Foundations of Mathematics

Mohammad Safdari

Contents

Preface

These notes present a rigorous approach for laying the foundations of mathematics using mathematical logic and set theory. There are many books on mathematical logic and/or set theory. However, the usual approach in the study of logic is to use some basic (and sometimes advanced) notions of set theory. For example, a formula is usually defined as a finite sequence of symbols, with some specific properties. This requires the notions of function and natural numbers. Likewise, to start the study of set theory one needs some elementary tools of logic. For example, one needs to know the concept of formula in order to express some of the axioms of set theory.

In these notes, our aim is to develop mathematical logic without using set theory or natural numbers, and then to employ the developed logic to study elementary set theory. Of course, we cannot build a large portion of logic without the tools of other branches of mathematics; nevertheless, we can build a sufficient amount to be able to develop set theory. The basic idea for building logic without referring to sets is to treat formulas as primitive notions, and to accept their elementary properties as axioms. Hence, intuitively, we start from the notion of finite strings of symbols, and their manipulation, rather than the notion of set. This idea is not new, but to the best of author's knowledge, it has not been rigorously implemented in a text.

In the process of carrying out the above idea, we have been careful to avoid using complicated combinatorial properties of finite strings of symbols. So, only those properties of formulas which are more or less obvious have been postulated as axioms. Fortunately, it was possible to develop a sufficient portion of logic despite this restriction.

The notes are organized as follows. In Chapter 1 we construct a mathematical theory of logic as described in the above lines. In Chapter 2 we use logic to construct the Zermelo–Fraenkel axiomatic theory of sets. In Chapter 3 we use the standard approach to study relations and functions using sets. In Chapters 4-6 we construct natural numbers, integers, and rational, real, and complex numbers. The constructions in these chapters are standard. We also included some material about the combinatorics of finite sets and finite sequences, and algebraic structures like rings and fields. In Chapter 7, we examine the axiom of choice, and use it to

study countable sets. Finally, there are two appendices on factorization in integral domains, and decimal expansion of real numbers.

The intended audience of these notes are undergraduate students, or early graduate students, that have a reasonable amount of mathematical maturity. However, the first two chapters are intended to be also suitable for mathematicians interested in the subject. The more technical and advanced remarks in these two chapters can be safely ignored by students.

These notes, as you can see, are a work in progress. All suggestions, comments, and corrections are most welcome.

Chapter 1

Logic

Our goal in this chapter is to build mathematical logic without any reference to set theory, or any other mathematical theory. We want to start from scratch, and we want to avoid any presumptions. So, part of our work is to make our assumptions clear, and to state them precisely. We should mention that in this chapter, any use of the words "set, number, sequence, ..." refers to the ordinary meaning of them, not to their meaning as mathematical objects.

We communicate through speaking and writing. We cannot build any theory of logic without using these tools, since we need to somehow communicate our ideas with each other, even with ourselves. So when we talk about starting from scratch, we do not mean that we will not use anything at all; although that would have been preferable, if it was possible. The **meta-language** is the language in which we are communicating, which in our case is English. It can be any other language too, like French, Persian, or Japanese. But as we said, we cannot create our theory without a meta-language, or some other way of communication.

However, we do not want to use the meta-language for stating results in our theory of logic, or later in our development of mathematics. There are several reasons for this. For example, our meta-language, which is an ordinary language spoken by people, is not precise enough. Another problem with ordinary languages is that they allow self-referencing. A famous example of this phenomena is the following sentence

"This sentence is false."

If the above sentence is true then it must be false, and vice versa!

Therefore we need to create a language for stating our logical and mathematical results, which has absolute precision, and does not allow self-referencing. This is our first step in this chapter. When we consider this language in contrast to the meta-language, we just call it the **language**. The next step after the construction of the language, is to develop our rules of logic. First we have to accept some basic

rules as **axioms**, i.e. we have to accept them without any justification. This is the common approach in mathematics; and it is essentially unavoidable. Because if we want to deduce every statement from other statements, then we have to continue this process indefinitely; and this is not feasible. But it should be noted that although we do not logically justify the axioms from other valid statements, we have strong intuitive reasons regarding our choice of axioms.

A similar situation occurs when we define new notions. We cannot define every notion in terms of simpler notions; because then we have to continue this defining process indefinitely. So we have to work with some notions which do not have a definition. These notions are called **primitive notions**. But we still need to have some intuition about these undefined primitive notions, in order to be able to study them and other notions defined in terms of them. The axioms provide such intuitions. So in a sense, we can consider the axioms as the defining properties of the primitive notions.

After we constructed the language, and postulated our logical axioms, we need to deduce some properties from the axioms. But there are two problems here. Either we have not fully developed our logic yet; or we simply cannot apply it, because we want to avoid self-referencing, i.e. applying the logic to itself. However, we still need to be able to argue, so that we can transmit the fact that our choices are sound; and may be more importantly, to persuade ourselves that our choices are sound! In order to do that, we use argumentation in the meta-language. Note that we are actually doing this in this very paragraph! When we use reasoning in the meta-language, we are in fact using some rules of logic, or at least some elementary forms of them. We refer to these rules as **meta-logic**.

Note that the rules of the meta-logic are not different than the rules of the logic that we are going to construct. The difference lies at the level that we are applying them. So in some sense, we are using logic to construct logic! Although this is true, the situation is not as bad as it seems. Firstly, because we will only use very basic rules at the level of meta-logic (for more on this topic, see the discussion at the beginning of Section 1.7). Secondly, we can accept everything that we will prove about logic using meta-logic as true axioms. And we can consider those reasonings in meta-logic merely as convincing rationale for our choice of axioms. This is in fact one of the ways that mathematicians chose and continue to choose a set of axioms for a theory. They use tools outside the theory to convince themselves, and others, that those sets of axioms are appropriate. They often rely on their intuitions in this process; and as we will see, the reasonings in the meta-logic provide valuable intuitions for us. However, after we developed our logic, we must stop using the meta-logic completely.

1.1 Formulas

Let us start by constructing the language. This language is called the **language of set theory**. It is powerful enough to express all of mathematics. We will use it first to construct set theory. The language of set theory is an example of the so-called **formal languages**. There are many other formal languages; and their study is a major part of mathematical logic. But their proper study requires tools of set theory and other parts of mathematics. This is also true about the language of set theory itself. So we need to develop set theory before studying formal languages. However, we do not need to understand every aspect of the language of set theory, when we use it to construct set theory. Hence, we will not prove many results about the language of set theory at this point. We mainly develop it enough so that we can express our rules of logic, and then be able to construct our theory of sets.

Later, when we study formal languages, we can treat the language that we are going to construct now as a *mid-level meta-language*, i.e. a language between the ordinary language that we speak, and the formal language that we want to study. This way, we also avoid any circular reasoning, when we further study the formal language of set theory. Thus we will have two distinct copies of the language of set theory; one that we construct in this chapter, and another one which is a specific instance of formal languages. Although, we informally know that the two copies of the language of set theory are essentially the same.

Now, any language has an alphabet, i.e. a collection of letters. These are the symbols that we write on paper in order to communicate through that language. We do not define the alphabet though; we simply treat it as a primitive notion.

Primitive Notion 1.1. The **letters** are abstract notions that we represent by symbols on paper. We assume that we are able to recognize the letters from their symbols, and we can distinguish between them through their symbols. We refer to the collection of all letters as the **alphabet**. In the language of set theory, the letters are of the following types:

(i) Variables: $a, b, c, \dots, x, y, z, a_0, b_0, \dots, y_0, z_0, a_1, \dots, z_1, a_2, \dots$
(ii) Logical symbols: $\wedge, \vee, \to, \leftrightarrow, \neg, \bot, \forall, \exists$
(iii) Special symbols: $=, \in$
(iv) Parentheses: $(,)$

Remark. Note that the numbers used as an index in the variables are just symbols, and do not have any specific mathematical meaning at this point.

Remark. We tacitly assume that we have infinitely many distinct variables. Although in practice we will only use finitely many of them.

Primitive Notion 1.2. A **formula** is a primitive notion, which intuitively, is a finite sequence of letters that have some specific structure. These structures will

be discussed in the next axiom. We represent formulas by writing the symbols of their letters successively. We assume that we are able to recognize formulas from their representations, and we can distinguish between them through their representations.

Before proceeding any further, let us mention an important process in meta-language, and also in mathematics. We are talking about the process of **naming** objects, and assigning **notations** to them. For example the symbol "x" is a name, or a notation, for the variable x in our alphabet. Later we will see that the variable x itself can be a name for a set in our universe of sets! We will also use names for formulas. For example we will see that $\forall x(x = x)$ is a formula in our language of set theory. We can call and denote this formula by ϕ. This process of naming formulas, and other objects, has many advantages. For example it makes it easier to talk about and refer to complicated formulas. Another important benefit is that we can use a name as a placeholder for many, or all, formulas; and state a general property about them. We will see many instances of this in the rest of the chapter.

Sometimes we assign a new notation to an object that already has a notation, in order to make it easier to read and comprehend the text. For example in arithmetic we write n^4 to denote the number $n \cdot n \cdot n \cdot n$. This type of notation is called **abbreviated notation**. There are many different ways of abbreviation in mathematics, and it is not feasible to try to formalize them. However, they usually do not create any confusion. Whenever the need arises, we will use abbreviated notations in these notes too.

Axiom 1.1.

(i) $\perp$ *is a formula.*

(ii) *For every variables like* x, y *the following are formulas*

$$x = y, \qquad x \in y.$$

Note that x, y *can also be the same variable.*

(iii) *If* ϕ *is a formula then*

$$\neg(\phi)$$

is also a formula.

(iv) *Let* ϕ, ψ *be two formulas, that are not necessarily distinct. Then the following are also formulas*

$$(\phi) \wedge (\psi), \quad (\phi) \vee (\psi), \quad (\phi) \to (\psi), \quad (\phi) \leftrightarrow (\psi).$$

(v) *If* ϕ *is a formula and* x *is a variable, then the following are also formulas*

$$\forall x(\phi), \qquad \exists x(\phi).$$

Remark. We assume that every formula is constructed after several applications of the above rules, and there is no other way to construct a formula. But in order to make this precise we need some basic mathematical tools; so we do not do this here. However, we will not state any other axiom about construction of formulas; so if something is not constructed in the above ways, it cannot be shown that it is a formula in the language of set theory.

Hence, every formula is built in the above ways by starting from $\perp$, or formulas of the form $x = y$ and $x \in y$, for some variables x, y. For this reason, $\perp$ and formulas of the form $x = y$ and $x \in y$ are called *atomic formulas*.

Remark. The above axiom is actually an **axiom schema**. This means that it is actually an infinite collection of axioms. For example in part (i), x, y can be any variables; so for example $a = b$, $u = c_1$, $z_2 \in z_2$, ... are all formulas. Thus we actually have an axiom for every pair of variables. Also in the other parts of the axiom, ϕ, ψ can be any formulas. So we actually have an axiom for every pair of formulas.

Notation. When there is no risk of confusion, we will usually omit some of the parentheses in the notation introduced in the above axiom. For example we may write

$$\neg\phi, \quad \phi \wedge \psi, \quad \phi \vee \psi, \quad \phi \to \psi, \quad \phi \leftrightarrow \psi, \quad \forall x\phi, \quad \exists x\phi.$$

When we drop some of the parentheses, some different formulas may look like the same. For example if we drop the parentheses in $(\phi)\wedge(\psi \to \tau)$ and $(\phi\wedge\psi) \to (\tau)$, we will get $\phi \wedge \psi \to \tau$. In order to avoid the confusions that arise in this way we will follow the following *order of precedence*. We assume that $=, \in$ bind stronger than $\neg$, and $\neg$ binds stronger than $\forall, \exists$, and $\forall, \exists$ bind stronger than $\wedge, \vee$, and $\wedge, \vee$ bind stronger than $\to, \leftrightarrow$. So by our convention, $\phi \wedge \psi \to \tau$ means $(\phi \wedge \psi) \to (\tau)$. As another example, consider

$$\forall x\, \neg\phi \wedge x \in z \leftrightarrow (\psi \vee \tau \to \exists y\sigma),$$

which is an abbreviation of the following formula

$$\big((\forall x(\neg\phi)) \wedge (x \in z)\big) \leftrightarrow \big((\psi \vee \tau) \to (\exists y(\sigma))\big).$$

In practice, we usually keep some of the parentheses to make reading the expression easier.

