x = \sup B$. Indeed, for any $b \in B$ and $c \in C$ we have $b \leq c$ (by the definition of C), hence every $b \in B$ is a minorant of C; it follows that $\forall b \in B : b \leq x$ (because $x = \inf C$), hence x is a majorant of B.

Furthermore, let $x' \in A$ be a majorant of B, that is, we have $b \leq x', \forall b \in B$. Then $x' \in C$ (by bthe definition of C), so $x \leq x'$ (because $x = \inf C$), hence x is the least majorant of B, that is $x = \sup B$.

Similarly, one shows that (iii) $\Rightarrow$ (i). ■

Exercise 73 Let $(A, \leq)$ be a complete lattice, and let $f : A \to A$ be an increasing function. Prove that there exists $a \in A$ such that $f(a) = a$. (We say that a is **fixed point** of f.)

5.3 Well-ordered sets

Definition 5.3.1 Let $(A, \leq)$ be an ordered set. We say that A is **well-ordered** if every nonempty subset of A has the least element (that is, for any $B \subseteq A, B \neq \emptyset$, $\exists \min B \in B$).

Example 5.3.2 a) $(\mathbb{N}, \leq)$ is well-ordered.

b) If $(A, \leq)$ is well-ordered, then $(A, \leq)$ is totally ordered. The converse is not true: $(\mathbb{R}, \leq)$ is not well-ordered, because for instance, the interval $(0, 1)$ has no least element. Similarly, $(\mathbb{Z}, \leq)$ is totally ordered, but it is not well-ordered.

c) Any finite totally ordered set is well-ordered.

Indeed, we must show that $\forall\ B \subseteq A,\ B \neq \varnothing,\ \ \exists\ \min B$. Let $B \subseteq A,\ B \neq \varnothing$. Since A is finite, $\Rightarrow$ B is finite, hence let $B = \{b_1, b_2, \dots, b_n\}$. Since $(A, \leq)$ is totally ordered, it follows that every two elements of A (and also of B) are comparable. We compare the first two elements of B, we keep the least of them, and then we compare it with the third element of B, and we keep the least of them. By induction, after n steps, we find $\min B$.

The following theorem shows that on well-ordered sets one may apply the method of mathematical induction.

Theorem 5.3.3 (The characterization of well-ordered sets) *If $(A, \leq)$ is a nonempty ordered set, then the following statements are equivalent:*

(i) $(A, \leq)$ *is well-ordered.*

(ii) A *is totally ordered, there exists* $a_0 = \min A$*, and for any* $B \subseteq A$*, if* B *satisfies the properties:*

a) $a_0 \in B$*,*

b) *for any* $a \in A$*,* $\{x \in A \mid x < a\} \subseteq B \Rightarrow a \in B$*,*

then $B = A$*.*

Proof. (i) $\Rightarrow$ (ii) We assume that $(A, \leq)$ is well-ordered. Then A is totally ordered, and there exists $a_0 = \min A$. We assume that the second condition from (ii) is not true, that is, there exists $B \subseteq A$, such that a) and b) hold, and $B \neq A$.

Hence $A \setminus B \neq \emptyset$, and by hypothesis there exists $x = \min A \setminus B$. Here $x \in A \setminus B$, that is, $x \notin B$. Furthermore, $\forall y \in A : y < x \Rightarrow y \in B$ (because if $y \in A \setminus B$, then this contradicts the definition of x), hence $\{y \in A \mid y < x\} \subseteq B$, and from here $x \in B$ by b), which is a contradiction.

(ii) $\Rightarrow$ (i) We assume that (ii) is true and assume by contradiction that A it is not well-ordered, that is, there exists $B \subseteq A, B \neq \emptyset$, which has no least element. Then:

α) $a_0 \in A \setminus B$, because if $a_0 = \min A \in B$, then $a_0 = \min B$, contradiction.

β) For any $a \in A$, $\{x \in A \mid x < a\} \subseteq A \setminus B \Rightarrow a \in A \setminus B$. Indeed, if this would not be true, then $a \in B$, and because A is totally ordered, the elements x smaller than a are in $A \setminus B$, hence we get $a = \min B$, contradiction.

From α and β we deduce that the subset $A \setminus B$ satisfies hypotheses a) and b, so $A \setminus B = A$, that is $B = \emptyset$, contradiction. ■

Corollary 5.3.4 *Let* $(A, \leq)$ *a set nonempty well-ordered,* $a_0 = \min A$ *and let* P *a predicate of one variable defined on* A*. We assume that:*

1. $P(a_0)$ *is true,*

2. *For any* $a \in A$*, if* $P(x)$ *is true for any* $x < a$*, then* $P(a)$ *is true.*

Then $P(a)$ *is true for any* $a \in A$*.*

Proof. Let

$$B = \{a \in A \mid P(a) \text{ is true}\} \subseteq A,$$

which satisfies the hypotheses a) and b) of the previous theorem, hence $B = A$. ■

Exercise 74 a) Prove that:

1) If $(A, \leq)$ is a well-ordered set, and $f : A \to A$ is strictly increasing, then for any $a \in A$ we have $a \leq f(a)$.

2) Between two well-ordered sets there exists at most one isomorphism.

5.4 The axiom of choice

In mathematics we often encounter the statement: “we choose an element of the set ...”, or more precisely:

(AC$_0$) For any set $X \neq \emptyset$ there exists an element $x \in X$, hence $\{x\} \subseteq X$.

This is the simplest formulation of the axiom of choice. At a first sight, the statement (A_0) loks obvious, because $X \neq \emptyset$ means that that there exists at least one element $x \in X$. But what does it mean that “there exists an element $x \in X$”. Let us discuss two examples:

a) Let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(a) \cdot f(b) \leqslant 0$. We define the set

$$X = \{x \in [a, b] \mid f(x) = 0\}.$$

The Darboux Theorem shows that there exists $x_0 \in [a, b]$ such that $f(x_0) = 0$. One of the the proofs of the theorem gives a method to find the least element $x_0 \in [a, b]$ such that $f(x_0) = 0$.

b) Let $P(x)$ be a polynomial with complex coefficients. We consider the set

$$X = \{x \mid x \text{ complex number such that } P(x) = 0\}.$$

The Gauss-d’Alembert Theorem states that the set X is nonempty and finite, but no proof gives a method to find the roots of an arbitrary polynomial. This is a pure existence theorem.

In these examples we see that the expression “there exists an element $x \in X$” has a narrow meaning (one gives a method to find the element x) and a wide sense.

5.4.1 *The general form of the axiom of chioce is the following:*

(AC) *Let $F \neq \emptyset$ be a set of nonempty pairwise disjoint sets. Then there exists a set A with the following properties:*

(1) $A \subseteq \bigcup_{X \in F} X$;

(2) for any $X \in F$, $A \bigcap X$ *contains exactly one element.*

The set A is called a **selection** for F. Note that **(AC$_0$)** is a particular case of (AC).

The axiom of choice was formulated by Ernst Zermelo in 1904. It is independent of the other axioms, and there are various equivalent formulations, as we will see below. Considering set theory without the axiom of choice, and regarding this statement as a closed formula, Kurt Gödel has constructed a model for set theory in which the axiom of choice is true. On the other hand, Paul Cohen constructed in 1963 another model for set theory in which the axiom of choice is not true. In other words, set theory without the axiom of choice is undecidable.

Many proofs in mathematics effectively use the axiom of choice, that is, no proofs without the axiom of choice are known. Since the axiom of choice leads to some surprising consequences (for instance, the paradoxes of Hausdorff, Banach-Tarski, von Neumann), there are “constructivist” approaches which avoid its use.

Theorem 5.4.2 *The following statements are equivalent:*

1) *The axiom of choice* **(A)**.

2) *If $(X_i)_{i \in I}$ is a family of sets such that $I \neq \emptyset$ and $X_i \neq \emptyset$ for any $i \in I$, then the direct product $\prod_{i \in I} X_i$ is nonempty (that is, there exists a* **choice function** $f : I \to \bigcup_{i \in I} X_i$ *such that for any $i \in I$ we have $f(i) \in X_i$).*

3) **(Zorn’s Lemma)** *Let $(A, \leq)$ a nonempty ordered set. If every chain (totally ordered subset) $L \subseteq A$ has a majorant, then for any $a \in A$ there exists a maximal element $m \in A$ such that $a \leq m$.*

4) **(Zermelo’s Well-Ordering Principle)** *For any set A, there exists an order relation “$\leq$” such that $(A, \leq)$ is a well-ordered set.*

5) *Any surjective function has at least a section (right inverse).*

Exercise 75 Let A be a set, and consider the ordered set $(\mathcal{O}(A), \subseteq)$ of all order relations on A. By using Zorn’s lemma, prove that:

a) ρ is a maximal element of $\mathcal{O}(A)$ if and only if ρ is a total order.

b) For any $\rho \in \mathcal{O}(A)$ there exists a total order $\bar{\rho} \in \mathcal{O}(A)$ such that $\rho \subseteq \bar{\rho}$.

Chapter 6

LATTICES AND BOOLE ALGEBRAS

6.1 The lattice as an algebraic structure

In the previous chapter we have defined the lattice as an ordered set with additional properties. The existence of the infimum and the supremum for every pair of elements allow us to define two operations on the respective set.

Definition 6.1.1 a) The algebraic structure $(A, \wedge, \vee)$ with two binary operations "$\wedge$" and "$\vee$" is called a **lattice**, if the following axioms are satisfied:

1. both operations are associative,
2. both operations are commutative,
3. for any $x, y \in A$ we have $x \wedge (x \vee y) = x$ and $x \vee (x \wedge y) = x$ (**absorbtion**).

b) We say that A has **identity (unit) element** 1, if 1 is a neutral element with respect to $\wedge$, that is $x \wedge 1 = x$ for any $x \in A$. We say that A has **zero element** 0, if 0 is a neutral element with respect to $\vee$, that is $x \vee 0 = x$ for any $x \in A$.

b) Let $(A, \wedge, \vee)$ and $(A', \wedge, \vee)$ be lattices. The function $f : A \to A'$ is called a **morphism of lattices**, if for any $a, b \in A$ we have

$$f(a \vee b) = f(a) \vee f(b), \qquad f(a \wedge b) = f(a) \wedge f(b).$$

Furthermore, f is an **isomorphism of lattices**, if is a bijective morphism of lattices.