Let us mention that at this point, we merely consider formulas as sequences of symbols, and we do not assign any meaning to them. In other words, we are only concerned with the **syntax** of the language, i.e. the formal rules of constructing well-formed expressions, aka formulas. Later when we develop set theory, we can assign meanings and provide interpretations for the formulas, i.e. we can study the **semantics** of the language. However, it is illuminating to informally introduce

some of the notions related to the semantics earlier. An important notion related to semantics is the notion of **truth**. We will not discuss this notion at length now; instead, we will use our intuitive understanding that a statement can be **true** or **false**, depending on whether its interpretation really happens or not. This also applies to the statements in the meta-language.

Next consider the logical symbols. The symbols $\bot, \neg, \wedge, \vee, \rightarrow, \leftrightarrow$ are called **logical connectives**. The symbol $\bot$ is called **falsum** or **absurdum**, and represents a false formula. It is included in the language mainly because it is more convenient to have a special notation for a formula which is always false. The symbol $\neg$ is called **negation**. Let ϕ be a formula. Then $\neg\phi$ means "not ϕ". The formula $\neg\phi$ is true when ϕ is false, and it is false when ϕ is true. The symbol $\wedge$ is called **conjunction**. For two formulas ϕ, ψ, the formula $\phi \wedge \psi$ means "ϕ and ψ". It is only true when both ϕ, ψ are true. The symbol $\vee$ is called **disjunction**. For two formulas ϕ, ψ, the formula $\phi \vee \psi$ means "ϕ or ψ". It is only true when at least one of ϕ, ψ is true; and it is false when both ϕ, ψ are false. Note that unlike some uses of the word "or" in ordinary language, the "or" in mathematical logic is *inclusive*. In other words, the truth of $\phi \vee \psi$ does not mean that exactly one of ϕ, ψ is true.

The symbol $\rightarrow$ is called **implication** or **conditional**. For two formulas ϕ, ψ, the formula $\phi \rightarrow \psi$ means "if ϕ then ψ". In the conditional formula $\phi \rightarrow \psi$, the formula ϕ is called the **antecedent**, and the formula ψ is called the **consequent**. The formula $\phi \rightarrow \psi$ is only false when ϕ is true and ψ is false. In particular, when both ϕ, ψ are false, then $\phi \rightarrow \psi$ is true. Note that unlike the usual use of "if ... then ..." in ordinary language, the truth of $\phi \rightarrow \psi$ does not mean that there is a causal relationship between ϕ, ψ. It merely means that if ϕ is true then ψ is true. Hence when ϕ is false, we cannot deduce anything about the truth of ψ, i.e. ψ can be true or false. To distinguish between the two concepts of implication and causal relationship, $\rightarrow$ is also called *material implication*. The symbol $\leftrightarrow$ is called **biconditional**. For two formulas ϕ, ψ, the formula $\phi \leftrightarrow \psi$ means "ϕ if and only if ψ" i.e. "if ϕ then ψ, and if ψ then ϕ". It is only true when ϕ, ψ are both true, or both false. Intuitively, $\phi \leftrightarrow \psi$ is true when both $\phi \rightarrow \psi$ and $\psi \rightarrow \phi$ are true. We will prove this fact later.

The following table summarizes the above information. It is called the *truth table*. Here T and F are abbreviations for true and false respectively. Note that at this moment, the truth table is just a tool to represent the informal meanings of logical connectives.

ϕ	ψ	$\neg\phi$	$\neg\psi$	$\phi \wedge \psi$	$\phi \vee \psi$	$\phi \rightarrow \psi$	$\phi \leftrightarrow \psi$	$\bot$
T	T	F	F	T	T	T	T	F
T	F	F	T	F	T	F	F	F
F	T	T	F	F	T	T	F	F
F	F	T	T	F	F	T	T	F

The symbols $\forall, \exists$ are called **quantifiers**. The symbol $\forall$ is called **universal quantifier**. Let ϕ be a formula. Suppose ϕ states some property about the variable x. To emphasize this we will write $\phi(x)$. Then $\forall x\phi(x)$ means that "for every x, $\phi(x)$ holds". Instead of "for every" we can also use "for all" or "for each". The formula $\forall x\phi(x)$ is only true when $\phi(x)$ is true for every choice of x. The symbol $\exists$ is called **existential quantifier**. The formula $\exists x\phi(x)$ means that "there exists x such that $\phi(x)$ holds". It is only true if $\phi(x)$ is true for at least one choice of x. An important concept regarding the quantifiers is their **domain of discourse**, which is also referred to as the **universe**. This is the collection of all objects x that we have to consider in order to examine the truth of $\forall x\phi$ or $\exists x\phi$. In the language of set theory, the domain of discourse is always the universe of all sets. But in meta-language, the domain of discourse can vary for different sentences. For example, we can express properties about all formulas; or we can have statements about all rules of logic.

Finally, consider the special symbols $=, \in$. The symbol $\in$ denotes the **set membership** relation. So $x \in y$ means that "x is an element of y", or equivalently "y contains x". The symbol $=$ denotes **equality**. So $x = y$ means that "x is equal to y". There is an important point here that should be noted. We have two notions of equality: one at the level of language, and one at the level of meta-language. When we say "x and y are equal" in the meta-language, we mean that x, y denote the same letter. But when we say "$x = y$" in the language, we mean that x, y denote the same set.

Notation. Instead of $\neg(x \in y)$ we usually write $x \notin y$. This means that "x does not belong to y". Also, instead of $\neg(x = y)$ we usually write $x \neq y$; which means that "x, y are not equal". Finally, instead of $\neg\bot$ we usually write $\top$; which represents a formula that is always true.

Remark. Let us emphasize again that the above semantic interpretations for the formulas of the language are all informal at this point. However, when we use connectives or quantifiers in meta-language, we assume that they have the above interpretations. In particular, our use of "or" is always inclusive; and our use of "if ... then ..." always indicates a material implication.

Remark. Another point that we wish to emphasize again is that a formula which is syntactically well-formed does not need to be true from a semantic viewpoint. In other words, there are formulas which are false. For example there might not be a set x such that $x \in x$, nevertheless, the formula $x \in x$ is syntactically well-formed.

Remark. The **converse** of the conditional formula $\phi \to \psi$ is the formula $\psi \to \phi$, and the **inverse** of $\phi \to \psi$ is the formula $\neg\phi \to \neg\psi$. If $\phi \to \psi$ is true, then its converse and inverse are not necessarily true, nor necessarily false. The **contrapositive** of $\phi \to \psi$ is the formula $\neg\psi \to \neg\phi$. We will see that the contrapositive of

a conditional formula is equivalent to it, i.e. the contrapositive is true if and only if the original formula is true.

***Example* 1.1.** Let ϕ be a formula. Consider the formula $\phi \wedge \neg\phi$. Note that semantically, this formula is always false. Because if ϕ is true then $\neg\phi$ is false, and hence $\phi \wedge \neg\phi$ is false. And if ϕ is false then $\phi \wedge \neg\phi$ is false. For any formula like ϕ, we call $\phi \wedge \neg\phi$ a **contradiction**. We also consider $\perp$ as a contradiction.

As another example consider $\phi \vee \neg\phi$. We can similarly see that from a semantic viewpoint, this formula is always true.

1.2 Rules of Inference

Primitive Notion 1.3. The primitive notion of **entailment** is a relation between several formulas $\phi_1, \phi_2, \ldots, \phi_n$ and another formula ψ. This relation is denoted by

$$\phi_1, \phi_2, \ldots, \phi_n \vdash \psi.$$

The symbol $\vdash$ is called **turnstile**. The formulas $\phi_1, \phi_2, \ldots, \phi_n$ are called **premises**, and the formula ψ is called **conclusion**. Intuitively, the above relation means that ϕ_1 and ϕ_2 and ... and ϕ_n together logically imply ψ. An entailment is allowed to have no premises; so we can have

$$\vdash \psi.$$

In this case we say that ψ is a **theorem**.

As we have done above, we represent an entailment by writing the representations of its formulas successively, and we separate the conclusion by a turnstile from the premises, and we separate the premises by " , ". We assume that we are able to recognize entailments from their representations, and we can distinguish between them through their representations. We also assume that we can recognize the formulas in an entailment from the representation of that entailment, and we can also figure out whether a given formula is a premise or the conclusion.

Remark. Let us emphasize again that the numbers which appear as indices in the names of formulas are just symbols, and do not have any specific mathematical meaning at this point.

Let us provide an informal semantic interpretation for the entailment relation. The entailment $\phi_1, \phi_2, \ldots, \phi_n \vdash \psi$ means that if $\phi_1, \phi_2, \ldots, \phi_n$ are all true, then ψ must also be true. It does not mean that ψ is necessarily true, because some of the $\phi_1, \phi_2, \ldots, \phi_n$ might not be true. But if we have $\vdash \psi$, then ψ must be true. In other words, theorems are true statements.

Notation. To simplify the notation, we will use capital Greek letters to denote several formulas. For example if we denote $\phi_1, \phi_2, \ldots, \phi_n$ by Γ, then we can write

$$\Gamma \vdash \psi$$

to denote $\phi_1, \phi_2, \ldots, \phi_n \vdash \psi$. In this notation, we allow Γ to be empty too; so $\Gamma \vdash \psi$ may also denote $\vdash \psi$.

The process of showing that an entailment such as $\Gamma \vdash \psi$ exists, is called **deduction**. In order to perform deductions, we need some rules to know when an entailment exists. These rules are called **rules of inference**. They are the rules of logic that we use to prove theorems. We will express them in the next few axioms.

Axiom 1.2. *Suppose $\Gamma, \Delta, \Lambda_0, \Lambda_1, \Lambda_2$ denote collections of several formulas, which can be empty too. Let ψ be a formula. Then we have*

(i)
$$\psi \vdash \psi.$$

(ii) *If $\Gamma \vdash \psi$ then*
$$\Gamma, \Delta \vdash \psi.$$

(iii) *If $\Gamma, \Delta, \Delta \vdash \psi$ then*
$$\Gamma, \Delta \vdash \psi.$$

(iv) *If $\Lambda_0, \Gamma, \Lambda_1, \Delta, \Lambda_2 \vdash \psi$ then*
$$\Lambda_0, \Delta, \Lambda_1, \Gamma, \Lambda_2 \vdash \psi.$$

(v) *If $\Gamma \vdash \phi$ and $\Delta, \phi \vdash \psi$ then*
$$\Gamma, \Delta \vdash \psi.$$

The above rules are called **structural rules**. They do not refer to any logical connective or quantifier; rather, they depend on the structure of the entailments themselves. The first rule is self-evident; it states that any formula implies itself. The second rule is called **weakening**. It means that if ψ can be deduced from a collection of formulas Γ, then ψ can also be deduced from a larger collection of formulas Γ, Δ. In other words, the entailment can be weakened by adding extra formulas to the premises. The third rule is called **contraction**. It means that repetition of formulas in the premises is superfluous, and extra occurrences of repeated formulas can be eliminated. The forth rule is called **exchange**. It means that the order of formulas in the premises is irrelevant.

The last rule is called the **cut rule**. It states that entailment is a transitive relation. In other words, it formalizes the intuitive fact that if several formulas imply several other formulas, and those other formulas imply another formula,

then the initial formulas also imply the last formula. Although we have stated the axiom in a way that there is only one formula in the middle, we can deduce the more general versions from this axiom. For example if $\Gamma \vdash \phi_1$ and $\Gamma \vdash \phi_2$, then $\Delta, \phi_1, \phi_2 \vdash \psi$ implies that $\Gamma, \Delta \vdash \psi$. To see this note that by using the cut rule we get $\Gamma, \Delta, \phi_1 \vdash \psi$. Notice that here we considered Δ, ϕ_1 in place of Δ. Now if we use the cut rule again we obtain $\Gamma, \Gamma, \Delta \vdash \psi$. Notice that here we considered Γ, Δ in place of Δ. Finally by using the contraction and exchange rules we get $\Gamma, \Delta \vdash \psi$, as desired. We can similarly extend the cut rule to the case of three or more formulas in the middle. But if we want to prove a general cut rule with an arbitrary number of formulas in the middle, we need mathematical induction. However, we usually do not have more than a few formulas in the middle; so we do not need the general version here.