Theorem 6.1.2 a) *If the ordered set* $(A, \leq)$ *is a lattice, then the operations*

$$a \wedge b = \inf\{a, b\}, \qquad a \vee b = \sup\{a, b\}, \quad \forall\ a, b \in A$$

define on the set A *a lattice structure* $(A, \wedge, \vee)$.

b) *Conversely, if the algebraic structure* $(A, \wedge, \vee)$ *is a lattice, then the relation*

$$a \leq b \Longleftrightarrow a \wedge b = a, \quad \forall\ a, b \in A$$

defined on the set A *is an order relation such that* $(A, \leq)$ *is a lattice; moreover, for any* $a, b \in A$ *we have*

$$a \vee b = \sup\{a, b\}, \quad a \wedge b = \inf\{a, b\}.$$

Proof. a) The commutativity of the operations $\wedge$ and $\vee$ is clear by definition. We prove that $\vee$ is associative: let $x = (a \vee b) \vee c, y = a \vee (b \vee c)$. We have $a \vee b \leq x, c \leq x \Longrightarrow a \leq x, b \leq x, c \leq x \Longrightarrow a \leq x, b \vee c \leq x \Longrightarrow a \vee (b \vee c) \leq x$, of where $y \leq x$. Analogously, we get that $x \leq y$, hence $x = y$.

Let $v = a \vee (a \wedge b)$, so $a \leq v$. On the other hand, $a \wedge b \leq a, a \leq a \Longrightarrow a \vee (a \wedge b) \leq a \Longrightarrow v \leq a \Longrightarrow v = a$.

b) Note that we have

$$(*) \qquad a \vee b = b \Longleftrightarrow a \wedge b = a.$$

Indeed, by absorbtion, we have $a \vee b = b \Longrightarrow a = a \wedge (a \vee b) = a \wedge b$; further, $a \wedge b = a \Longrightarrow b = b \vee (b \wedge a) = b \vee (a \wedge b) = b \vee a = a \vee b$.

Furthermore, observe thatthe two operations are idempotent, that is, we have

$$(**) \qquad a \vee a = a \wedge a = a.$$

Indeed, by absorbtion, for any $a \in A$ we have $a = a \wedge (a \vee a)$, and the $a \vee a = a \vee (a \wedge (a \vee a)) = a$. The dual property is checked analogously.

We show that $\leq$ is an order relation. We have seen that $a \vee a = a$, from where $a \leq a$, hence the relation is reflexive.

Antisimmetry: let $a \leq b, b \leq a$. It follows that $a \vee b = b$, $b \vee a = a \Longrightarrow a = b$.

Transitivity: for any $a, b, c \in A$ we have $a \leq b, b \leq c \Longrightarrow a \vee b = b, b \vee c = c \Longrightarrow a \vee c = a \vee (b \vee c) = (a \vee b) \vee c = b \vee c = c \Longrightarrow a \leq c$.

We show that $a \vee b = \sup\{a, b\}$. Indeed, we may write $a \vee (a \vee b) = (a \vee a) \vee b = a \vee b$, of where $a \leq a \vee b$; analogously we have $b \leq a \vee b$, hence $a \vee b$ is a majorant of a and b. If c is a majorant, that is $a \leq c, b \leq c$, then $a \vee c = c, b \vee c = c \Longrightarrow (a \vee b) \vee c = a \vee (b \vee c) = a \vee c \Longrightarrow a \vee b \leq c$, hence $a \vee b$ is the least majorant.

The equality $a \wedge b = \inf\{a, b\}$ follows from (*). ■

Example 6.1.3 1) $(\mathbb{N}, \wedge, \vee)$ is a lattice with zero element and identity element, where $x \wedge y = (x, y)$, is the greatest common divisor of x and y, while $x \vee y = [x, y]$ is the least common multiple of x and y. The zero elementul is the natural number 1, because $x \vee 1 = [x, 1] = x$ for any x. The identity element is the number natural 0, because $x \wedge 0 = (x, 0) = x$ for any x. This lattice corresponds to the ordered set $(\mathbb{N}, |)$.

2) If M is a set, then $(\mathcal{P}(M), \cap, \cup)$ is a lattice with zero element and unit element. The zero elementul is the empty set $\emptyset$, while the unit element is M. This lattice corresponds to the ordered set $(\mathcal{P}(M), \subseteq)$.

Exercise 76 Prove that:

a) If $f : A \to B$, then $f^* : \mathcal{P}(B) \to \mathcal{P}(A)$, $f^*(Y) = f^{-1}(Y)$ is a morphism of lattices.

b) The function $f_* : \mathcal{P}(A) \to \mathcal{P}(B)$, $f_*(X) = f(X)$ is a morphism of lattices if and only if f is injective.

Exercise 77 Consider the sets $A = \{1, 2, 3\}$ and $B = \{d > 0 \mid d|30\}$. Find all the lattice isomorphisms $f : (\mathcal{P}(A), \subseteq) \to (B, |)$.

Exercise 78 Let $(A, \leq, \wedge, \vee)$ and $(B, \leq, \wedge, \vee)$ be two lattices, and let $f : A \to B$ be a function. Prove that:

a) If f is morphism of lattices, then f is increasing.

b) The converse statement is not true, that is, there exist increasing functions which are not morphisms of lattices.

c) If A is totally ordered, and f is increasing, then f is morphism of lattices.

d) If f is bijective morphism of lattices, then $f^{-1} : B \to A$ is also a morphism of lattices.

e) f is an isomorphism of lattices $\Longleftrightarrow$ f is order isomorphism.

Definition 6.1.4 The lattice $(A, \wedge, \vee)$ is **distributive**, if for any $a, b, c \in A$,

$$(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c).$$

Remark 6.1.5 1) One can show that lattice $(A, \wedge, \vee)$ is distributive if and only if for any $a, b, c \in A$, $(a \wedge b) \vee c = (a \vee c) \wedge (b \vee c)$.

2) The lattices from Examples 6.1.3 of are distributive.

Exercise 79 Prove that:

a) In a lattice $(A, \wedge, \vee)$ we have $a \leq a'$, $b \leq b' \Longrightarrow a \vee b \leq a' \vee b'$ and $a \wedge b \leq a' \wedge b'$.

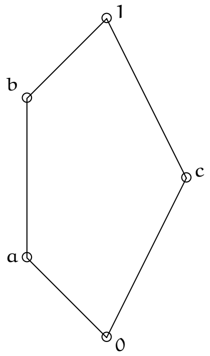
b) The lattice $(A, \wedge, \vee)$ is distributive if and only if for any $a, b, c \in A$ we have $(a \wedge b) \vee c = (a \vee c) \wedge (b \vee c)$.

c) If A is distributive, then for any $a, b, c \in A$ we have

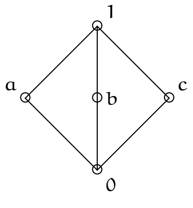
$$a \vee c = b \vee c, \ a \wedge c = b \wedge c \Longrightarrow a = b.$$

d) The following lattices are not distributive:

(1)



(2)



Exercise 80 Prove that:

a) If $(A, \leq)$ is totally ordered, then A is a distributive lattice.

b) $(\mathbb{N}, |)$ is a distributive lattice.

6.2 Boole lattices and Boole rings

Definition 6.2.1 The lattice A is called a **Boole lattice** (or **Boole algebra**), if A is distributive, there exists the least element $0 = \min A$, there exists greatest element $1 = \max A$, and for any $a \in A$ there exists a **complement** $a' \in A$ such that $a \wedge a' = 0$ and $a \vee a' = 1$. We denote this algebraic structure by $(A, \vee, \wedge, 0, 1, ')$.

Example 6.2.2 $(\mathcal{P}(M), \cap, \cup)$ is a Boole lattice, where $\min \mathcal{P}(M) = \emptyset$, $\max \mathcal{P}(M) = M$, while the complement of $X \subseteq M$ is $\complement_M X = M \setminus X$.

Theorem 6.2.3 *If A is a Boole lattice, then:*
a) *For any $a \in A$ there exists a unique complement $a' \in A$ such that $a \wedge a' = 0$ and $a \vee a' = 1$.*
b) $0' = 1, \quad 1' = 0, \quad (a')' = a,$
c) *For any $a, b \in A$,*

$$(a \wedge b)' = a' \vee b', \qquad (a \vee b)' = a' \wedge b'$$

(De Morgan laws).

Proof. a) If $a \vee a' = a \vee \bar{a} = 1$ and $a \wedge a' = a \wedge \bar{a} = 0$, then

$$\begin{aligned} a' &= a' \vee 0 = a' \vee (a \wedge \bar{a}) = (a' \vee a) \wedge (a' \vee \bar{a}) = \\ &= (\bar{a} \vee a) \wedge (a' \vee \bar{a}) = \bar{a} \vee (a \wedge a') = \bar{a} \vee 0 = \bar{a}. \end{aligned}$$

b) We have $0 \vee 1 = 1$, $0 \wedge 1 = 0$, hence $0' = 1$ and $1' = 0$. Furthermore, $a' \wedge a = 0$, $a' \vee a = 1$, hence $(a')' = a$.

c) $(a \vee b) \vee (a' \wedge b') = (a \vee b \vee a') \wedge (a \vee b \vee b') = (1 \vee b) \wedge (1 \vee a) = 1 \wedge 1 = 1$ and $(a \vee b) \wedge (a' \wedge b') = (a \wedge a' \wedge b') \vee (b \wedge a' \wedge b') = 0 \vee 0 = 0$, hence $(a \vee b)' = a' \wedge b'$; analogously one shows that $(a \wedge b)' = a' \vee b'$. ■

Definition 6.2.4 The associative ring with unit $(A, +, \cdot)$ is called a **Boole ring** if $x^2 = x$ for any $x \in A$ (that is, every element of A is idempotent).

Theorem 6.2.5 *If $(A, +, \cdot)$ is a Boole ring, then*
a) $1 + 1 = 0$ *(hence $x + x = 0$ for any $x \in A$).*
b) *A is commutative.*

Proof. a) $1 + 1 = (1 + 1)^2 = 1 + 1 + 1 + 1$, hence $1 + 1 = 0$.
b) If $x, y \in A$, then

$$x + y = (x + y)^2 = x^2 + xy + yx + y^2 = x + y + xy + yx,$$

hence $xy = -yx$; because $1 = -1$, it follows that $xy = yx$. ■

The following theorem discovered by Marshall H. Stone (1903 – 1989) says that the notions of Boole lattice and of Boole ring are equivalent.