Remark. A special case of the cut rule is when Δ is empty. In this case we have: If $\Gamma \vdash \phi$ and $\phi \vdash \psi$, then $\Gamma \vdash \psi$. This special case makes the transitivity of entailment more apparent.

Axiom 1.3 (Rules of inference for logical connectives)**.** *Suppose* Γ *denotes a collection of several formulas, which can be empty too. Let* ϕ, ψ, τ *be formulas. Then we have*

(i) *Introduction of* $\wedge$:

$$\phi, \psi \vdash \phi \wedge \psi.$$

(ii) *Elimination of* $\wedge$:

$$\phi \wedge \psi \vdash \phi, \qquad \textit{and} \qquad \phi \wedge \psi \vdash \psi.$$

(iii) *Introduction of* $\rightarrow$:

$$\textit{If} \quad \Gamma, \phi \vdash \psi \qquad \textit{then} \qquad \Gamma \vdash \phi \rightarrow \psi.$$

(iv) *Elimination of* $\rightarrow$*, or Modus ponens* :

$$\phi \rightarrow \psi,\ \phi \vdash \psi.$$

(v) *Introduction of* $\vee$:

$$\phi \vdash \phi \vee \psi, \qquad \textit{and} \qquad \psi \vdash \phi \vee \psi.$$

(vi) *Elimination of* $\vee$*, or Proof by cases* :

$$\textit{If} \quad \Gamma, \phi \vdash \tau \quad \textit{and} \quad \Gamma, \psi \vdash \tau \qquad \textit{then} \qquad \Gamma,\ \phi \vee \psi \vdash \tau.$$

(vii) *Introduction of* $\leftrightarrow$:

$$\textit{If} \quad \Gamma, \phi \vdash \psi \quad \textit{and} \quad \Gamma, \psi \vdash \phi \qquad \textit{then} \qquad \Gamma \vdash \phi \leftrightarrow \psi.$$

(viii) *Elimination of* $\leftrightarrow$:

$$\phi \leftrightarrow \psi,\ \phi \vdash \psi, \qquad \textit{and} \qquad \phi \leftrightarrow \psi,\ \psi \vdash \phi.$$

(ix) *Introduction of* $\neg$:

$$\textit{If} \quad \Gamma, \phi \vdash \bot \qquad \textit{then} \qquad \Gamma \vdash \neg\phi.$$

(x) *Elimination of* $\neg$:

$$\phi, \neg\phi \vdash \bot.$$

(xi) *Reductio ad absurdum* (*RAA*) :

$$\textit{If} \quad \Gamma, \neg\phi \vdash \bot \qquad \textit{then} \qquad \Gamma \vdash \phi.$$

Remark. Note that the axioms that we state about entailment are all axiom schemas. For example in the above axiom, ϕ, ψ can be any formulas. Similarly, Γ can be any collection of formulas. In other words, as we said before, for every formulas ϕ, ψ, and every collection of formulas Γ, we have an axiom as above.

The process of deduction is purely syntactic, i.e. it is a set of rules which tell us how to derive a formula from several other formulas, by looking only at their structure as sequences of symbols. However, when we consider the informal semantic interpretations for the formulas, we will see that the rules of inference are compatible with our intuitive understanding of the notion of reasoning. This is actually the reason that we choose them as rules of inference. For example we know that the truth of $\phi \wedge \psi$ is equivalent to the truth of both ϕ and ψ. Hence, the truth of ϕ and ψ implies the truth of $\phi \wedge \psi$, i.e. $\phi \wedge \psi$ is derivable from ϕ and ψ. This is exactly the rule of introduction of $\wedge$. Similarly, the truth of ϕ and ψ follows from the truth of $\phi \wedge \psi$, i.e. $\phi \wedge \psi$ implies both ϕ and ψ. And this is the rule of elimination of $\wedge$. The other rules also mirror our intuitive understanding of the logical connectives, and they can be justified similarly. Note that justification here means a heuristic argument which motivates the choice of some rule as an axiom, not a rigorous proof of the axiom!

Let us inspect the rules more closely. First consider the introduction of $\rightarrow$. It says that if we can deduce ψ from ϕ and some other hypotheses Γ, then we can infer from Γ alone that $\phi \rightarrow \psi$. In other words, Γ implies that "ϕ implies ψ". This is in agreement with our intuitive understandings of implication "$\rightarrow$" and entailment "$\vdash$". Next consider elimination of $\rightarrow$. It is also called **modus ponens**, which literally means "mood that affirms". It says that if $\phi \rightarrow \psi$ is true, and ϕ is true, then ψ must also be true. Note that this is in agreement with our intuition regarding material implication; we can even say that this is exactly the meaning of a conditional statement. The introduction and elimination of $\leftrightarrow$ can be interpreted

similarly. Note that they both reflect the fact that $\phi \leftrightarrow \psi$ is regarded as "$\phi \to \psi$ and $\psi \to \phi$".

The introduction of $\vee$ has a simple meaning. It says that if one of the ϕ or ψ is true then $\phi \vee \psi$ is true. Note that this rule implies that $\vee$ is the inclusive "or". The elimination of $\vee$ says that if we can deduce τ by assuming ϕ, and we can deduce τ by assuming ψ, then we can also deduce τ by assuming "either ϕ or ψ", i.e. we can deduce τ by assuming $\phi \vee \psi$. This rule is also called **proof by cases**, because it says that if ϕ or ψ is true, and we can show that τ is true in the case that ϕ is true, and we can also show that τ is true in the case that ψ is true, then we have showed that τ is true.

Finally consider the rules concerning $\neg$. The introduction of $\neg$ says that if the assumption of ϕ leads to a contradiction, then ϕ must be false, which means $\neg\phi$ must be true. And, the elimination of $\neg$ says that $\phi, \neg\phi$ lead to a contradiction. The **reductio ad absurdum** (RAA), which literally means "reduction to absurdity", is a special rule among the above rules. It says that if the assumption of "not ϕ" leads to a contradiction, then ϕ must be true. Intuitively, this rule presuppose that every formula is either true or false. So if "not ϕ" is false then ϕ must be true. However, this fact does not follow from the other axioms. In fact we will prove it using the RAA.

In addition, we should mention that although RAA and introduction of $\neg$ look similar, they are different rules. To see this note that if we apply the introduction of $\neg$ to the premises of RAA, then we obtain $\neg\neg\phi$. And from a syntactic viewpoint, there is no reason that the formula $\neg\neg\phi$ must imply ϕ. In fact it can be shown that the other axioms do not imply that ϕ follows from $\neg\neg\phi$. Thus we have to prove this fact using RAA.

Remark. The application of the rules I$\neg$ and RAA in a deduction is also known as **proof by contradiction**.

***Example* 1.2.** Let us demonstrate a simple application of the rules of inference. We know that $\psi, \phi \vdash \psi \wedge \phi$, due to the introduction of $\wedge$. Now we can use the exchange rule to switch the order of ψ, ϕ in the premises, and conclude that

$$\phi, \psi \vdash \psi \wedge \phi.$$

A question that arises is that why do we have two notions of implication denoted by $\to$ and $\vdash$? To answer this question, first note that $\vdash$ is a symbol in meta-language that we introduced; and it is not part of the language of set theory. Whereas $\to$ is a symbol in the language of set theory. So the two notions lie at different levels.

Another distinction is that $\vdash$ denotes a process of deduction, but $\to$ is just a syntactic symbol which we use to construct formulas, and a priori it does not have a meaning. In other words, $\phi \to \psi$ is a formula that says "if ϕ then ψ", which might be true or false. But $\phi \vdash \psi$ means that we can deduce ψ from ϕ in a process of

deduction. However, as we will see below, the two notions are closely related for conditional formulas which are true.

We know that if $\Gamma, \phi \vdash \psi$ then $\Gamma \vdash \phi \to \psi$, due to the introduction of $\to$. In other words, we can say that "entailment" implies "implication". Let us show that the converse also holds, i.e. if $\Gamma \vdash \phi \to \psi$ then $\Gamma, \phi \vdash \psi$. The reason is that by elimination of $\to$ we know that $\phi \to \psi$, $\phi \vdash \psi$. Hence by the exchange rule we get ϕ, $\phi \to \psi \vdash \psi$. Now we get the desired by the cut rule. Therefore we have shown that

$$\Gamma, \phi \vdash \psi \qquad \text{if and only if} \qquad \Gamma \vdash \phi \to \psi.$$

In particular we have

$$\phi \vdash \psi \qquad \text{if and only if} \qquad \vdash \phi \to \psi.$$

Informally, this equivalence means that in order to show that $\phi \to \psi$ is true, it suffices to deduce ψ by assuming ϕ, i.e. to show that $\phi \vdash \psi$. In mathematics, conditional statements are usually proved in this way.

Remark. When $\vdash \phi \to \psi$, we say that ϕ is a **sufficient condition** for ψ, because in order for ψ to hold it suffices that ϕ holds. We also say that ψ is a **necessary condition** for ϕ, because if ϕ holds then ψ must necessarily hold too. In addition, when $\vdash \phi \leftrightarrow \psi$, we say that ϕ is a necessary and sufficient condition for ψ, and vice versa.

Remark. Suppose we know that $\vdash \phi$, i.e. ϕ is a theorem. Also, suppose we want to show that $\Gamma \vdash \psi$. Then it suffices to show that $\Gamma, \phi \vdash \psi$. Because by the cut rule, from $\vdash \phi$ and $\Gamma, \phi \vdash \psi$ we can conclude that $\Gamma \vdash \psi$. This argument shows that we may use theorems in the premises of entailments to deduce other formulas, and then we can discard those theorems.

Notation. For simplicity, we will denote the introduction and elimination rules by the letters "I" and "E" followed by the respective connectives. For example we will denote the introduction of $\to$ by I$\to$, and the elimination of $\vee$ by E$\vee$.

***Example* 1.3.** For every formula like ϕ we have $\phi, \neg\phi \vdash \bot$ by E$\neg$. Hence by RAA we get

$$\phi \vdash \phi.$$

Thus we can prove the above structural rule from the inference rules for connectives. However, it seems more natural to treat the above rule as an axiom. As another example note that we also have

$$\vdash \phi \to \phi.$$

Because we know that $\phi \vdash \phi$. Thus by I$\to$ we get $\vdash \phi \to \phi$, as desired.

***Example* 1.4.** By E$\wedge$ we know that $\phi \wedge \neg\phi \vdash \phi$ and $\phi \wedge \neg\phi \vdash \neg\phi$. On the other hand, by E$\neg$ we have $\phi, \neg\phi \vdash \bot$. Hence by the cut rule we get

$$\phi \wedge \neg\phi \vdash \bot.$$

More generally, suppose $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$. Then by the cut rule and E$\neg$ we obtain $\Gamma \vdash \bot$.

In our first theorem, we present a few additional rules of inference. Note that here the meaning of "theorem" is different than its meaning in the Primitive Notion 1.3. Here, "theorem" is a statement in the meta-language which states a valid fact about logic; in contrast, "theorem" in Primitive Notion 1.3 is a formula in the language which states a true fact about sets. A better term for theorems in the meta-language could be "meta-theorem", but for simplicity we will keep calling them theorems.

Another important point is that the theorems that we state about entailment are also schemas, similar to the axioms. For example in the following theorem, ϕ, ψ, τ, σ can be any formulas. Similarly, Γ can be any collection of formulas. In other words, as we said before, for every formulas ϕ, ψ, τ, σ, and every collection of formulas Γ, we have a theorem as below.

Remark. We will usually try to cite every rule that we use inside a proof, but for simplicity, sometimes we will not mention our uses of structural rules, especially the exchange rule.