Theorem 6.2.6 (Stone) a) *Let $(A, \vee, \wedge, 0, 1, ')$ be a Boole lattice, and define the operations:*

$$\begin{aligned} a + b &= (a \wedge b') \vee (a' \wedge b) = (a \vee b) \wedge (a' \vee b') \\ a \cdot b &= a \wedge b. \end{aligned}$$

Then $(A, +, \cdot)$ is Boole ring with zero element 0 and unit element 1.
b) *Let $(A, +, \cdot, 0, 1)$ be a Boole ring, and define the operations:*

$$a \vee b = a + b + ab, \qquad a \wedge b = ab.$$

Then $(A, \vee, \cdot)$ is Boole lattice, in which $a' = 1 + a$, $\min A = 0$ and $\max A = 1$.
c) *The correspondences defined by* a) *and* b) *are inverses of each other.*
d) *If $f : A \to A'$ is a morphism of Boole lattices, then f is also a morphism of Boole rings, while if $g : B \to B'$ is a morphism of Boole rings, then g is and morphism of Boole lattices.*

Proof. a) Obviously, "+" is commutative. If $a, b, c \in A$, then

$$\begin{aligned} a+(b+c) &= (a\wedge(b+c)')\vee(a'\wedge(b+c)) = \\ &= (a\wedge((b\wedge c')\vee(b'\wedge c'))')\vee(a'\wedge((b\wedge c')\vee(b'\wedge c))) = \\ &= (a\wedge(b\wedge c')'\wedge(b'\wedge c)')\vee(a'\wedge b\wedge c')\vee(a'\wedge b'\wedge c) = \\ &= (a\wedge(b'\vee c)\wedge(b\vee c'))\vee(a'\wedge b\wedge c')\vee(a'\wedge b'\wedge c) = \\ &= (a\wedge b'\wedge c')\vee(a\wedge b\wedge c)\vee(a'\wedge b\wedge c')\vee(a'\wedge b'\wedge c). \end{aligned}$$

It follows that $(a+b)+c = c+(a+b) = a+(b+c)$; further,

$$\begin{aligned} a+0 &= (a\wedge 0')\vee(a'\wedge 0) = a\vee 0 = a \\ a+a &= (a\wedge a')\vee(a'\wedge a) = 0\vee 0 = 0, \end{aligned}$$

hence $-a = a$ for any $a \in A$

The operation "$\cdot$" is commutative and associative, $a\cdot 1 = a\wedge 1 = a$, $a^2 = a\wedge a = a$; we verify the distributivity:

$$\begin{aligned} a(b+c) &= a\wedge((b'\wedge c)\vee(b\wedge c')) = \\ &= (a\wedge b'\wedge c)\vee(a\wedge b\wedge c') \\ ab+ac &= ((a\wedge b)\wedge(a\wedge c)')\vee((a\wedge b)'\wedge(a\wedge c)) = \\ &= (ab\wedge(a'\vee c'))\vee((a'\vee b')\wedge a\wedge c) = \\ &= (a\wedge b\wedge a')\vee(a\wedge b\wedge c')\vee(a'\wedge a\wedge c)\vee(b'\wedge a\wedge c) = \\ &= (a\wedge b\wedge c')\vee(b'\wedge a\wedge c); \end{aligned}$$

it follows that $(A, +, \cdot, 0, 1)$ is a Boole ring.

b) One shows easily that "$\vee$" and "$\wedge$" are commutative and associative, the properties of distributivity and absorbtion hold, and for any $a \in A$, $a \vee 0 = = a + 0 + a\cdot 0 = a$; $a\wedge 1 = a\cdot 1 = a$; $a\wedge(1+a) = a(1+a) = a + a^2 = a + a = 0$ and $a\vee(1+a) = a+1+a+a(1+a) = 1+a+a^2 = 1$.

c) Let $(A, \vee, \wedge, 0, 1, ')$ a Boole lattice, let $(A, +, \cdot, 0, 1)$ be the Boole ring corresponding to it, and let $a\cup b = a+b+ab$, $a\cap b = a\cdot b$, $\bar{a} = a+1$. Then one shows that $a\cup b = a\vee b$, $a\cap b = a\wedge b$ and $a' = \bar{a}$.

Conversely, let $(A, +, \cdot, 0, 1)$ be a Boole ring, let $(A, \vee, \wedge, 0, 1, ')$ be the Boole lattice corresponding to it, and let $a\oplus b = (a\wedge b')\vee(a'\wedge b)$, $a\odot b = a\wedge b$. Then $a\oplus b = a+b$ and $a\odot b = ab$.

d) The statement concerning morphisms is left to the reader. ■

Example 6.2.7 1) $(\mathbb{Z}_2, +, \cdot)$ is a Boole ring, while the corresponding Boole lattice is given by

$\vee$	$\hat{0}$	$\hat{1}$
$\hat{0}$	$\hat{0}$	$\hat{1}$
$\hat{1}$	$\hat{1}$	$\hat{1}$

$\wedge$	$\hat{0}$	$\hat{1}$
$\hat{0}$	$\hat{0}$	$\hat{0}$
$\hat{1}$	$\hat{0}$	$\hat{1}$

$'$	
$\hat{0}$	$\hat{1}$
$\hat{1}$	$\hat{0}$

2) $(\mathcal{P}(M), \cup, \cap, \emptyset, M, \complement)$ is a Boole lattice to which it corresponds the Boole ring $(\mathcal{P}(M), \Delta, \cap)$, where recall that $A\Delta B = (A\setminus B)\cup(B\setminus A)$ is symmetric difference of A and B.

Exercise 81 a) Complete the details of the proof of Stone's Theorem.

b) If A is a Boole lattice, and $a, b \in A$, then

$$a \le b \iff b' \le a' \iff a\wedge b' = 0 \iff a'\vee b = 1.$$

Exercise 82 Using the structure of Boole ring of $\mathcal{P}(U)$, solve the following system of equations, where $A, B, C \in \mathcal{P}(U)$ are given, while $X \in \mathcal{P}(U)$ is the unknown:

a) $A\cap X = B$, $A\cup X = C$.

b) $A\setminus X = B$, $X\setminus A = C$.

6.3 The Lyndenbaum–Tarski algebra

The propositional logic provides us with an important example of Boole lattice.

Definition 6.3.1 a) Let $\mathcal{F}$ be the set of propositional formulas over a given set of atomic formulas. We consider the algebraic structure $(\mathcal{F}, \wedge, \vee, ^-)$. In Chapter 1 we have defined on $\mathcal{F}$ the relations "$\Rightarrow$" (*it follows*) respectively "$\Leftrightarrow$" (*equivalent*). It is clear that $\Rightarrow$ is a preorder, while $\Leftrightarrow = (\Rightarrow \cap \Rightarrow^{-1})$ is an equivalence on $\mathcal{F}$, compatible with the operations $\wedge$, $\vee$ and $^-$.

b) We construct the quotient set $\hat{\mathcal{F}} = \mathcal{F}/\Leftrightarrow$, hence $\hat{\mathcal{F}} = \{\hat{A} \mid A \in \mathcal{F}\}$, where

$$\hat{A} = \{A' \in \mathcal{F} \mid A \Leftrightarrow A'\}.$$

On the set $\hat{\mathcal{F}}$ we define the operations

$$\hat{A} \wedge \hat{B} = \widehat{A \wedge B}, \quad \hat{A} \vee \hat{B} = \widehat{A \vee B}, \quad \bar{\hat{A}} = \hat{\bar{A}}.$$

These definitions do not depend on the choice of representatives.

c) The class of tautologies is denoted by 1, while the class of contradictions by 0. Thus we have

$$1 = \{A \in \mathcal{F} \mid A \text{ tautology}\}, \qquad 0 = \{A \in \mathcal{F} \mid A \text{ contradiction}\}.$$

d) On the quotient set $\hat{\mathcal{F}}$, one can define the following order relation: $\hat{A} \Rightarrow \hat{B}$ if and only if $A \Rightarrow B$.

The proof of the next theorem is left to the reader.

Theorem 6.3.2 *a) The algebraic structure* $(\hat{\mathcal{F}}, \wedge, \vee, {}^{-}, 0, 1)$ *is a Boole lattice.*

b) *The following statements are equivalent:*

(i) $\hat{A} \Rightarrow \hat{B}$; (ii) $\hat{A} \wedge \hat{B} = \hat{A}$; (iii) $\hat{A} \vee \hat{B} = \hat{B}$.

The algebraic structure $(\hat{\mathcal{F}}, \wedge, \vee, {}^{-}, 0, 1)$ is called the *Lyndenbaum–Tarski algebra.* The above theorem gives the possibility to use the methods of algebra in mathematical logic.

Exercise 83 a) Prove that the relations "$\Rightarrow$" and "$\Leftrightarrow$" are compatible with the operations $\wedge$, $\vee$ and ${}^{-}$.

b) Prove Theorem 6.3.2.

6.4 Boole formulas and Boole functions. Normal forms

Let B a finite set.

Definition 6.4.1 a) A **Boole formula (polynomial) over** B is a sequence of symbols constructed as follows:

1. If $x \in B$, then x is a Boole formula;
2. If x, y are Boole formulas, then the following sequences of symbols are Boole formulas:

   $$(x \vee y), \quad (x \wedge y), \quad \text{si} \quad (\bar{x});$$

3. There are no other Boole formulas.

b) If x is a Boole formula, then the **dual** of x (notation: x^*) is obtained by switching the symbols "$\wedge$" and "$\vee$".

c) We may sometimes also use the symbols "$\to$" and "$\leftrightarrow$", but these can be reduced to the above symbols by using the rules known from propositional logic.

Remark 6.4.2 a) If, in addition, we assume that B is a Boole lattice, and if x is a Boole formula over B, then to x corresponds a unique element din B, which is also denoted by x. Since the axioms of the Boole lattice are symmetric, we immediately obtain the **duality principle**:

$(*)$ *If x and y are Boole formulas, and $x = y$ in B, then we have and the equality $x^* = y^*$ in B.*

b) A Boole formula can be transformed into many other equivalent formulas by using the axioms of Boole lattices. However, there are some more important formulas, called **normal forms**.

We first introduce several notations:

- If $\alpha \in V = \{0, 1\}$, let $x^\alpha = \begin{cases} x, & \text{if } \alpha = 1, \\ \bar{x}, & \text{if } \alpha = 0. \end{cases}$
- If $\alpha = (\alpha_1, \ldots, \alpha_n) \in V^n$, then the formulas

  $$x_1^{\alpha_1} \wedge x_2^{\alpha_2} \wedge \cdots \wedge x_n^{\alpha_n} \quad \text{si} \quad x_1^{\alpha_1} \vee x_2^{\alpha_2} \vee \cdots \vee x_n^{\alpha_n}.$$

  are called **elementary conjunctions**, respectively **elementary disjunctions** .

Definition 6.4.3 a) If $c_1, \dots, c_m$ are elementary conjunctions, then the formula $\bigvee_{i=1}^m c_k$ is called a **disjunctive normal form**.

b) If $d_1, \dots, d_m$ elementary disjunctions, then the formula $\bigwedge_{i=1}^m d_k$ is called a **conjunctive normal form**.

It is not difficult to prove that for every Boole formula there is an equivalent disjunctive (respectively conjunctive) normal form. These normal forms are nor unique.

Example 6.4.4 *We consider the formula* $\bar{x}_1 \to (x_1 \wedge x_2)$*, and we transform it into a disjunctive (respectively conjunctive) normal form:*

$$\begin{aligned}\bar{x}_1 \to (x_1 \wedge x_2) &= \bar{\bar{x}}_1 \vee (x_1 \wedge x_2) = x_1 \vee (x_1 \wedge x_2) = x_1 = \\ &= (x_1 \vee x_1) \wedge (x_1 \vee x_2) = x_1 \wedge (x_1 \vee x_2).\end{aligned}$$

Definition 6.4.5 a) We consider the Boole lattice $B = V = \{0, 1\}$. If $x = x(x_1, \dots, x_n)$ is a Boole formula, then assigning values $x_i \in V$, to the formula x corresponds a unique function $x : V^n \to V$. A function obtained in this way is called **function Boole**.

b) Let $f : V^n \to V$ be a function, and define the **truth sets** T_f (*true*) and F_f (*false*) as follows:

$$\begin{aligned}T_f &= \{\alpha = (\alpha_1, \dots, \alpha_n) \in V^n \mid f(\alpha_1, \dots, \alpha_n) = 1\},\\ F_f &= \{\alpha = (\alpha_1, \dots, \alpha_n) \in V^n \mid f(\alpha_1, \dots, \alpha_n) = 0\}.\end{aligned}$$

In what follows, we show that every formula Boole has certain special disjunctive or conjunctiva normal form, called *perfect*. We also deduce that every function $f : V^n \to V$ is a Boole function.

Theorem 6.4.6 *Let* $f : V^n \to V$ *be a Boole function.*

1) *If* $T_f \neq \emptyset$*, then*

$$f(x_1, \dots, x_n) = \bigvee_{\alpha \in T_f} \bigwedge_{i=1}^n x_i^{\alpha_i}.$$

2) *If* $F_f \neq \emptyset$*, then*

$$f(x_1, \dots, x_n) = \bigwedge_{\alpha \in F_f} \bigvee_{i=1}^n x_i^{\bar{\alpha}_i}.$$

Proof. 1) If $(\alpha_1, \dots, \alpha_n) \in T_f$, then $f(\alpha_1, \dots, \alpha_n) = 1$ and $\bigwedge_{i=1}^n \alpha_i^{\alpha_i} = 1$; if $(\beta_1, \dots, \beta_n) \neq (\alpha_1, \dots, \alpha_n)$, then $\bigwedge_{i=1}^n \beta_i^{\alpha_i} = 1$, because $\beta_i^{\alpha_i} = 0$ if $\beta_i \neq \alpha_i$; it follows that $\bigvee_{\alpha \in T_f} \bigwedge_{i=1}^n x_i^{\alpha_i} = 1$.

Conversely, if $\bigvee_{\alpha \in T_f} \bigwedge_{i=1}^n x_i^{\alpha_i} = 1$, then there exists $\alpha \in T_f$, such that $\bigwedge_{i=1}^n x_i^{\alpha_i} = 1$, hence $x_i^{\alpha_i} = 1$ for any $i = 1, \dots, n$. It follows that $x_i = \alpha_i$, $i = 1, \dots, n$, hence $(x_1, \dots, x_n) = (\alpha_1, \dots, \alpha_n) \in T_f$ and $f(x_1, \dots, x_n) = 1$.

2) is proven analogously. ■

Definition 6.4.7 The formula from 1) (respectively 2)) is called the **perfect disjunctive normal form (PDNF)** (respectively **perfect conjunctive normal (CNFP)**) of f.

Note that the constant function 0 does not have PDNF, while the constant function 1 does not have PCNF.

Example 6.4.8 Let $f(x_1, x_2) = x_1 \to x_2$; then we have $T_f = \{(0, 0), (0, 1), (1, 1)\}$ and $F_f = \{(1, 0)\}$, hence

$$f(x_1, x_2) = (x_1^0 \wedge x_2^0) \vee (x_1^0 \wedge x_2^1) \vee (x_1^1 \wedge x_2^1) = (\bar{x}_1 \wedge \bar{x}_2) \vee (\bar{x}_1 \wedge x_2) \vee (x_1 \wedge x_2); \qquad \text{(PDNF)}$$

$$f(x_1, x_2) = x_1^{\bar{1}} \vee x_2^{\bar{0}} = \bar{x}_1 \vee x_2. \qquad \text{(PCNF)}$$

Exercise 84 Prove that $f(x_1, x_2) = \overline{x_1 \wedge x_2} \to (\bar{x}_1 \vee \bar{x}_2)$ is equal to the constant function 1, using:

a) truth tables ; b) Boole rings.

Exercise 85 Find the PDNF and the PCNF for $f(x_1, x_2, x_3) = \bar{x}_1 \to (x_2 \wedge \bar{x}_3)$.

Exercise 86 Let $f : V^3 \to V$ such that $T_f = \{(1, 1, 1), (1, 1, 0), (1, 0, 1), (1, 0, 0)\}$.

a) Find the PDNF and the PCNF for $f(x_1, x_2, x_3)$.

b) Prove that $f(x_1, x_2, x_3) = x_1$.

Chapter 7

NUMBER SETS

7.1 The set of natural numbers

7.1.1 Peano Axioms

Definition 7.1.1 The axiom of infinity 3.1.2 states that there exists a set y such that $\emptyset \in y$ and $x \in y$, $x^+ \in y$, where $x^+ = x \cup \{x\}$.

Denote by $\mathcal{A}$ the class of sets satisfying the above property, that is,

$$\mathcal{A} = \{A \mid \emptyset \in A;\ \text{if } x \in A, \text{ then } x^+ \in A\}.$$

We call $\mathcal{A}$ the class of **inductive sets**.

Denote by $\mathbb{N} := \bigcap \mathcal{A}$ the intersection of all inductive sets. Then $\mathbb{N}$ is a set, which is called **the set of natural numbers**. Notations: $0 := \emptyset$, $1 := 0^+ = \{0\}$, $2 := 1^+ = \{0, 1\}$, $3 := 2^+ = \{0, 1, 2\}, \dots$. We also denote $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$.

The element $s(n) = n^+$ is called **the successor** of n.

Theorem 7.1.2 (Peano Axioms) *The triple formed by the set of natural numbers* $\mathbb{N}$*, the element* 0*, and the successor function* $s : \mathbb{N} \to \mathbb{N}$ *satisfies the* **Peano Axioms***:*

1) $0 \in \mathbb{N}$.

2) *If* $n \in \mathbb{N}$*, then* $n^+ \in \mathbb{N}$ *(that is,* s *is well defined).*

3) **(The principle of mathematical induction)** *If* $S \subseteq \mathbb{N}$*,* $0 \in S$ *and* $n \in S$*,* $n^+ \in S$*, then* $S = \mathbb{N}$ *(that is, every inductive subset of* $\mathbb{N}$ *coincides with* $\mathbb{N}$*).*

4) *If* $n \in \mathbb{N}$*, then* $n^+ \neq 0$*.*

5) *If* $n, m \in \mathbb{N}$ *and* $n^+ = m^+$*, then* $n = m$*.*

Remark 7.1.3 a) Note that $n \neq n^+$, because the *axiom of regularity* excludes the anomaly $n \in n$.

b) By the principle of mathematical induction we easily see that every nonzero natural number is the successor of a natural number, that is $\forall n \in \mathbb{N}^*$, $\exists m \in \mathbb{N}$ such that $n = m^+$.

c) The axioms 2), 4) and 5), together with the above remark say that **the successor function**

$$s : \mathbb{N} \to \mathbb{N}, \qquad s(n) = n^+$$

is well defined, is injective, but is not surjective, because we have $\operatorname{Im} s = \mathbb{N}^*$.

The following theorem gives the possibility of **recursive (inductive) definitions**.

Theorem 7.1.4 (The recurrence theorem) *Let* X *be a set,* $a \in X$ *a fixed element, and* $f : X \to X$ *a function. Then there exists a unique function* $u : \mathbb{N} \to X$ *such that* $u(0) = a$*, and* $u(n^+) = f(u(n))$ *for any* $n \in \mathbb{N}$*.*

The next corollary says that the Peano axioms determine the triple $(\mathbb{N}, 0, s)$ uniquely, up to a unique isomorphism.

Corollary 7.1.5 *If the triple* $(N', 0', s')$ *satisfies the Peano axioms, then it is isomorphic to the triple* $(\mathbb{N}, 0, s)$*, that is there exists a unique function* $u : \mathbb{N} \to N'$ *which satisfies the following properties:*

(1) $u(0) = 0'$, (2) $u \circ s = s' \circ u$, (3) u *is bijective.*

7.1.2 Operations, and the relation of order on the set of natural numbers

Definition 7.1.6 (operations with natural numbers) a) The addition of natural numbers is defined inductively by:

$$m + 0 = m, \qquad m + s(n) = s(m + n).$$

Remark that $s(n) = n^+ = n + 1$.

b) The multiplication of natural numbers is defined inductively by:

$$m \cdot 0 = 0, \qquad ms(n) = mn + m.$$

Observe that $n \cdot 1 = n$.