Theorem 1.1. *Suppose Γ denotes a collection of several formulas, which can be empty too. Let ϕ, ψ, τ, σ be formulas. Then we have*

(i) *Ex falso quodlibet (EFQ)* :

$$\phi \wedge \neg\phi \vdash \psi, \qquad \text{and} \qquad \bot \vdash \psi.$$

(ii) *Law of excluded middle* :

$$\vdash \phi \vee \neg\phi.$$

(iii) *Law of non-contradiction* :

$$\vdash \neg(\phi \wedge \neg\phi), \qquad \text{and} \qquad \vdash \top.$$

(iv) *Modus tollens* :

$$\phi \to \psi,\ \neg\psi \vdash \neg\phi.$$

(v) *Hypothetical syllogism* :

$$\phi \to \psi,\ \psi \to \tau \vdash \phi \to \tau.$$

(vi) *Disjunctive syllogism, or Modus tollendo ponens* :

$$\phi \vee \psi,\ \neg\phi \vdash \psi, \qquad \textit{and} \qquad \phi \vee \psi,\ \neg\psi \vdash \phi.$$

(vii) *Modus ponendo tollens* :

$$\neg(\phi \wedge \psi),\ \phi \vdash \neg\psi, \qquad \textit{and} \qquad \neg(\phi \wedge \psi),\ \psi \vdash \neg\phi.$$

(viii) *Constructive dilemma* :

$$\textit{If} \quad \Gamma, \phi \vdash \tau \quad \textit{and} \quad \Gamma, \psi \vdash \sigma \qquad \textit{then} \qquad \Gamma,\ \phi \vee \psi \vdash \tau \vee \sigma.$$

(ix) *Destructive dilemma* :

$$\textit{If} \quad \Gamma, \phi \vdash \tau \quad \textit{and} \quad \Gamma, \psi \vdash \sigma \qquad \textit{then} \qquad \Gamma,\ \neg\tau \vee \neg\sigma \vdash \neg\phi \vee \neg\psi.$$

(x)

$$\textit{If} \quad \Gamma, \phi \vdash \neg\phi \qquad \textit{then} \qquad \Gamma \vdash \neg\phi.$$

Proof. **(i)** We know that $\bot \vdash \bot$. Hence by the weakening rule we get $\bot, \neg\psi \vdash \bot$. Now by RAA we obtain $\bot \vdash \psi$, as desired. In addition we know that $\phi \wedge \neg\phi \vdash \bot$. Therefore by the cut rule we also get $\phi \wedge \neg\phi \vdash \psi$.

(ii) By I$\vee$ we have $\phi \vdash \phi \vee \neg\phi$. We also know that $\neg(\phi \vee \neg\phi) \vdash \neg(\phi \vee \neg\phi)$. So by the weakening rule we get $\neg(\phi \vee \neg\phi), \phi \vdash \phi \vee \neg\phi$, and $\neg(\phi \vee \neg\phi), \phi \vdash \neg(\phi \vee \neg\phi)$. Hence by the cut rule and E$\neg$ we obtain

$$\neg(\phi \vee \neg\phi), \phi \vdash \bot.$$

Thus by I$\neg$ we get $\neg(\phi \vee \neg\phi) \vdash \neg\phi$. Now by I$\vee$ and the cut rule we obtain $\neg(\phi \vee \neg\phi) \vdash \phi \vee \neg\phi$. Therefore by the cut rule and E$\neg$ we get

$$\neg(\phi \vee \neg\phi) \vdash \bot.$$

Hence by RAA we get $\vdash \phi \vee \neg\phi$, as desired.

(iii) We know that $\phi \wedge \neg\phi \vdash \bot$. Hence by I$\neg$ we get $\vdash \neg(\phi \wedge \neg\phi)$, as desired. Similarly, we know that $\bot \vdash \bot$. So by I$\neg$ we get $\vdash \neg\bot$, which is the same as saying $\vdash \top$.

(iv) By E$\to$ and the weakening rule we have $\phi \to \psi, \phi, \neg\psi \vdash \psi$. We also know that $\phi \to \psi, \phi, \neg\psi \vdash \neg\psi$. Thus by the exchange and cut rules, and E$\neg$, we get $\phi \to \psi, \neg\psi, \phi \vdash \bot$. Hence by I$\neg$ we obtain $\phi \to \psi, \neg\psi \vdash \neg\phi$, as desired.

(v) By E$\to$ we have $\phi \to \psi, \phi \vdash \psi$, and $\psi \to \tau, \psi \vdash \tau$. Hence by the cut and exchange rules we have $\phi \to \psi, \psi \to \tau, \phi \vdash \tau$. Thus by I$\to$ we get $\phi \to \psi, \psi \to \tau \vdash \phi \to \tau$, as desired.

(vi) By E$\neg$ we have $\phi, \neg\phi \vdash \bot$, and by EFQ rule we have $\bot \vdash \psi$. Hence by the cut rule we get $\phi, \neg\phi \vdash \psi$. Also, by the weakening rule we have $\psi, \neg\phi \vdash \psi$.

Therefore by E$\vee$ and exchange rule we get $\phi \vee \psi, \neg\phi \vdash \psi$, as desired. The other case can be proved similarly.

(vii) By the weakening rule and I$\wedge$ we have $\neg(\phi \wedge \psi), \phi, \psi \vdash \phi \wedge \psi$. We also know that $\neg(\phi \wedge \psi), \phi, \psi \vdash \neg(\phi \wedge \psi)$. Thus by the cut rule and E$\neg$ we get

$$\neg(\phi \wedge \psi), \phi, \psi \vdash \bot.$$

Hence by I$\neg$ we obtain $\neg(\phi \wedge \psi), \phi \vdash \neg\psi$; and by exchange rule and I$\neg$ we obtain $\neg(\phi \wedge \psi), \psi \vdash \neg\phi$, as desired.

(viii) By I$\vee$ we have $\Gamma, \phi \vdash \tau \vdash \tau \vee \sigma$, and $\Gamma, \psi \vdash \sigma \vdash \tau \vee \sigma$. Hence by the cut rule and E$\vee$ we get $\Gamma, \phi \vee \psi \vdash \tau \vee \sigma$.

(ix) By the weakening and exchange rules we have $\Gamma, \neg\tau, \phi \vdash \tau$. We also know that $\Gamma, \neg\tau, \phi \vdash \neg\tau$. Hence by the cut rule and E$\neg$ we get $\Gamma, \neg\tau, \phi \vdash \bot$. Thus by I$\neg$ we obtain $\Gamma, \neg\tau \vdash \neg\phi$. Therefore by I$\vee$ and the cut rule we get $\Gamma, \neg\tau \vdash \neg\phi \vee \neg\psi$. Similarly we can show that $\Gamma, \neg\sigma \vdash \neg\phi \vee \neg\psi$. Thus by E$\vee$ we get $\Gamma, \neg\tau \vee \neg\sigma \vdash \neg\phi \vee \neg\psi$, as desired.

(x) The assumption is that $\Gamma, \phi \vdash \neg\phi$. We also know that $\Gamma, \phi \vdash \phi$. Thus by the cut rule and E$\neg$ we get $\Gamma, \phi \vdash \bot$. Hence by I$\neg$ we obtain $\Gamma \vdash \neg\phi$, as desired. ■

The rule ex falso quodlibet (EFQ), which literally means "from falsehood anything (follows)", states that a contradiction like $\phi \wedge \neg\phi$, or $\bot$, can imply any formula like ψ. The law of excluded middle states that for every formula like ϕ, either ϕ or $\neg\phi$ must be true. Hence, a formula is either true or false. And the law of non-contradiction states that a contradiction like $\phi \wedge \neg\phi$, or $\bot$, cannot be true. In other words, a formula cannot be both true and false.

The rule modus tollens, which literally means "mood that denies", informally says that if ϕ implies ψ, and ψ is false, then ϕ must be false too. It is closely related to the law of contraposition, which is stated in Theorem 1.6. The hypothetical syllogism says that implication is transitive. The word *syllogism* is the name of inference rules in Aristotelean logic. Finally, let us mention that modus tollendo ponens literally means "mood that affirms by denying", and modus ponendo tollens literally means "mood that denies by affirming".

The last part of the above theorem says that if a formula implies its negation, then that formula must be false, i.e. its negation must be true. We can similarly show that if the negation of a formula implies the formula, then the negation must be false, i.e. the formula must be true. In other words:

$$\text{If} \quad \Gamma, \neg\phi \vdash \phi \qquad \text{then} \qquad \Gamma \vdash \phi.$$

This rule is known as *consequentia mirabilis*, which literally means "admirable consequence". The proof of this fact is similar to the above, but we have to use RAA instead of I$\neg$. We can also prove it by using the last part of the above theorem and the double negation law.

***Example* 1.5.** Let ϕ, ψ be formulas. Then we have

$$\text{If} \quad \vdash \neg\phi \qquad \text{then} \qquad \vdash \phi \to \psi.$$

This confirms our intuitive understanding of material implication, namely, the fact that $\phi \to \psi$ is true when ϕ is false. To prove it, note that by weakening rule we have $\phi \vdash \neg\phi$. We also know that $\phi \vdash \phi$. Hence by E$\neg$ and the cut rule we get $\phi \vdash \bot$. Thus by EFQ rule we obtain $\phi \wedge \neg\phi \vdash \psi$. Therefore by the cut rule we get $\phi \vdash \psi$. Hence by I$\to$ we have $\vdash \phi \to \psi$, as desired. It is also easy to show that

$$\text{If} \quad \vdash \psi \qquad \text{then} \qquad \vdash \phi \to \psi.$$

In other words, $\phi \to \psi$ is true when ψ is true. Because by the weakening rule we have $\phi \vdash \psi$. Hence by I$\to$ we obtain $\vdash \phi \to \psi$, as desired.

1.3 Equivalent Formulas

Definition 1.1. We say two formulas ϕ, ψ are **equivalent** if $\phi \vdash \psi$ and $\psi \vdash \phi$. In this case we write

$$\phi \equiv \psi.$$

Remark. An important point to keep in mind is that the "if" in definitions actually means "if and only if". Thus the above definition actually says "ϕ, ψ are equivalent if and only if $\phi \vdash \psi$ and $\psi \vdash \phi$". But the tradition in mathematics is to use "if" in definitions instead, and we will adhere to this convention in these notes.

Remark. Note that by the introduction of $\leftrightarrow$, if $\phi \equiv \psi$ then we have

$$\vdash \phi \leftrightarrow \psi.$$

Thus from a semantic viewpoint, equivalent formulas are either both true, or both false.

Theorem 1.2. *Suppose ϕ, ψ, τ are formulas, and $\phi \equiv \psi$. Let Γ denote a collection of several formulas, which can be empty too. Then the entailment $\Gamma, \phi \vdash \tau$ holds if and only if the entailment $\Gamma, \psi \vdash \tau$ holds.*

Proof. To see this, suppose $\Gamma, \psi \vdash \tau$ holds. Then since $\phi \vdash \psi$, the cut rule implies that $\phi, \Gamma \vdash \tau$ holds too. Hence by the exchange rule we get $\Gamma, \phi \vdash \tau$, as desired. The converse can be proved similarly. ■

Remark. The above theorem means that we can replace a formula in the premises of an entailment by an equivalent formula, and the new entailment is equivalent to the original one.

Notation. Sometimes we may write $\Gamma \vdash \phi \vdash \psi$ to denote "$\Gamma \vdash \phi$ and $\phi \vdash \psi$". Note that due to the cut rule if $\Gamma \vdash \phi \vdash \psi$ then $\Gamma \vdash \psi$ too. In addition, we may write $\phi \equiv \psi \equiv \tau$ to denote "$\phi \equiv \psi$ and $\psi \equiv \tau$". Note that by the next theorem we can also conclude that $\phi \equiv \tau$. Both of these abbreviated notations can be used for more than three formulas.