Theorem 7.1.7 (the basic properties of the operations) *If* $m, n, p \in \mathbb{N}$, *then*

1) $(m + n) + p = m + (n + p)$;
2) $m + 0 = 0 + m$;
3) $m + 1 = 1 + m$;
4) $m + n = n + m$;
5) *If* $m + p = n + p$, *then* $m = n$. *In particular, if* $m + p = m$, *then* $p = 0$.
6) *If* $m + n = 0$, *then* $m = n = 0$;
7) **(Trichotomy)** *Of the following three statements, exactly one is true:*
(i) $m = n$, (ii) $\exists p \in \mathbb{N}^*$ *such that* $m = n + p$, (iii) $\exists p \in \mathbb{N}^*$ *such that* $n = m + p$;
8) $(m + n)p = mp + np$; $p(m + n) = pm + pn$;
9) $m(np) = (mn)p$;
10) $0 \cdot m = 0$;
11) $1 \cdot m = m$;
12) $mn = nm$;
13) *If* $mn = 0$, *then* $m = 0$ *or* $n = 0$;
14) *If* $mp = np$ *and* $p \neq 0$, *then* $m = n$;
15) *If* $mn = 1$, *then* $m = n = 1$.

Exercise 87 Prove Theorem 7.1.7.

Definition 7.1.8 (ordering natural numbers) Let $m, n \in \mathbb{N}$. We say that m **is less than** n, notation $m < n$, if there exists $p \in \mathbb{N}^*$ such that $m + p = n$. If $m = n$ or $m < n$, then we say that m **is less than or equal to** n, and we denote $m \leq n$.

Theorem 7.1.9 (the basic properties of the order relation) *Let* $m, n, p \in \mathbb{N}$. *Then:*

1) *"$\leq$" is a total order;*
2) $0 \leq n$;
3) *If* $n \neq 0$, *then* $1 \leq n$;
4) $m < n$ *if and only if* $m^+ \leq n$;
5) $m \leq n$ *if and only if* $m < n^+$;
6) *Nu there exists* $n \in \mathbb{N}$ *such that* $m < n < m^+$;
7) $(\mathbb{N}, \leq)$ *is well-ordered;*
8) **(The principle of mathematical induction, 2nd variant: complete or strong induction)** *If* $P(n)$ *is a predicate on the set of natural numbers such that* $P(0)$ *is true, and if* $P(k)$ *true for any* $k < n$, *then* $P(n)$ *is true;*
9) *If* $m < n$, *then* $m + p < n + p$;
10) *If* $m < n$ *and* $p \neq 0$, *then* $mp < np$
11) **(Archimedes axiom)** *If* $m \in \mathbb{N}$ *and* $n \in \mathbb{N}^*$, *then there exists* $p \in \mathbb{N}$ *such that* $pn > m$;
12) **(The division theorem)** *If* $m \in \mathbb{N}$ *and* $n \in \mathbb{N}^*$, *then there exist unique* $q, r \in \mathbb{N}$ *such that* $m = nq + r$ *and* $r < n$.

Proof. 7) We assume that $(\mathbb{N}, \leq)$ is not well-ordered, that is, there exists a subset $A \neq \emptyset$ which does not have a least element. Let S be the set of strict minorants of A, that is,

$$S = \{n \in \mathbb{N} \mid n < a \ \forall a \in A\}.$$

It is clear that $0 \in S$, because A does not have a least element. If $n \in S$, then $n^+ \leq a$ for any $a \in A$. But $n^+ \notin A$ (because otherwise it would be the least element of A), hence $n^+ < a$ for any $a \in A$, that is $n^+ \in S$. By induction, it follows that $S = \mathbb{N}$, hence $A = \emptyset$, contradiction. ■

Exercise 88 Prove Theorem 7.1.9.

7.2 The set of integers

Theorems 7.1.7 and 7.1.9 say that the structure $(\mathbb{N}, +, \cdot, \leq)$ is an associative, commutative semiring with unit, without divisors of zero, well-ordered and Archimedean. One of the problems is that $(\mathbb{N}, +)$ is not a group. We solve this by enlarging the set $\mathbb{N}$. We will construct below the set of integers starting with the set of natural numbers, an then we define the addition, the multiplication and the ordering of integers.

Definition 7.2.1 a) On the set $\mathbb{N} \times \mathbb{N}$ we define the homogeneous relation:

$$(m, n) \sim (p, q) \text{ if } m + q = n + p,$$

which is easily seen to be an equivalence. We denote by $\widetilde{(m, n)}$ the equivalence class of the pair (m, n), so

$$\widetilde{(m, n)} = \{(p, q) \in \mathbb{N} \times \mathbb{N} \mid (p, q) \sim (m, n)\}.$$

The quotient set

$$\mathbb{Z} := \mathbb{N} \times \mathbb{N}/ \sim = \{\widetilde{(m, n)} \mid m, n \in \mathbb{N}\}$$

is called the **set of integers**, while the classes $\widetilde{(m, n)}$ are called **integers**.

b) The addition and the multiplication of integers are defined as follows:

$$\widetilde{(m, n)} + \widetilde{(p, q)} := \widetilde{(m + p, n + q)}, \qquad \widetilde{(m, n)}\widetilde{(p, q)} := \widetilde{(mp + nq, np + mq)}.$$

These definitions do not depend on the choice of representatives.

Theorem 7.2.2 $(\mathbb{Z}, +, \cdot)$ *is an integral domain, in which the zero element is* $0 := \widetilde{(0, 0)} = \{(m, m) \mid m \in \mathbb{N}\}$*, the unit element is* $1 := \widetilde{(1, 0)} = \{(m + 1, m) \mid m \in \mathbb{N}\}$*, and the opposite of the integer* $\widetilde{(m, n)}$ *is* $-\widetilde{(m, n)} := \widetilde{(n, m)}$*.*

Exercise 89 a) Prove that the relation "$\sim$" is an equivalence.

b) Prove that the definitions of the addition and of the multiplication do not depend on the choice of representatives.

c) Prove Theorem 7.2.2.

Definition 7.2.3 The integers are ordered by the relation:

$$\widetilde{(m, n)} < \widetilde{(p, q)} \text{ if and only if } q + m < p + n.$$

Theorem 7.2.4 1) *The definition of the relation "$\leq$" do not depend on the choice of representatives.*

2) $(\mathbb{Z}, \leq)$ *is totally ordered.*

3) *The function* $\alpha : \mathbb{N} \to \mathbb{Z}_+$, $\alpha(n) = \widetilde{(n, 0)}$ *is well defined, strictly increasing, and it is an isomorphism of semirings.*

4) *The ordering of integers is compatible with the addition and the multiplication, that is, for any* $a, b, c, d \in \mathbb{Z}$ *we have*

$$a < b,\ c \leq d \Rightarrow a + c < b + d, \qquad a < b,\ c > 0 \Rightarrow ac < bc, \qquad a < b,\ c < 0 \Rightarrow ac > bc.$$

5) **(Archimedes axiom)** *For any* $a \in \mathbb{Z}_+^*$ *and* $b \in \mathbb{Z}$*, there exists* $n \in \mathbb{N}$ *such that* $na > b$*.*

Remark 7.2.5 We will identify: n with $\alpha(n) = \widetilde{(n, 0)}$, $\mathbb{N}$ with $\mathbb{Z}_+$, where

$$\mathbb{Z}_+ = \{a \in \mathbb{Z} \mid a \geq 0\} = \{\widetilde{(m, n)} \mid m \geq n\}.$$

Taking this into account, for any $m, n \in \mathbb{N}$ we have

$$m - n = m + (-n) = \widetilde{(m, 0)} + (-\widetilde{(n, 0)}) = \widetilde{(m, 0)} + \widetilde{(0, n)} = \widetilde{(m, n)}.$$

Exercise 90 Prove Theorem 7.2.4.

7.3 The set of rational numbers

We have seen that $(\mathbb{Z},+,\cdot,\leq)$ is a totally ordered Archimedean integral domain. We extend this structure to obtain a totally ordered field.

Definition 7.3.1 a) On the set $\mathbb{Z}\times\mathbb{Z}^*$ we define the homogeneous relation

$$(a,b)\sim(c,d),\quad \text{if }\ ad=bc,$$

which is easily seen to be an equivalence. We denote by $\widetilde{(a,b)}$ the equivalence class of the pair (a,b), hence

$$\widetilde{(a,b)}=\{(c,d)\in\mathbb{Z}\times\mathbb{Z}^*\mid(c,d)\sim(a,b)\}.$$

The quotient set

$$\mathbb{Q}:=\mathbb{Z}\times\mathbb{Z}^*/\sim=\{\widetilde{(a,b)}\mid(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\}$$

is called the **set of rational numbers**. A rational number is usually denoted on the form of a **fraction**:

$$\widetilde{(a,b)}=\frac{a}{b}.$$

Note that for $a\in\mathbb{Z}$ and $b,c\in\mathbb{Z}^*$, we have $\frac{a}{b}=\frac{ac}{bc}$. In particular, $\frac{a}{b}=\frac{-a}{-b}$, hence may always choose a representative with positive denominator, that is, we may assume that $b\in\mathbb{N}^*$.

b) The addition and the multiplication of rational numbers are defined as follows:

$$\widetilde{(a,b)}+\widetilde{(c,d)}:=\widetilde{(ad+bc,bd)},\qquad\widetilde{(a,b)}\widetilde{(c,d)}:=\widetilde{(ac,bd)},$$

that is,

$$\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd},\qquad\frac{a}{b}\frac{c}{d}=\frac{ac}{bd}.$$

These definitions do not depend on the choice of representatives.

Theorem 7.3.2 *The algebraic structure $(\mathbb{Q},+,\cdot)$ is a field, in which the zero element is*

$$0:=\widetilde{(0,1)}=\{(0,b)\mid b\in\mathbb{Z}^*\}=\frac{0}{a},\quad a\in\mathbb{Z}^*,$$

the unit element is

$$1:=\widetilde{(1,1)}=\{(a,a)\mid a\in\mathbb{Z}^*\}=\frac{a}{a},\quad a\in\mathbb{Z}^*,$$

the opposite of the rational number $\widetilde{(a,b)}$ is

$$-\widetilde{(a,b)}:=\widetilde{(-a,b)}=\widetilde{(a,-b)},\qquad\textit{that is}\ -\frac{a}{b}=\frac{-a}{b}=\frac{a}{-b},$$

and if $a,b\in\mathbb{Z}^*$, then the inverse of $\widetilde{(a,b)}$ is

$$\widetilde{(a,b)}^{-1}=\widetilde{(b,a)},\qquad\textit{that is}\ \left(\frac{a}{b}\right)^{-1}=\frac{b}{a}.$$

Exercise 91 a) Prove that the relation "$\sim$" is an equivalence.

b) Prove that the definitions of the addition and of the multiplication do not depend on the choice of representatives.

c) Prove Theorem 7.3.2.