Theorem 1.3. *Suppose ϕ, ψ, τ are formulas. Then we have*

(i) *Reflexivity* : $\phi \equiv \phi$.
(ii) *Symmetry* : *If* $\phi \equiv \psi$ *then* $\psi \equiv \phi$.
(iii) *Transitivity* : *If* $\phi \equiv \psi$ *and* $\psi \equiv \tau$, *then* $\phi \equiv \tau$.

Proof. **(i)** This is a trivial consequence of the fact that $\phi \vdash \phi$.

(ii) If $\phi \equiv \psi$ then by definition we have "$\phi \vdash \psi$ and $\psi \vdash \phi$". However this the same as saying "$\psi \vdash \phi$ and $\phi \vdash \psi$". Therefore we get $\psi \equiv \phi$ as desired. Note that here we are using the fact that the connective "and" in meta-language is commutative, i.e. the order of the phrases which are connected by "and" does not affect the truth of the compound statement.

(iii) If $\phi \equiv \psi$ and $\psi \equiv \tau$, then by definition we have $\phi \vdash \psi$ and $\psi \vdash \phi$, and $\psi \vdash \tau$ and $\tau \vdash \psi$. In other words we have $\phi \vdash \psi \vdash \tau$ and $\tau \vdash \psi \vdash \phi$. Hence by the cut rule we obtain $\phi \vdash \tau$ and $\tau \vdash \phi$. Thus $\phi \equiv \tau$ as desired. Note that we feel free to use our hypotheses in any order we want. In other words, we are using the exchange rule at the level of meta-logic. ■

Theorem 1.4. *Suppose ϕ, ψ, τ, σ are formulas. Also suppose $\phi \equiv \psi$ and $\tau \equiv \sigma$. Then we have*

(i) $\neg\phi \equiv \neg\psi$.
(ii) $\phi \wedge \tau \equiv \psi \wedge \sigma$.
(iii) $\phi \vee \tau \equiv \psi \vee \sigma$.
(iv) $\phi \to \tau \equiv \psi \to \sigma$.
(v) $\phi \leftrightarrow \tau \equiv \psi \leftrightarrow \sigma$.

Proof. **(i)** We know that $\phi \vdash \psi$ and $\psi \vdash \phi$. We have to show that $\neg\phi \vdash \neg\psi$ and $\neg\psi \vdash \neg\phi$. Note that $\neg\phi \vdash \neg\phi$. Thus by the weakening and exchange rules we have $\neg\phi, \psi \vdash \phi$ and $\neg\phi, \psi \vdash \neg\phi$. Therefore by E$\neg$ and cut rule we have $\neg\phi, \psi \vdash \bot$. Hence by I$\neg$ we get $\neg\phi \vdash \neg\psi$. Similarly we can show that $\neg\psi \vdash \neg\phi$.

(ii) By E$\wedge$ we have $\phi \wedge \tau \vdash \phi \vdash \psi$ and $\phi \wedge \tau \vdash \tau \vdash \sigma$. On the other hand, by I$\wedge$ we have $\psi, \sigma \vdash \psi \wedge \sigma$. Hence by the cut rule we have $\phi \wedge \tau \vdash \psi \wedge \sigma$. Similarly we can show that $\psi \wedge \sigma \vdash \phi \wedge \tau$. Thus $\phi \wedge \tau \equiv \psi \wedge \sigma$, as desired.

(iii) By I$\vee$ we have $\phi \vdash \psi \vdash \psi \vee \sigma$ and $\tau \vdash \sigma \vdash \psi \vee \sigma$. Hence by E$\vee$ we have $\phi \vee \tau \vdash \psi \vee \sigma$. Similarly we can show $\psi \vee \sigma \vdash \phi \vee \tau$, and conclude the desired result.

(iv) By E$\to$ we have $\phi \to \tau, \phi \vdash \tau \vdash \sigma$. Hence by Theorem 1.2 we also have $\phi \to \tau, \psi \vdash \sigma$, since $\phi \equiv \psi$. Thus by I$\to$ we get $\phi \to \tau \vdash \psi \to \sigma$. Similarly we can show that $\psi \to \sigma \vdash \phi \to \tau$.

(v) By E$\leftrightarrow$ we have $\phi \leftrightarrow \tau, \phi \vdash \tau \vdash \sigma$ and $\phi \leftrightarrow \tau, \tau \vdash \phi \vdash \psi$. Hence by Theorem 1.2 we also have $\phi \leftrightarrow \tau, \psi \vdash \sigma$ and $\phi \leftrightarrow \tau, \sigma \vdash \psi$, because $\phi \equiv \psi$ and $\tau \equiv \sigma$. Thus by I$\leftrightarrow$ we get $\phi \leftrightarrow \tau \vdash \psi \leftrightarrow \sigma$. Similarly we can show that $\psi \leftrightarrow \sigma \vdash \phi \leftrightarrow \tau$. ■

Suppose a formula is composed of several other formulas. Then the above theorem enables us to replace some of the components by equivalent formulas to obtain a formula equivalent to the original compound formula. We will not prove the general case of this fact here, but let us demonstrate it by an example. Suppose $\phi \equiv \psi$ and $\tau \equiv \sigma$. Consider the following formula

$$(\phi \wedge \phi_1) \rightarrow (\phi_2 \rightarrow \tau \vee \phi_3).$$

We claim that it is equivalent to $(\psi \wedge \phi_1) \rightarrow (\phi_2 \rightarrow \sigma \vee \phi_3)$. To see this note that any formula is equivalent to itself. Therefore by the above theorem we have

$$\phi \wedge \phi_1 \equiv \psi \wedge \phi_1, \qquad \tau \vee \phi_3 \equiv \sigma \vee \phi_3.$$

Now if we apply the theorem again we obtain $\phi_2 \rightarrow \tau \vee \phi_3 \equiv \phi_2 \rightarrow \sigma \vee \phi_3$, and therefore we get

$$(\phi \wedge \phi_1) \rightarrow (\phi_2 \rightarrow \tau \vee \phi_3) \equiv (\psi \wedge \phi_1) \rightarrow (\phi_2 \rightarrow \sigma \vee \phi_3),$$

as desired.

Theorem 1.5. *Suppose ϕ, ψ, τ are formulas. If $\phi \vdash \psi \vdash \tau \vdash \phi$ then we have $\phi \equiv \psi \equiv \tau$.*

Remark. This theorem provides a shortcut for proving that three formulas are equivalent with each other. It says that instead of proving that each formula is deduced from each of the other formulas, we can just show that we have a cycle of entailments. This is how such equivalences are usually proved in mathematics. Similarly we can show that the analogous statements are true for more than three formulas.

Proof. We know that $\psi \vdash \tau \vdash \phi$, so by the cut rule we have $\psi \vdash \phi$. On the other hand we know that $\phi \vdash \psi$. Hence we get $\phi \equiv \psi$. Similarly we know that $\psi \vdash \tau$. We also know that $\tau \vdash \phi \vdash \psi$. Thus by the cut rule we get $\tau \vdash \psi$. Therefore we have $\psi \equiv \tau$ too. Note that the equivalence $\phi \equiv \tau$ is not formally part of the statement $\phi \equiv \psi \equiv \tau$, but as we said before, it follows from the transitivity of $\equiv$. ■

Theorem 1.6. *Suppose ϕ, ψ, τ are formulas. Then we have*

(i) *Commutativity* :

$$\phi \wedge \psi \equiv \psi \wedge \phi, \qquad \textit{and} \qquad \phi \vee \psi \equiv \psi \vee \phi.$$

(ii) *Associativity* :

$$(\phi \wedge \psi) \wedge \tau \equiv \phi \wedge (\psi \wedge \tau), \qquad \textit{and} \qquad (\phi \vee \psi) \vee \tau \equiv \phi \vee (\psi \vee \tau).$$

(iii) *Distributivity* :

$$\tau \vee (\phi \wedge \psi) \equiv (\tau \vee \phi) \wedge (\tau \vee \psi), \qquad \textit{and} \qquad \tau \wedge (\phi \vee \psi) \equiv (\tau \wedge \phi) \vee (\tau \wedge \psi).$$

(iv) *Idempotency* :

$$\phi \wedge \phi \equiv \phi, \qquad \textit{and} \qquad \phi \vee \phi \equiv \phi.$$

(v) *Absorption* :

$$\phi \wedge (\phi \vee \psi) \equiv \phi, \qquad \textit{and} \qquad \phi \vee (\phi \wedge \psi) \equiv \phi.$$

(vi) *De Morgan's laws* :

$$\neg(\phi \wedge \psi) \equiv \neg\phi \vee \neg\psi, \qquad \textit{and} \qquad \neg(\phi \vee \psi) \equiv \neg\phi \wedge \neg\psi.$$

(vii) *Double negation law* :

$$\neg\neg\phi \equiv \phi.$$

(viii) *Law of material implication* :

$$\phi \to \psi \equiv \neg\phi \vee \psi.$$

(ix) *Law of contraposition* :

$$\phi \to \psi \equiv \neg\psi \to \neg\phi.$$

(x) *Negation of a conditional* :

$$\neg(\phi \to \psi) \equiv \phi \wedge \neg\psi.$$

(xi)

$$\phi \leftrightarrow \psi \equiv (\phi \to \psi) \wedge (\psi \to \phi),$$
$$\psi \leftrightarrow \phi \equiv \phi \leftrightarrow \psi \equiv \neg\phi \leftrightarrow \neg\psi.$$

(xii)

$$\phi \wedge \neg\phi \equiv \bot, \qquad \textit{and} \qquad \phi \vee \neg\phi \equiv \top.$$

(xiii)

$$\phi \wedge \top \equiv \phi, \qquad \textit{and} \qquad \phi \vee \bot \equiv \phi,$$
$$\phi \vee \top \equiv \top, \qquad \textit{and} \qquad \phi \wedge \bot \equiv \bot.$$

(xiv)

$$\neg\phi \equiv \phi \to \bot.$$

Proof. **(i)** By E$\wedge$ we have $\phi \wedge \psi \vdash \psi$ and $\phi \wedge \psi \vdash \phi$. On the other hand by I$\wedge$ we have $\psi, \phi \vdash \psi \wedge \phi$. Hence by the cut rule we get $\phi \wedge \psi \vdash \psi \wedge \phi$. Similarly we can show that $\psi \wedge \phi \vdash \phi \wedge \psi$.

Next, by I$\vee$ we have $\phi \vdash \psi \vee \phi$ and $\psi \vdash \psi \vee \phi$. Hence by E$\vee$ we get $\phi \vee \psi \vdash \psi \vee \phi$. Similarly we can show that $\psi \vee \phi \vdash \phi \vee \psi$.

(ii) By E$\wedge$ we have

$$(\phi \wedge \psi) \wedge \tau \vdash \phi \wedge \psi \vdash \phi, \qquad (\phi \wedge \psi) \wedge \tau \vdash \phi \wedge \psi \vdash \psi,$$

and $(\phi \wedge \psi) \wedge \tau \vdash \tau$. Thus by the cut rule and I$\wedge$ we get $(\phi \wedge \psi) \wedge \tau \vdash \psi \wedge \tau$, and therefore $(\phi \wedge \psi) \wedge \tau \vdash \phi \wedge (\psi \wedge \tau)$. Similarly we can show that $\phi \wedge (\psi \wedge \tau) \vdash (\phi \wedge \psi) \wedge \tau$.

Next, by I$\vee$ we have $\phi \vdash \phi \vee (\psi \vee \tau)$, and

$$\psi \vdash \psi \vee \tau \vdash \phi \vee (\psi \vee \tau), \qquad \tau \vdash \psi \vee \tau \vdash \phi \vee (\psi \vee \tau).$$

Thus by E$\vee$ we get $\phi \vee \psi \vdash \phi \vee (\psi \vee \tau)$, and therefore $(\phi \vee \psi) \vee \tau \vdash \phi \vee (\psi \vee \tau)$. Similarly we can show that $\phi \vee (\psi \vee \tau) \vdash (\phi \vee \psi) \vee \tau$.

(iii) By E$\wedge$ and I$\vee$ we have $\phi \wedge \psi \vdash \phi \vdash \tau \vee \phi$, and $\phi \wedge \psi \vdash \psi \vdash \tau \vee \psi$. Hence by the cut rule and I$\wedge$ we get $\phi \wedge \psi \vdash (\tau \vee \phi) \wedge (\tau \vee \psi)$. Now by I$\vee$ we have $\tau \vdash \tau \vee \phi$, and $\tau \vdash \tau \vee \psi$. Hence by the cut rule and I$\wedge$ we get $\tau \vdash (\tau \vee \phi) \wedge (\tau \vee \psi)$. Therefore by E$\vee$ we get

$$\tau \vee (\phi \wedge \psi) \vdash (\tau \vee \phi) \wedge (\tau \vee \psi).$$

On the other hand, by I$\wedge$ and I$\vee$ we have $\phi, \psi \vdash \phi \wedge \psi \vdash \tau \vee (\phi \wedge \psi)$. Also, by the weakening rule and I$\vee$ we have $\phi, \tau \vdash \tau \vee (\phi \wedge \psi)$. Hence by E$\vee$ we get $\phi, \tau \vee \psi \vdash \tau \vee (\phi \wedge \psi)$. If we use the exchange rule we get $\tau \vee \psi, \phi \vdash \tau \vee (\phi \wedge \psi)$. Now by the weakening rule and I$\vee$ we also have $\tau \vee \psi, \tau \vdash \tau \vee (\phi \wedge \psi)$. Therefore by E$\vee$ we obtain $\tau \vee \psi, \tau \vee \phi \vdash \tau \vee (\phi \wedge \psi)$. Thus by E$\wedge$ and the cut rule we get

$$(\tau \vee \phi) \wedge (\tau \vee \psi) \vdash \tau \vee (\phi \wedge \psi).$$

Next, by E$\wedge$ and I$\vee$ we have $\tau \wedge \phi \vdash \phi \vdash \phi \vee \psi$, and $\tau \wedge \psi \vdash \psi \vdash \phi \vee \psi$. Also, by E$\wedge$ we have $\tau \wedge \phi \vdash \tau$, and $\tau \wedge \psi \vdash \tau$. Hence by the cut rule and I$\wedge$ we get $\tau \wedge \phi \vdash \tau \wedge (\phi \vee \psi)$, and $\tau \wedge \psi \vdash \tau \wedge (\phi \vee \psi)$. Thus by E$\vee$ we obtain

$$(\tau \wedge \phi) \vee (\tau \wedge \psi) \vdash \tau \wedge (\phi \vee \psi).$$

Conversely, by I$\wedge$ and I$\vee$ we have $\tau, \phi \vdash \tau \wedge \phi \vdash (\tau \wedge \phi) \vee (\tau \wedge \psi)$, and $\tau, \psi \vdash \tau \wedge \psi \vdash (\tau \wedge \phi) \vee (\tau \wedge \psi)$. Hence by E$\vee$ we get $\tau, \phi \vee \psi \vdash (\tau \wedge \phi) \vee (\tau \wedge \psi)$. Therefore by E$\wedge$ and the cut rule we obtain

$$\tau \wedge (\phi \vee \psi) \vdash (\tau \wedge \phi) \vee (\tau \wedge \psi).$$

(iv) By I$\wedge$ we know that $\phi, \phi \vdash \phi \wedge \phi$. Thus by the contraction rule we get $\phi \vdash \phi \wedge \phi$. On the other hand, by E$\wedge$ we have $\phi \wedge \phi \vdash \phi$.

Next, by I$\vee$ we have $\phi \vdash \phi \vee \phi$. Conversely, we know that $\phi \vdash \phi$. If we use this entailment twice, and apply E$\vee$, we get $\phi \vee \phi \vdash \phi$.

(v) We know that $\phi \vdash \phi$. Also by I$\vee$ we have $\phi \vdash \phi \vee \psi$. Thus by the cut rule and I$\wedge$ we get $\phi \vdash \phi \wedge (\phi \vee \psi)$. Conversely, by E$\wedge$ we have $\phi \wedge (\phi \vee \psi) \vdash \phi$.

Next, by I$\vee$ we have $\phi \vdash \phi \vee (\phi \wedge \psi)$. On the other hand, we know that $\phi \vdash \phi$. Also by E$\wedge$ we have $\phi \wedge \psi \vdash \phi$. Hence by E$\vee$ we get $\phi \vee (\phi \wedge \psi) \vdash \phi$.

(vi) By the weakening rule and E$\wedge$ we have $\neg\phi, \phi \wedge \psi \vdash \phi$. We also have $\neg\phi, \phi \wedge \psi \vdash \neg\phi$. Hence by the cut rule and E$\neg$ we get $\neg\phi, \phi \wedge \psi \vdash \bot$. Thus by I$\neg$ we obtain $\neg\phi \vdash \neg(\phi \wedge \psi)$. Similarly we have $\neg\psi \vdash \neg(\phi \wedge \psi)$. Therefore by E$\vee$ we get

$$\neg\phi \vee \neg\psi \vdash \neg(\phi \wedge \psi).$$

On the other hand, by the weakening rule and I$\vee$ we have $\neg(\neg\phi\vee\neg\psi), \neg\phi \vdash \neg\phi\vee\neg\psi$. Thus by the cut rule and E$\neg$ we get

$$\neg(\neg\phi \vee \neg\psi), \neg\phi \vdash \bot.$$

Hence by RAA we get $\neg(\neg\phi \vee \neg\psi) \vdash \phi$. Similarly we obtain $\neg(\neg\phi \vee \neg\psi) \vdash \psi$. Thus by the cut rule and I$\wedge$ we get $\neg(\neg\phi \vee \neg\psi) \vdash \phi \wedge \psi$. Therefore by the weakening and cut rules, and E$\neg$, we obtain

$$\neg(\phi \wedge \psi), \neg(\neg\phi \vee \neg\psi) \vdash \bot.$$

Hence by RAA we get $\neg(\phi \wedge \psi) \vdash \neg\phi \vee \neg\psi$, as desired.

Next, by the weakening rule and I$\vee$ we have $\neg(\phi \vee \psi), \phi \vdash \phi \vee \psi$. Hence by the cut rule and E$\neg$ we get $\neg(\phi \vee \psi), \phi \vdash \bot$. Thus by I$\neg$ we get $\neg(\phi \vee \psi) \vdash \neg\phi$. Similarly we obtain $\neg(\phi \vee \psi) \vdash \neg\psi$. Therefore by the cut rule and I$\wedge$ we get

$$\neg(\phi \vee \psi) \vdash \neg\phi \wedge \neg\psi.$$

Conversely, by the weakening rule and E$\wedge$ we have $\neg\phi \wedge \neg\psi, \phi \vdash \neg\phi$. Hence by the cut rule and E$\neg$ we get $\neg\phi \wedge \neg\psi, \phi \vdash \bot$. Thus by I$\neg$ we get

$$\phi \vdash \neg(\neg\phi \wedge \neg\psi).$$

Similarly we obtain $\psi \vdash \neg(\neg\phi \wedge \neg\psi)$. Therefore by E$\vee$ we get $\phi \vee \psi \vdash \neg(\neg\phi \wedge \neg\psi)$. Hence by the weakening and cut rules, and E$\neg$, we obtain

$$\neg\phi \wedge \neg\psi, \phi \vee \psi \vdash \bot.$$

Thus by I$\neg$ we get $\neg\phi \wedge \neg\psi \vdash \neg(\phi \vee \psi)$, as desired.

(vii) By E$\neg$ we have $\phi, \neg\phi \vdash \bot$. Hence by I$\neg$ we get $\phi \vdash \neg\neg\phi$. On the other hand, by the exchange rule and E$\neg$ we have $\neg\neg\phi, \neg\phi \vdash \bot$. Thus by RAA we get $\neg\neg\phi \vdash \phi$, as desired.

(viii) By the weakening rule we have $\phi, \psi \vdash \psi$. Also, by E$\neg$ and the EFQ rule (stated in Theorem 1.1) we have $\phi, \neg\phi \vdash \bot \vdash \psi$. Hence by E$\vee$ we obtain $\phi, \neg\phi \vee \psi \vdash \psi$. Thus by the exchange rule and I$\to$ we get $\neg\phi \vee \psi \vdash \phi \to \psi$.

Conversely, by E$\to$ and I$\vee$ we have $\phi \to \psi, \phi \vdash \psi \vdash \neg\phi \vee \psi$. Also, by the weakening rule and I$\vee$ we have $\phi \to \psi, \neg\phi \vdash \neg\phi \vdash \neg\phi \vee \psi$. Hence by E$\vee$ we get $\phi \to \psi, \phi \vee \neg\phi \vdash \neg\phi \vee \psi$. But we know that $\vdash \phi \vee \neg\phi$; so in particular we have $\phi \to \psi \vdash \phi \vee \neg\phi$, due to the weakening rule. Thus by the cut and contraction rules we get $\phi \to \psi \vdash \neg\phi \vee \psi$, as desired.

(ix) By modus tollens (stated in Theorem 1.1) we have $\phi \to \psi, \neg\psi \vdash \neg\phi$. Thus by I$\to$ we get

$$\phi \to \psi \vdash \neg\psi \to \neg\phi.$$

Conversely, note that if we repeat the above argument, and replace ϕ, ψ by $\neg\phi, \neg\psi$ respectively, we obtain $\neg\phi \to \neg\psi \vdash \neg\neg\psi \to \neg\neg\phi$. However, we have shown that $\neg\neg\phi \equiv \phi$ and $\neg\neg\psi \equiv \psi$. Hence by Theorem 1.4 we have $\neg\neg\psi \to \neg\neg\phi \equiv \psi \to \phi$. In particular we have $\neg\neg\psi \to \neg\neg\phi \vdash \psi \to \phi$. Thus by the cut rule we get $\neg\phi \to \neg\psi \vdash \psi \to \phi$, as desired.

(x) We have shown that $\phi \to \psi \equiv \neg\phi \vee \psi$. Thus by Theorem 1.4 and De Morgan's law we have $\neg(\phi \to \psi) \equiv \neg(\neg\phi \vee \psi) \equiv \neg\neg\phi \wedge \neg\psi$. However, we also know that $\neg\neg\phi \equiv \phi$. Therefore by using Theorem 1.4 again we obtain $\neg\neg\phi \wedge \neg\psi \equiv \phi \wedge \neg\psi$. Hence by the transitivity of $\equiv$ we get $\neg(\phi \to \psi) \equiv \phi \wedge \neg\psi$, as desired.

(xi) By the weakening rule and E$\to$ we have $\phi \to \psi, \psi \to \phi, \phi \vdash \psi$, and $\phi \to \psi, \psi \to \phi, \psi \vdash \phi$. Thus by I$\leftrightarrow$ we get $\phi \to \psi, \psi \to \phi \vdash \phi \leftrightarrow \psi$. Hence by E$\wedge$ and the cut rule we obtain

$$(\phi \to \psi) \wedge (\psi \to \phi) \vdash \phi \leftrightarrow \psi.$$

Conversely, by E$\leftrightarrow$ we have $\phi \leftrightarrow \psi, \phi \vdash \psi$, and $\phi \leftrightarrow \psi, \psi \vdash \phi$. Thus by I$\to$ we get $\phi \leftrightarrow \psi \vdash \phi \to \psi$, and $\phi \leftrightarrow \psi \vdash \psi \to \phi$. Therefore by the cut rule and I$\wedge$ we obtain

$$\phi \leftrightarrow \psi \vdash (\phi \to \psi) \wedge (\psi \to \phi).$$

Next, consider $\psi \leftrightarrow \phi$. By the above argument and the commutativity of $\wedge$ we have

$$\psi \leftrightarrow \phi \equiv (\psi \to \phi) \wedge (\phi \to \psi) \equiv (\phi \to \psi) \wedge (\psi \to \phi) \equiv \phi \leftrightarrow \psi.$$

The desired result follows from transitivity of $\equiv$. Finally consider $\neg\phi \leftrightarrow \neg\psi$. By the above argument, law of contraposition, and Theorem 1.4 we have

$$\begin{aligned} \neg\phi \leftrightarrow \neg\psi &\equiv (\neg\phi \to \neg\psi) \wedge (\neg\psi \to \neg\phi) \\ &\equiv (\psi \to \phi) \wedge (\phi \to \psi) \equiv \psi \leftrightarrow \phi \equiv \phi \leftrightarrow \psi. \end{aligned}$$

And again the desired result follows from transitivity of $\equiv$.