Definition 7.3.3 The rational numbers are ordered by the relation:

$$\frac{a}{b}\leq\frac{c}{d},\quad\text{if }\ (bc-ad)bd\geq 0.$$

b) The **absolute value (modulus)** of the rational number $a\in\mathbb{Q}$ is $|a|:=\begin{cases}a, & \text{if } a\geq 0,\\ -a, & \text{if } a<0\end{cases}.$

The next theorem says that the structure $(\mathbb{Q}, +, \cdot, \leq)$ is a **totally ordered Archimedean field**, in which the totally ordered integral domain $(\mathbb{Z}, +, \cdot, \leq)$ of integers is embedded.

Theorem 7.3.4 1) *The definition of the relation "$\leq$" does not depend on the choice of representatives.*

2) $(\mathbb{Q}, \leq)$ *is totally ordered.*

3) *The function* $\alpha : \mathbb{Z} \to \mathbb{Q}$, $\alpha(a) = \widetilde{(a, 1)} = \frac{a}{1}$ *is well defined, strictly increasing (hence injective), and it is a unital morphism of rings.*

4) *The ordering of rational numbers is compatible with the addition and the multiplication, that is, if* $x, y, z, t \in \mathbb{Q}$, *then*

$$x < y, z \leq t \Rightarrow x + z < y + t, \qquad x < y, z > 0 \Rightarrow xz < yz, \qquad x < y, z < 0 \Rightarrow xz > yz.$$

5) **(Archimedes axiom)** $\forall x \in \mathbb{Q}_+^*$, $\forall y \in \mathbb{Q}$, $\exists n \in \mathbb{N}$ *such that* $nx > y$.

Remark 7.3.5 We will identify the integer a with the rational number $\alpha(a) = \frac{a}{1}$, so in this way, $\mathbb{Z} \subset \mathbb{Q}$. Note that $\frac{a}{b} = \alpha(a)\alpha(b)^{-1}$.

Exercise 92 Prove Theorem 7.3.4.

Exercise 93 Prove that for any $x, y \in \mathbb{Q}$

$$|x| = |-x|, \qquad |xy| = |x||y|, \qquad |x + y| \leq |x| + |y|, \qquad ||x| - |y|| \leq |x - y|.$$

Chapter 8

CARDINAL NUMBERS

The results presented in this chapter have been discovered by the German mathematician Georg Cantor (1845 – 1918). He is the creator of set theory, and has shown the importance of bijective functions. Cantor has introduced the definition of infinite sets and of well-ordered sets, and has shown that there exists a "hierarchy" a of infinite sets. He has also introduced cardinal numbers and ordinal numbers, and has studied their arithmetic.

8.1 Cardinal number. Operations with cardinal numbers

Definition 8.1.1 We say that the sets A and B are **equipotent** (notation: $A \sim B$), if there exists a bijective function $f : A \to B$.

Remark 8.1.2 "$\sim$" is an equivalence on the class of sets, hence we get a partition of this class.

Indeed, if A a set, then $1_A : A \to A$ is bijective, hence "$\sim$" is reflexive. If $A \sim B$ and $B \sim C$, then there exists the bijective functions $f : A \to B$ and $g : B \to C$. Since $g \circ f : A \to C$ is bijective, it follows that $A \sim C$, hence "$\sim$" is transitive. If $f : A \to B$ is bijective, then and $f^{-1} : B \to A$ is bijective, hence "$\sim$" is symmetric.

Definition 8.1.3 a) The **cardinal** of the set A is the equipotence class of A. Notation: $|A|$, hence $A \sim B$ if and only if $|A| = |B|$, and we say that the set A is a **representative** of the number cardinal $\alpha = |A|$.

(Note that this is the 'naive' approach, the precise axiomatic definition of $|A|$ is difficult, and requires the introduction of ordinal numbers.)

b) The **addition**, the *multiplication*, and the *exponentiation* of cardinal numbers are defined as follows:

1. $\sum_{i\in I} \alpha_i = |\coprod_{i\in I} A_i|$;
2. $\prod_{i\in I} \alpha_i = |\prod_{i\in I} A_i|$;
3. $\beta^\alpha = |B^A| = |\mathrm{Hom}(A, B)|$.

Remark 8.1.4 The above definitions do not depend on the choice of representatives. Indeed, if $\alpha_i = |A_i| = |A_i'|$, and $f_i : A \to A_i'$ is bijective for any $i \in I$, then $\coprod_{i\in I} f_i : \coprod A_i \to \coprod_{i\in I} A_i'$ and $\prod_{i\in I} f_i : \prod_{i\in I} A_i \to \prod_{i\in I} A_i'$ are bijective. If $f : A' \to A$ and $g : B \to B'$ are bijective, then $\mathrm{Hom}(f, g) : \mathrm{Hom}(A, B) \to \mathrm{Hom}(A', B')$ is also bijective.

Theorem 8.1.5 *Let* $(A_i)_{i\in I}$ *be a family of sets.*

a) $\phi : \coprod_{i\in I} A_i \to \bigcup_{i\in I} A_i$, $\phi(a_i, i) = a_i$ *is surjective, and* ϕ *is injective if and only if* $A_i \cap A_j = \emptyset$ *for any* $i, j \in I$, $i \neq j$.

b) $|A_1 \cup A_2| + |A_1 \cap A_2| = |A_1| + |A_2|$.

Proof. a) If $a \in \bigcup_{i\in I} A_i$, then there exists $i \in I$ such that $a \in A_i$ and $\phi(a, i) = a$, hence ϕ is surjective.

We assume that ϕ is injective, and that there exists $i, j \in I$ and $a \in A_i \cap A_j$. Since $\phi(a, i) = \phi(a, j) = a$, it follows that $(a, i) = (a, j)$, hence $i = j$.

Conversely, assume that for any $i \neq j$, $A_i \cap A_j = \emptyset$, and let $(a_i, i), (a_j, j) \in \coprod_{i\in I} A_i$ such that $\phi(a_i, i) = \phi(a_j, j)$; it follows that $a_i = a_j \in A_i \cap A_j$, hence $i = j$ and $(a_i, i) = (a_j, j)$.

b) If $A_1 \cap A_2 = \emptyset$, then by a) it follows that $A_1 \cup A_2 \sim A_1 \coprod A_2$, hence $|A_1 \cup A_2| = |A_1| + |A_2|$.

In general, $A_1 \cup A_2 = A_1 \cup (A_2 \setminus A_1)$ and $A_2 = (A_2 \setminus A_1) \cup (A_1 \cap A_2)$, where $A_1 \cap (A_2 \setminus A_1) = \emptyset$ and $(A_2 \setminus A_1) \cap A_1 \cap A_2) = \emptyset$; it follows ca

$$|A_1 \cup A_2| + |A_1 \cap A_2| = |A_1| + |A_1 \setminus A_2| + |A_1 \cap A_2| = |A_1| + |A_2|.$$

Theorem 8.1.6 *The following identities involving cardinal numbers hold:*
a) $\alpha_1 + \alpha_2 = \alpha_2 + \alpha_1$*;* $\alpha_1\alpha_2 = \alpha_2\alpha_1$*;*
b) $(\alpha_1 + \alpha_2) + \alpha_3 = \alpha_1 + (\alpha_2 + \alpha_3)$*;* $(\alpha_1\alpha_2)\alpha_3 = \alpha_1(\alpha_2\alpha_3)$*;*
c) $(\sum_{i\in I} \alpha_i)(\sum_{j\in J} \beta_j) = \sum_{(i,j)\in I\times J} \alpha_i\beta_j$*;*
d) $\beta^{\sum_{i\in I} \alpha_i} = \prod_{i\in I} \beta^{\alpha_i}$*;*
e) $(\prod_{i\in I} \alpha_i)^\beta = \prod_{i\in I} \alpha_i^\beta$*;*
f) $\gamma^{\alpha\beta} = (\gamma^\beta)^\alpha$.

Proof. a) Let $\alpha_1 = |A_1|$, $\alpha_2 = |A_2|$ and remark that the functions

$$\phi : A_1 \coprod A_2 \to A_2 \coprod A_1, \quad \phi(a_1, 1) = (a_1, 2), \ \phi(a_2, 2) = (a_2, 1),$$
$$\psi : A_1 \times A_2 \to A_2 \times A_1, \quad \psi(a_1, a_2) = (a_2, a_1)$$

are bijective.

b) Note that the sets $(A_1 \coprod A_2) \coprod A_3 = \{((a_1, 1), 1'),((a_2, 2), 1'),(a_3, 2') \mid a_i \in A_i\}$ and $A_1 \coprod (A_2 \coprod A_3) = \{(a_1, 1'),((a_2, 1), 2'),((a_3, 2), 2') \mid a_i \in A_i\}$ are equipotent.

c) If $\alpha_i = |A_i|$ and $\beta_j = |B_j|$, then

$$\phi : (\coprod_{i\in I} A_i) \times (\coprod_{j\in J} B_j) \to \coprod_{(i,j)\in I\times J} (A_i \times B_j), \qquad ((a_i, i),(b_j, j)) \mapsto ((a_i, b_i),(i, j))$$

is bijective.

d) Let $\alpha_i = |A_i|$, $i \in I$ and $\beta = |B|$. Then

$$\phi : \mathrm{Hom}(\coprod_{i\in I} A_i, B) \to \prod_{i\in I} \mathrm{Hom}(A_i, B), \qquad \phi(\alpha) = (\alpha \circ q_i)_{i\in I}$$

is bijective, where $q_i : A_i \to \coprod_{i\in I} A_i$ is the canonical injection of the direct sum.

e) With the above notations, we have that the function

$$\psi : \mathrm{Hom}(B, \prod_{i\in I} A_i) \to \prod_{i\in I} \mathrm{Hom}(B, A_i), \qquad \psi(\alpha) = (p_i \circ \alpha)_{i\in I}$$

is bijective, where $p_i : \prod_{i\in I} A_i \to A_i$ is the canonical projection of the direct product.

f) Let $\alpha = |A|$, $\beta = |B|$, $\gamma = |C|$, and consider the functions

$$\phi : \mathrm{Hom}(A \times B, C) \to \mathrm{Hom}(A, \mathrm{Hom}(B, C)), \qquad \phi(f)(a)(b) = f(a, b),$$
$$\psi : \mathrm{Hom}(A, \mathrm{Hom}(B, C)) \to \mathrm{Hom}(A \times B, C), \qquad \psi(g)(a, b) = g(a)(b),$$

where $a \in A$ and $b \in B$. One easily shows that $\psi = \phi^{-1}$. ■

Theorem 8.1.7 (Cantor) *For any set* A *we have* $|\mathcal{P}(A)| = 2^{|A|}$.