(xii) By the EFQ rule (stated in Theorem 1.1) we know that a contradiction implies any formula. So we have $\phi \wedge \neg\phi \vdash \bot$ and $\bot \vdash \phi \wedge \neg\phi$.

Next, note that by the law of excluded middle and the law of non-contradiction we respectively have $\vdash \phi \vee \neg\phi$ and $\vdash \top$. Hence by the weakening rule we get $\top \vdash \phi \vee \neg\phi$ and $\phi \vee \neg\phi \vdash \top$, as desired.

(xiii) By E$\wedge$ and I$\vee$ we respectively obtain

$$\phi \wedge \top \vdash \phi, \qquad \top \vdash \phi \vee \top.$$

Conversely, by the law of non-contradiction we have $\vdash \top$. Thus by the weakening rule we get

$$\phi \vee \top \vdash \top,$$

and $\phi \vdash \top$. We also know that $\phi \vdash \phi$. Therefore by the cut rule and I$\wedge$ we get $\phi \vdash \phi \wedge \top$.

Next, note that by I$\vee$ and E$\wedge$ we respectively obtain

$$\phi \vdash \phi \vee \bot, \qquad \phi \wedge \bot \vdash \bot.$$

On the other hand, by the EFQ rule (stated in Theorem 1.1) we know that

$$\bot \vdash \phi \wedge \bot,$$

and $\bot \vdash \phi$. We also know that $\phi \vdash \phi$. Therefore by E$\vee$ we get $\phi \vee \bot \vdash \phi$ as desired.

(xiv) We have $\phi \to \bot \equiv \neg\phi \vee \bot \equiv \neg\phi$. ■

Remark. Consider the formula $(\phi \wedge \psi) \wedge \tau$. We know that it is equivalent to $\phi \wedge (\psi \wedge \tau)$. Sometimes we abuse the notation, and denote these equivalent formulas simply by $\phi \wedge \psi \wedge \tau$. Similarly, we may write $\phi \vee \psi \vee \tau$ to denote the equivalent formulas $(\phi \vee \psi) \vee \tau$ and $\phi \vee (\psi \vee \tau)$. The associativity also implies that if several formulas are all connected by conjunction or disjunction, then the arrangement of parentheses between them does not alter the truth of the compound formula. In other words, all the possible arrangements of parentheses result in equivalent formulas. So for example, $\phi \vee ((\psi \vee \tau) \vee \sigma)$ is equivalent to $(\phi \vee \psi) \vee (\tau \vee \sigma)$. We may denote these equivalent formulas by $\phi \vee \psi \vee \tau \vee \sigma$. Similar abbreviated notations can be used when we have more formulas. However, we do not have the tools to state the general version of this fact precisely, and to prove it rigorously.

Remark. The equivalence $\phi \to \psi \equiv \neg\phi \vee \psi$ affirms the fact that $\phi \to \psi$ is only false if ϕ is true and ψ is false. Note that although the truth value of the material implication is not completely evident, we deduced the above equivalence from the inference rules which are intuitively more obvious. So we can consider this as an informal justification of the truth value of the material implication.

Also as we saw, the above equivalence implies that the negation of the conditional formula $\phi \to \psi$ is equivalent to $\phi \wedge \neg\psi$. Informally, this means that if ϕ does not imply ψ, then ϕ must be true while ψ is false.

Remark. The application of the law of contraposition in a deduction is also known as **proof by contraposition**. Namely, in order to show that $\phi \to \psi$ is true, sometimes it is easier to show that $\neg\psi \to \neg\phi$ is true. Then we get the desired by the law of contraposition. Informally, in this method, instead of showing that if ϕ is true then ψ must be true too, we show that if ψ is false then ϕ must be false too. Thus we can conclude that if ϕ is true then ψ cannot be false, so ψ must be true.

Remark. The equivalence $\phi \leftrightarrow \psi \equiv (\phi \to \psi) \wedge (\psi \to \phi)$ confirms our initial intuition about biconditional statements. It also implies that if we have $\vdash \phi \leftrightarrow \psi$ then by the cut rule and E$\wedge$ we also have $\vdash \phi \to \psi$ and $\vdash \psi \to \phi$. But as we saw before, this means that $\phi \vdash \psi$ and $\psi \vdash \phi$. Hence $\phi \equiv \psi$. We have seen that the converse of this fact holds too. Therefore we get

$$\phi \equiv \psi \qquad \text{if and only if} \qquad \vdash \phi \leftrightarrow \psi.$$

Thus, we could have also used the above as the definition of the equivalence of two formulas.

Remark. In the law of material implication, we have seen that we can express $\to$ in terms of $\vee, \neg$. More generally, it is possible to express each connective in terms of the other connectives. To see this, we just need to replace ϕ, ψ by $\neg\phi, \neg\psi$ in some parts of the above theorem, and use the double negation law and Theorem 1.4, to conclude

$$\begin{aligned} \phi \wedge \psi &\equiv \neg(\neg\phi \vee \neg\psi) &&\equiv \neg(\phi \to \neg\psi), \\ \phi \vee \psi &\equiv \neg(\neg\phi \wedge \neg\psi) &&\equiv \neg\phi \to \psi \equiv \neg\psi \to \phi, \\ \phi \to \psi &\equiv \neg(\phi \wedge \neg\psi) &&\equiv \neg\phi \vee \psi. \end{aligned}$$

We have also seen that we can express $\neg$ in terms of $\to, \bot$.

Also note that the above equivalences for $\phi \vee \psi$ give us useful tools for proving it. Namely, as explained in the following example, in order to prove $\phi \vee \psi$ we assume that one of ϕ, ψ is false, and then we show that the other one must be true.

***Example* 1.6.** Suppose Γ denotes a collection of several formulas, which can be empty too. Let ϕ, ψ be formulas. Then we have

$$\text{If} \quad \Gamma, \neg\phi \vdash \psi \qquad \text{then} \qquad \Gamma \vdash \phi \vee \psi.$$

Because by I$\to$ we have $\Gamma \vdash \neg\phi \to \psi$. Hence by the law of material implication we get $\Gamma \vdash \neg\neg\phi \vee \psi$. Finally, by the double negation law and Theorem 1.4 we obtain $\Gamma \vdash \phi \vee \psi$, as desired.

Theorem 1.7. *Suppose ϕ, ψ, τ are formulas. Then we have*

(i)
$$(\phi \wedge \psi) \to \tau \equiv \phi \to (\psi \to \tau) \equiv \psi \to (\phi \to \tau).$$

(ii)
$$\tau \to (\phi \wedge \psi) \equiv (\tau \to \phi) \wedge (\tau \to \psi).$$

(iii)
$$(\phi \vee \psi) \to \tau \equiv (\phi \to \tau) \wedge (\psi \to \tau).$$

(iv)
$$\tau \to (\phi \vee \psi) \equiv (\tau \to \phi) \vee (\tau \to \psi).$$

Proof. **(i)** By I$\wedge$ and E$\to$ we know that $\phi, \psi \vdash \phi \wedge \psi$ and $\phi \wedge \psi \to \tau, \phi \wedge \psi \vdash \tau$. Thus by the cut and exchange rules we get $\phi \wedge \psi \to \tau, \phi, \psi \vdash \tau$. Now by applying I$\to$ twice we obtain $\phi \wedge \psi \to \tau, \phi \vdash \psi \to \tau$, and therefore

$$\phi \wedge \psi \to \tau \vdash \phi \to (\psi \to \tau).$$

Conversely, by applying E$\to$ twice we get $\phi \to (\psi \to \tau), \phi \vdash \psi \to \tau$, and $\psi \to \tau, \psi \vdash \tau$. Thus by the cut and exchange rules we get $\phi \to (\psi \to \tau), \phi, \psi \vdash \tau$. Now by the cut rule and E$\wedge$ we obtain $\phi \to (\psi \to \tau), \phi \wedge \psi \vdash \tau$. Hence by I$\to$ we get

$$\phi \to (\psi \to \tau) \vdash \phi \wedge \psi \to \tau.$$

Finally note that by the above argument, commutativity of $\wedge$, and Theorem 1.4 we have

$$\psi \to (\phi \to \tau) \equiv \psi \wedge \phi \to \tau \equiv \phi \wedge \psi \to \tau \equiv \phi \to (\psi \to \tau).$$

The desired result follows from transitivity of $\equiv$.

(ii) By the law of material implication, distributivity of $\vee$ over $\wedge$, and Theorem 1.4 we have

$$\tau \to (\phi \wedge \psi) \equiv \neg\tau \vee (\phi \wedge \psi) \equiv (\neg\tau \vee \phi) \wedge (\neg\tau \vee \psi) \equiv (\tau \to \phi) \wedge (\tau \to \psi).$$

(iii) By the law of material implication, De Morgan's law, distributivity of $\vee$ over $\wedge$, and Theorem 1.4 we have

$$\begin{aligned}(\phi \vee \psi) \to \tau &\equiv \neg(\phi \vee \psi) \vee \tau \equiv (\neg\phi \wedge \neg\psi) \vee \tau \\ &\equiv (\neg\phi \vee \tau) \wedge (\neg\psi \vee \tau) \equiv (\phi \to \tau) \wedge (\psi \to \tau).\end{aligned}$$

(iv) By the law of material implication; idempotency, commutativity, and associativity of $\vee$; and Theorem 1.4 we have

$$\begin{aligned}\tau \to (\phi \vee \psi) &\equiv \neg\tau \vee (\phi \vee \psi) \equiv (\neg\tau \vee \neg\tau) \vee (\phi \vee \psi) \\ &\equiv \neg\tau \vee (\neg\tau \vee (\phi \vee \psi)) \equiv \neg\tau \vee ((\neg\tau \vee \phi) \vee \psi) \\ &\equiv (\neg\tau \vee (\neg\tau \vee \phi)) \vee \psi \equiv ((\neg\tau \vee \phi) \vee \neg\tau) \vee \psi \\ &\equiv (\neg\tau \vee \phi) \vee (\neg\tau \vee \psi) \equiv (\tau \to \phi) \wedge (\tau \to \psi).\end{aligned}$$

■

1.4 Variables and their Substitution

Primitive Notion 1.4. A variable can be a **variable in a formula**. This relation between variables and formulas is a primitive notion. But informally it means that the variable has appeared in the formula. There are two kinds of variables in a formula, namely **bound variables** and **free variables**. These are also primitive notions. We will provide their intuitive meanings after the next axiom.

Notation. Suppose ϕ is a formula, and $x_1, \ldots, x_n$ are its free variables. Then to denote this, we write

$$\phi(x_1, \ldots, x_n).$$

Note that $x_1, \ldots, x_n$ are not necessarily all the free variables of ϕ. We just want to emphasize that they are among the free variables of ϕ. For example we can just write $\phi(x_1)$ to state that x_1 is a free variable in ϕ.

Remark. Note that in the above notation, ϕ can be any formula. Also, $x_1, \ldots, x_n$ are just a notation for several variables, and any other collection of variables can be used in their place.

Axiom 1.4.

(i) *$\bot$ has no free or bound variables.*

(ii) *For every variable like x, the free variable of the formulas*

$$x = x, \qquad x \in x,$$

is x. Also for every distinct variables like x, y, the free variables of the formulas

$$x = y, \qquad x \in y,$$

are x, y. These formulas do not have any bound variables.

(iii) *Let ϕ be a formula. Then the free and bound variables of $\neg(\phi)$ are the free and bound variables of ϕ respectively.*

(iv) *Let ϕ, ψ be two formulas, that are not necessarily distinct. Consider the following formulas*

$$(\phi) \wedge (\psi), \quad (\phi) \vee (\psi), \quad (\phi) \to (\psi), \quad (\phi) \leftrightarrow (\psi).$$

A variable is a free variable of any of the above formulas if and only if it is a free variable of ϕ or a free variable of ψ. In other words, the collection of free variables of any of the above formulas is the union of the collection of free variables of ϕ and the collection of free variables of ψ. Similarly, a variable is a bound variable of any of the above formulas if and only if it is a bound variable of ϕ or a bound variable of ψ.