Proof. Let $\varphi_A : \mathcal{P}(A) \to \mathrm{Hom}(A, \{0, 1\})$, $\varphi_A(X) = \chi_X$, where

$$\chi_X : A \to \{0, 1\}, \qquad \chi_X(a) = \begin{cases} 1, & \text{if } a \in X \\ 0, & \text{if } a \notin X \end{cases}$$

is the **the characteristic function** of the subset X. Note that φ_A is bijective, because

$$\varphi_A^{-1}(\chi) = \chi^{-1}(1), \qquad \forall\, \chi : A \to \{0, 1\}$$

is the inverse of φ_A. ■

8.2 Ordering cardinal numbers

Definition 8.2.1 Let $\alpha = |A|$ and $\beta = |B|$ be cardinal numbers. We say that $\alpha \le \beta$ if there exists an injective function $\phi : A \to B$.

The definition does not depend on the choice of representatives. Indeed, if $\alpha = |A| = |A'|$, $f : A' \to A$ are bijective, $\beta = |B| = |B'|$, $g : B \to B'$ is bijective, then $\mathrm{Hom}(f, g)(\phi) = g \circ \phi \circ f : A' \to B'$ is injective.

Exercise 94 If $\alpha_i \leq \beta_i$, $\forall\ i \in I$, then:

a) $\sum_{i\in I} \alpha_i \leq \sum_{i\in I} \beta_i$;

b) $\prod_{i\in I} \alpha_i \leq \prod_{i\in I} \beta_i$;

c) if $0 \neq \alpha \leq \alpha'$ and $\beta \leq \beta'$, then $\beta^\alpha \leq \beta'^{\alpha'}$.

To show that "$\leq$" is an order relation, we need the following important lemma.

Lemma 8.2.2 (Cantor–Bernstein–Schröder) *If $A_2 \subseteq A_1 \subseteq A_0$ and $A_0 \sim A_2$, then $A_0 \sim A_1$.*

Theorem 8.2.3 *The relation "$\leq$" is a total order. (In particular, trichotomy: if α and β are cardinal numbers, then either $\alpha < \beta$ $\alpha = \beta$, or $\alpha > \beta$.)*

Theorem 8.2.4 (Cantor) *For any cardinal number α, we have $\alpha < 2^\alpha$.*

Proof. Let $\alpha = |A|$. Since $A \to \mathcal{P}(A)$, $a \mapsto \{a\}$ is injective, it follows that $\alpha \leq 2^\alpha$. We assume that $\phi : A \to \mathcal{P}(A)$ is bijective, and let

$$X = \{a \in A \mid a \notin \phi(a)\}.$$

Then there exists $x \in A$ such that $\phi(x) = X$. If $x \in X$, then $x \in \phi(x)$, hence $x \notin X$; if $x \notin X$, then $x \notin \phi(x)$, hence $x \in X$. In both cases we get a contradiction, hence $\alpha < 2^\alpha$. ■

8.3 Finite, infinite and countable sets

Definition 8.3.1 *Let A be a set.*

a) We say that A is **finite**, if it is equipotent to a natural number, that is $\exists n \in \mathbb{N}$ such that $A \sim n$. The set A is **infinite**, if it is not finite.

b) We say that A **infinite countable**, if is equipotent to the set of natural numbers, that is $A \sim \mathbb{N}$. We denote by $\aleph_0$ the cardinal $\mathbb{N}$.

c) We say that A is **countable** (or **at most countable**), if is finite or infinite countable.

d) We denote by $\mathfrak{c}$ the cardinal of the set $\mathbb{R}$ of real numbers, and we also say that $\mathfrak{c}$ is the **power of the continuum**.

Theorem 8.3.2 a) *Any subset of a finite set is finite.*

b) *Any finite set is equipotent to a unique natural number.*

c) *The addition of natural numbers defined in the previous chapter coincides with their addition as cardinal numbers.*

Theorem 8.3.3 *Let A be a set. The following statements are equivalent:*

(1) *A is infinite;*

(2) *There exists an injective function $f : \mathbb{N} \to A$;*

(3) *A has a proper subset equipotent to A.*

Corollary 8.3.4 a) *The set $\mathbb{N}$ a of natural numbers is infinite, moreover, $|\mathbb{N}| =: \aleph_0$ is the least numar cardinal infinit (or* **transfinit***).*

b) *Let A a set. The following statements are equivalent:*

(1) *A is finite;*

(2) *If $f : \mathbb{N} \to A$, then f nu is injectiv;*

(3) *If $B \subseteq A$ and $|B| = |A|$, then $B = A$.*

c) *The union of two finite sets is finite.*

Exercise 95 Let A a set. Prove that the following statements are equivalent:

(i) A is finite set;

(ii) If $f : A \to A$ is injective, then f is surjective;

(iii) If $f : A \to A$ is surjective, then f is injective.

Exercise 96 Let A an infinite set. Prove that:

a) $|A| + n = |A|$, $\forall\ n \in \mathbb{N}$;

b) $|A| + \aleph_0 = |A|$.

Exercise 97 Prove that:

a) $\aleph_0 + \aleph_0 = \aleph_0$; $\aleph_0 \cdot \aleph_0 = \aleph_0$;

b) If $A_n \sim \mathbb{N}$ for any $n \in \mathbb{N}$, then $\bigcup_{n\in\ \mathbb{N}} A_n \sim \mathbb{N}$ (that is, a countable union of countable sets is countable);

c) The set $\mathcal{P}_f(\mathbb{N}) = \{X \subset \mathbb{N} \mid X \text{ is finite}\}$ of finite subsets $\mathbb{N}$ is countable;

d) The set of rational numbers is countable;

e) The set $\mathbb{Q}[X]$ of polynomials with rational coefficients is countable;

f) The set $\mathbb{A} := \{z \in \mathbb{C} \mid (\exists)P \in \mathbb{Q}[X] \setminus \{0\},\ P(z) = 0\}$ a **algebraic numbers** is countable.

Theorem 8.3.5 a) *The set* $\mathbb{R}$ *of real numbers is uncountable, that is* $\mathfrak{c} > \aleph_0$*;*

b) $\mathfrak{c} = 2^{\aleph_0}$.

Proof. a) We us **Cantor's diagonal method**. Consider a function

$$f : \mathbb{N}^* \to [0, 1), \qquad f(n) = 0, a_{n1}a_{n2}\dots a_{nn}\dots,$$

where $a_{ni} \in \{0, \dots, 9\}$. Let $a_n \in \{0, \dots, 9\}$ such that $a_n \notin \{0, 9, a_{nn}\}$, and $a = 0, a_1, a_2, \dots a_n \dots$. Then clearly $f(n) \neq a$ for any $n \in \mathbb{N}$, hence f is not surjective. Thus we have shown that there is no bijective functions $f : \mathbb{N}^* \to \mathbb{R}$.

b) We know that

$$2^{\aleph_0} = |\mathrm{Hom}(\mathbb{N}^*, \{0, 1\})|,$$

therefore we shall use the representation of real numbers as infinite binary fractions (that is, in base-2 numeral system). Let $a \in [0, 1)$, $a = 0, a_1a_2\dots$ (in base-2), where $a_n \in \{0, 1\}$. We assume that 1 is not a period of the fractions. Consider the function

$$\phi : [0, 1) \to \mathrm{Hom}(\mathbb{N}^*, \{0, 1\}), \qquad \phi(a) = f, \text{ where } f(n) = a_n \ \ \forall\ n \geq 1.$$

Then ϕ is injective, because the representation of the real number a as infinite binary fraction without perioada 1 is unique. In addition, we have that

$$\complement(\mathrm{Im}\,\phi) = \{f : \mathbb{N}^* \to \{0, 1\} \mid\ \exists\ n_0 \text{ such that }\ f(n) = 1\ \ \forall\ n > n_0\}.$$

But a real number of the form $0, b_1b_2\dots b_{n_0}111\dots$ is rational, hence $\complement(\mathrm{Im}\,\phi)$ (that is, the set of numbers which may be represented with period 1) is countable. Since we have

$$\mathrm{Hom}(\mathbb{N}^*, \{0, 1\}) = \mathrm{Im}\,\phi \cup \complement(\mathrm{Im}\,\phi),$$

it follows that $2^{\aleph_0} = \mathfrak{c} + \aleph_0$, hence $\mathfrak{c} = 2^{\aleph_0}$. ■

Exercise 98 Prove that:

a) Any interval of real numbers has cardinality $\mathfrak{c}$, that is $\mathbb{R} \sim (0, 1) \sim (a, b) \sim [a, b) \sim [a, b] \sim (a, b]$ for any $a, b \in \mathbb{R}$ such that $a < b$;

b) The set $\mathbb{R} \setminus \mathbb{Q}$ of irrational numbers has cardinality $\mathfrak{c}$ (that is, $\mathbb{R} \sim \mathbb{R} \setminus \mathbb{Q}$).

Exercise 99 Prove that:

a) $\mathfrak{c}^2 = \mathfrak{c}^{\aleph_0} = \mathfrak{c}$;

b) $\mathfrak{c} + \mathfrak{c} = \mathfrak{c} \cdot \aleph_0 = {\aleph_0}^{\aleph_0} = \mathfrak{c}$.

Theorem 8.3.6 *If* α *is an infinite cardinal, then* $\alpha^2 = \alpha = \alpha\alpha'$ *for any* α'*, where* $0 \neq \alpha' \leq \alpha$*.*

8.4 Elements of combinatorics

We discuss several aspects regarding the computation of the number of elements of certain finite sets.

8.4.1 Arrangements, permutations, combinations

Definition 8.4.1 Let A and B be finite sets, with $|A| = k$ and $|B| = n$. We fix a total order for each set as follows: $A = \{a_1 < a_2 < \cdots < a_k\}$ and $B = \{b_1 < b_2 < \cdots < b_n\}$.

a) A sequence of length k of elements of B is called a k-**arrangement with repetition** of n elements. The number of k-arrangements with repetition of n elements is denoted by $\bar{A}_n^k$.

b) A sequence of length k of elements of B, in which every element appears at most once, is called a k-**arrangement** of n elements. The number of k-arrangements of n elements is denoted by A_n^k.

c) A sequence of length n of elements of B, in which every element appears exactly once, is called a **permutation** of n elements. The number of permutations of n elements is denoted by P_n.

d) A subset with k elements of B (where $k \leq n$) is called k-**combination** of n elements. The number k-combinations of n elements is denoted by $\binom{n}{k}$ or C_n^k.

Remark 8.4.2 Since the sequences are in fact functions, the above definitions may be reformulated as follows:

a) The number of k-arrangements with repetition of n elements is the number of all functions $f : A \to B$.

b) The number of k-arrangements of n elements is the number of injective functions $f : A \to B$.

c) The number of permutations of n elements is the number of bijective functions $f : A \to B$.

d) The number of k-combinations of n elements is the number of strictly increasing functions $f : A \to B$, or in other words, is the number of strictly increasing sequences of length k of elements of B.

Exercise 100 For $n = 5$ and $k = 2$, enumerate all

a) k-arrangementsle with repetition of n elements.

b) k-arrangements of n elements.

c) permutations of n elements.

d) k-combinations of n elements.

Exercise 101 Prove that:

a) $\bar{A}_n^k = n^k$.

b) If $k \le n$, then $A_n^k = \frac{n!}{(n-k)!}$.

c) $P_n = n!$.

d) If $k \le n$, then $C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}$.

e) $\bar{C}_n^k = (\binom{n}{k}) = \frac{(n+k-1)!}{(n-1)!k!}$.

Exercise 102 Prove that:

a) $C_n^k = C_n^{n-k}$; $C_n^k + C_n^{k+1} = C_{n+1}^{k+1}$.

b) $(X+Y)^n = \sum_{k=0}^n C_n^k X^{n-k} Y^k$ (*the binomial formula*).

c) $\sum_{k=0}^n C_n^k = 2^n$ (in two ways!).

Exercise 103 a) In how many ways may n be written as the sum of k nonzero natural numbers, if we take into account the order of the terms?

b) In how many ways may n be written as the sum of k natural numbers, if we take into account the order of the terms?

Exercise 104 Let $|A| = k$, $|B| = n$ and let $f : A \to B$.

a) If f is injective, find the number of left inverses of f.

b) If f is surjective, find the number of right inverses of f.

8.4.2 Inclusion–exclusion principle

Proposition 8.4.3 (Inclusion–exclusion principle) *If $A_1, \dots, A_n$ are finite sets, then the cardinal of their union is given by the formula*

$$|\bigcup_{i=1}^n A_i| = \sum_{i=1}^n |A_i| - \sum_{1\le i_1<i_2\le n} |A_{i_1} \cap A_{i_2}| + \sum_{1\le i_1<i_2<i_3\le n} |A_{i_1} \cap A_{i_2} \cap A_{i_3}| -$$

$$- \cdots + (-1)^{k+1} \sum_{1\le i_1<\cdots<i_k\le n} |A_{i_1} \cap \cdots \cap A_{i_k}| + \cdots + (-1)^{n+1} |\bigcap_{i=1}^n A_i|.$$

Proof. **1.** If $A \cap B = \emptyset$, then $|A \cup B| = |A| + |B|$. In general, $|A_1 \cup A_2| = |A_1| + |A_2 \setminus A_1|$ and $|A_2| = |A_2 \setminus A_1| + |A_1 \cap A_2 |$, hence

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|.$$

We continue by induction. We assume that the statement is true for n sets, and let $A_1, \dots, A_n, A_{n+1}$ be finite sets. Then

$$|\bigcup_{i=1}^{n+1} A_i| = |\bigcup_{i=1}^{n} A_i \cup A_{n+1}| = |\bigcup_{i=1}^{n} A_i| + |A_{n+1}| - |\bigcup_{i=1}^{n} A_i \cap A_{n+1}| =$$

$$= \sum_{i=1}^{n+1} |A_i| - \sum_{1\le i_1<i_2\le n} |A_{i_1} \cap A_{i_2}| + \cdots + (-1)^{n+1} |\bigcap_{i=1}^{n} A_i| + \bigcup_{i=1}^{n} (A_i \cap A_{n+1})| =$$

$$= \sum_{i=1}^{n+1} |A_i| - \sum_{1\le i_1<i_2\le n+1} |A_{i_1} \cap A_{i_2}| + \cdots + (-1)^{n+2} \, |\bigcap_{i=1}^{n+1} A_i|,$$

hence formula holds by the principle of mathematical induction. ■

Proof. 2. Another proof is based on the properties of the characteristic function, given in Exercise 54. We assume that $A_1, \dots, A_n \subseteq A$. Note in addition that for a finite subset $X \subseteq A$, we have

$$|X| = \sum_{x \in A} \chi_X(x).$$

We have $A_1 \cup \dots \cup A_n = \complement(\complement A_1 \cap \dots \cap \complement A_n)$ (De Morgan), hence

$$\begin{aligned}\chi_{A_1 \cup \dots \cup A_n} &= 1 - (1 - \chi_{A_1}) \dots (1 - \chi_{A_n}) = \\ &= \sum_{i=1}^{n} \chi_{A_i} - \sum_{i_1 < i_2} \chi_{A_{i_1}} \chi_{A_{i_2}} + \dots + (-1)^{k+1} \sum_{i_1 < \dots < i_k} \chi_{A_{i_1}} \dots \chi_{A_{i_k}} + \dots + (-1)^{n+1} \prod_{i=1}^{n} \chi_{A_i} = \\ &= \sum_{i=1}^{n} \chi_{A_i} - \sum_{i_1 < i_2} \chi_{A_{i_1} \cap A_{i_2}} + \dots + (-1)^{k+1} \sum_{i_1 < \dots < i_k} \chi_{A_{i_1} \cap \dots \cap A_{i_k}} + \dots + (-1)^{n+1} \chi_{\bigcap_{i=1}^{n} A_i}.\end{aligned}$$

By adding the values on all the elements $x \in A$ of the functions from both sides, we get the stated formula. ■

Exercise 105 *Applications of the inclusion–exclusion principle.*

a) Let $m = p_1^{k_1} \dots p_n^{k_n} \in \mathbb{N}$. Calculate the **Euler number** $\phi(m)$, where, by definition,

$$\phi(m) := |\{a \in \mathbb{N} \mid 1 \le a \le m, \ (a, m) = 1\}|.$$

b) Let $\sigma \in S_n$ be a permutation of degree n. We say that the element $i \in \{1, \dots, n\}$ is a **fixed point** of σ, if $\sigma(i) = i$. How many permutations of degree n do not have fixed points? (such permutations are also called *derangements.*)

Exercise 106 Let A and B be finite sets, with $|A| = k$ and $|B| = n$. If $k \ge n$, then the **the number of surjective functions** $f : A \to B$ is denoted by $s(k, n)$. Prove that:

$$s(k, n) = n^k - C_n^1 (n-1)^k + C_n^2 (n-2)^k + \dots + (-1)^{n-1} C_n^{n-1} 1^k.$$

Bibliography

[1] Adamson, I.: *A Set Theory Workbook.* Birkhäuser, Boston, 1998.

[2] Bilaniuk, S.: *A Problem Course in Mathematical Logic.* http://euclid.trentu.ca/math/sb/pcml/pcml-16.pdf. Trent University, Ontario, 2003.

[3] Breaz, S., Covaci, R.: *Elemente of logica, teoria of the sets and aritmetica.* Ed. Fundatiei for Studii Europene, Cluj-Napoca, 2006.

[4] Bloch, E.D.: *Proofs and Fundamentals.* 2nd ed. Springer, New York, 2011.

[5] Bloch, E.D.: *The Real Numbers and Real Analysis.* Springer, New York, 2011.

[6] Epp, S.: *Discrete Mathematics with Applications.* 4th ed. Brooks/Cole, Boston, 2011.

[7] Gallier, J.: *Discrete Mathematics.* 2nd ed. Springer Verlag, New York, 2011.

[8] Grätzer, G.: *Universal Algebra.* 2nd ed. Springer Verlag, Berlin, 2008.

[9] Grätzer, G.: *Lattice Theory: Foundation.* Birkhäuser, Basel, 2010.

[10] Halmos, P.: *Naive Set Theory.* D. Van Nostrand Company Inc., Princeton, 1974.

[11] Kneale, W., Kneale, M.: *The Development of Logic.* Oxford University Press, London, 1985.

[12] Krantz, S. G.: *Discrete Mathematics Demystified.* McGraw-Hill, New York, 2009.

[13] Krantz, S. G.: *The Proof is in the Pudding. The Changing Nature of Mathematical Proof.* Springer Verlag, New York, 2011.

[14] Lavrov, I.A., Maksimova, L.L.: *Probleme of teoria mulţimilor şi logică matematica.* Ed. Tehnica, Bucuresti, 1974.

[15] Levy, A.: *Basic Set Theory.* Dover Publications, New York, 1979.

[16] Lidl, R., Pilz, G.: *Applied Abstract Algebra.* Springer-Verlag, Berlin, 1998.

[17] Manin, Yu. I.: *A Course in Mathematical Logic for Mathematicians.* 2nd ed. Springer-Verlag, New York, 2010.

[18] Marcus, A., Szántó Cs., Tóth L.: *Logika és halmazelmélet.* Scientia, Cluj-Napoca, 2005.

[19] Nastasescu, C.: *Introducere în teoria mulţimilor.* Ed. Didactică and Pedagogică, Bucureşti, 1981.

[20] Purdea, I., Pic, Gh.: *Tratat of algebra moderna I.* Ed. Academiei, Bucuresti, 1977.

[21] Purdea, I.: *Culegere of probleme of algebră. Relations, functions and algebre universale.* Litografia Univ. Babeş-Bolyai, Cluj-Napoca, 1996.

[22] Ross, K. A., Wright Ch., *Discrete Mathematics.* Pearson Education, New Jersey, 2003.

Online resources:

- http://en.wikipedia.org/wiki/Set_theory
- http://en.wikipedia.org/wiki/Logic
- http://en.wikipedia.org/wiki/Foundations_of_mathematics
- http://en.wikipedia.org/wiki/Philosophy_of_mathematics
- http://en.wikipedia.org/wiki/History_of_mathematics
- http://en.wikipedia.org/wiki/History_of_logic