(v) *Let ϕ be a formula, and let x be a variable. Consider the following formulas*

$$\forall x(\phi), \qquad \exists x(\phi).$$

A variable is a free variable of any of the above formulas if and only if it is a free variable of ϕ, and it is not x. Also, a variable is a bound variable of any of the above formulas if and only if it is x, or it is a bound variable of ϕ.

Remark. The above axiom provides sufficient tools for finding the free and bound variables of every formula. We cannot prove this fact rigorously now, but informally it should be evident, since there is a rule for finding the free and bound variables corresponding to every rule for constructing a formula. Therefore in practice we can always find the free and bound variables of any formula we encounter.

Intuitively, a variable like x is a bound variable in a formula if and only if $\forall x$ and/or $\exists x$ appear somewhere in the formula. All other variables which appear in the formula are free. Note that a variable can be both free and bound in the same formula. For example, x is both free and bound in the following formulas:

$$(x \in y) \wedge (\forall x(x = x)), \qquad (\exists x(x \in y)) \rightarrow (\neg(x = y)).$$

This undesired phenomenon results from the fact that we do not impose any condition on two formulas like ϕ, ψ, when we combine them to construct other formulas like $\phi \vee \psi$. It is possible to exclude such expressions from being formulas, but it is easier to allow them to be formulas at this stage, and later develop some rules to deal with them.

Remark. We can rewrite the above formulas as follows:

$$(x \in y) \wedge (\forall z(z = z)), \qquad (\exists z(z \in y)) \rightarrow (\neg(x = y)).$$

Intuitively, these formulas have the same meaning as the previous ones. Also, no variable is both free and bound in them. We can always change the bound occurrences of a variable in a formula to produce another formula, as we did above, so that no variable is both free and bound in the new formula. We will not define this process rigorously, since we do not use it. Also, we do not have the mathematical tools to prove the above fact in general here. But we can easily check that it is true in every formula we encounter.

Another point to mention is that a variable can be the bound variable of several quantifiers in a formula. For example y in the formula

$$\forall x \exists y(x \in y) \vee \forall z \forall y(z \in y \rightarrow z \neq y)$$

is of this type. A related notion to this phenomenon is the notion of the **scope** of a quantifier in a formula. We will not try to define this notion precisely here, and

we do not need its precise definition now. But informally it indicates the part of the formula over which the quantifier has effect. For example in the above formula, the scope of $\forall z$ is $\forall y(z \in y \to z \neq y)$, and the scope of $\exists y$ is $x \in y$. More explicitly, when we use a formula like ϕ to construct another formula like $\forall x\phi$, then ϕ is the scope of $\forall x$. But if we use the formula $\forall x\phi$ to construct another formula like $\forall x\phi \to \psi \wedge \exists x\tau$. Then the scope of $\forall x$ does not change, and is still ϕ.

***Example* 1.7.** Consider the formula

$$\exists x \forall y(x = y \to y \in z) \wedge \big((x = a \leftrightarrow x \in v) \vee \forall z \exists x(x \neq z \to z \neq x)\big).$$

The free variables of this formula are z, x, a, v, and its bound variables are x, y, z. The scope of $\forall y$ is $x = y \to y \in z$, and the scope of $\forall z$ is $\exists x(x \neq z \to z \neq x)$. There are two occurrences of $\exists x$ in the formula. The scope of the first $\exists x$ is $\forall y(x = y \to y \in z)$, and the scope of the second $\exists x$ is $x \neq z \to z \neq x$.

Remark. In the above formula x occurs as a bound variable once in the scope of the first $\exists x$, and again in the scope of the second $\exists x$. In addition, x occurs as a free variable in $x = a$ and $x \in v$. We will not try to make the notions of *free and bound occurrences* precise, but the above example should be sufficient to convey their meanings. Similarly, we have several occurrences of $\exists x$. We will not try to give an exact meaning to this either. Mainly because we do not need, and will not use, the exact meaning of these notions. An informal and intuitive understanding of them is sufficient for our purposes.

Finally let us mention some other anomalies that can happen when we add quantifiers to a formula. We do not require x to be a free variable in ϕ, or even a variable in ϕ, when we construct $\forall x\phi$ or $\exists x\phi$. Thus for example $\exists a(x = y)$ is a well-formed formula, even though $\exists a$ is superfluous here. Even worse is the case that a variable is bounded successively by two quantifiers, like $\forall x(\exists x(x = y))$.

Remark. As another example consider

$$\forall x(x \neq z \wedge \exists x(x = y)).$$

Here x is bounded by two quantifiers successively, but unlike the previous example the outer quantifier is not superfluous here. The usual interpretation of these kinds of formulas is to change the inner x to another bound variable. So for example, we can rewrite the above formula as follows:

$$\forall x(x \neq z \wedge \exists u(u = y)).$$

We will not need the apparatus to deal with these kinds of formulas, so we will not introduce it rigorously here.

Although the above formulas are problematic, and sometimes meaningless, they do not create much problem for us, and we can deal with them by introducing some simple rules later. So, similarly to most texts, we will not exclude them from being formulas.

Primitive Notion 1.5. The notion of **substitution** of free occurrences of a variable in a formula (if any) by another variable is a primitive notion. Intuitively, it means that if x is a free variable in a formula ϕ, and y is a variable which is not bound in ϕ, then we replace every free occurrence of x in the written representation of ϕ by y, and we obtain a new formula which we denote by

$$\phi[y/x].$$

Note that x need not be a free variable in ϕ; and in this case we intuitively know that $\phi[y/x]$ is the same formula as ϕ. Also note that x, y can be the same variable.

Remark. Note that y can be a free variable in ϕ before substitution. But y cannot be a bound variable in ϕ. Also note that if x is a variable which is both free and bound in ϕ, then when we substitute it with y, we only substitute the occurrences of x as a free variable, and we do not change the bound occurrences of x.

Remark. The reason that we do not allow y to be a bound variable in ϕ is that otherwise the truth of ϕ can change after substitution. For example consider

$$\exists y(x \neq y).$$

This formula says that there is a set y different than x, which is a true statement in the standard universe of sets. However if we replace x by y we get $\exists y(y \neq y)$. Now this new formula says that there is a set y that is not equal to itself, which is obviously a false statement.

Axiom 1.5. *Suppose ϕ is a formula, and x is a variable which is not a free variable in ϕ. Let y be a variable which is not bound in ϕ. Then $\phi[y/x]$ is the same formula as ϕ.*

Remark. As a consequence, $\perp$ does not change under substitution of variables, since $\perp$ has no free variables.

Axiom 1.6. *Suppose ϕ is a formula, x is a variable, and y is a variable which is not bound in ϕ. Then $\phi[y/x]$ is a formula whose bound variables are exactly the bound variables of ϕ. Also, if x is a free variable of ϕ, then a variable is a free variable of $\phi[y/x]$ if and only if it is a free variable of ϕ other than x, or it is y.*

Remark. Note that if x is not a free variable of ϕ, then $\phi[y/x]$ is the same formula as ϕ. Hence, in this case, a variable is a free variable of $\phi[y/x]$ if and only if it is a free variable of ϕ. Thus in particular, x is not a free variable of $\phi[y/x]$ either. This observation, and the above axiom, imply that x is never a free variable in $\phi[y/x]$, regardless of whether x is a free variable in ϕ or not.

Notation. Suppose $\phi(x)$ is a formula that has x as a free variable, and y is a variable which is not bound in ϕ. Then we sometimes denote $\phi[y/x]$ simply by $\phi(y)$. Note that this usage of the notation $\phi(\cdot)$ is compatible with the previous one, since in this case, y is a free variable in $\phi[y/x]$.

Axiom 1.7. *Suppose ϕ is a formula, and x is a variable which is not a bound variable of ϕ. Also suppose that y is a variable which is not a free or bound variable of ϕ. Let us denote $\phi[y/x]$ by ψ. Then we have*

(i) *$\phi[x/x]$ is the same formula as ϕ.*

(ii) *$\psi[x/y]$ is the same formula as ϕ.*

Remark. Note that we are allowed to substitute x for y in ψ, since x is not bound in ψ. Because the bound variables of ψ are the same as ϕ, and we assumed that x is not bound in ϕ. For the same reason, we are allowed to substitute x for x in ϕ.

Remark. The above axiom says that if we substitute a variable, and then substitute that variable back, we arrive at the original formula. It also says that if we substitute a variable with itself, the formula does not change. These facts are intuitively obvious, but we do not have the mathematical tools to prove them, so we accept them as axioms.

Remark. Note that if y is allowed to be a free variable in ϕ, then $\psi[x/y]$ and ϕ will be different; because y is not a free variable in $\psi[x/y]$, while it is a free variable in ϕ.

Axiom 1.8. *Suppose ϕ, ψ are formulas, x, z are variables, and y is a variable which is not bound in ϕ, ψ.*

(i) *If ϕ is one of the formulas $x = x$ or $x \in x$, then $\phi[y/x]$ is respectively*

$$y = y, \qquad or \qquad y \in y.$$

Also if ϕ is one of the formulas $x = z$ or $x \in z$, where x, z are distinct, then $\phi[y/x]$ is respectively

$$y = z, \qquad or \qquad y \in z.$$

Note that in this case y, z can also be the same variable. Similarly, if ϕ is one of the formulas $z = x$ or $z \in x$, where x, z are distinct, then $\phi[y/x]$ is respectively

$$z = y, \qquad or \qquad z \in y.$$

Also, in this case y, z can be the same variable too.

(ii) *The formula $(\neg\phi)[y/x]$ is*

$$\neg(\phi[y/x]).$$

(iii) *The following formulas*

$$(\phi \wedge \psi)[y/x], \quad (\phi \vee \psi)[y/x], \quad (\phi \to \psi)[y/x], \quad (\phi \leftrightarrow \psi)[y/x],$$

are respectively

$$(\phi[y/x]) \wedge (\psi[y/x]), \qquad (\phi[y/x]) \vee (\psi[y/x]),$$
$$(\phi[y/x]) \to (\psi[y/x]), \qquad (\phi[y/x]) \leftrightarrow (\psi[y/x]).$$

(iv) *If z is a variable different from x, y, then the formulas*

$$(\forall z\phi)[y/x], \qquad (\exists z\phi)[y/x],$$

are respectively

$$\forall z(\phi[y/x]), \qquad \exists z(\phi[y/x]).$$

Remark. Note that by the above axiom and Axiom 1.5 we can perform the substitution on any atomic formula. Because $\perp$ does not change under substitution as it does not have any variable. In addition, any atomic formula of the form $w = z$ or $w \in z$ is either considered in the above axiom (if one or both of w, z is the same as x), or it does not contain x as a free variable (if w, z are different from x), in which case the formula does not change under substitution by Axiom 1.5.

Remark. Note that in the last part of the above axiom, if we had $\forall y\phi$ or $\exists y\phi$, then we could not substitute y for x, because by Axiom 1.4, y is a bound variable in these formulas. In addition, if we had $\forall x\phi$ or $\exists x\phi$, then x is not a free variable in these formulas; therefore, by Axiom 1.5, these formulas do not change under the substitution for x.

Remark. Note that by Axiom 1.4, if y is not a bound variable in ϕ, ψ, then it is not a bound variable in the following formulas

$$\neg\phi, \quad \phi \wedge \psi, \quad \phi \vee \psi, \quad \phi \to \psi, \quad \phi \leftrightarrow \psi, \quad \forall z\phi, \quad \exists z\phi.$$

Therefore, in the above axiom, the requirements for substitution are satisfied.

Remark. The above axiom and Axiom 1.5 provide sufficient tools to find $\phi[y/x]$ for every ϕ which has the necessary properties. We cannot prove this fact rigorously now, but informally it should be evident, since there is a rule for constructing $\phi[y/x]$ corresponding to every rule for constructing a formula. Therefore in practice we can always find $\phi[y/x]$ for any formula ϕ that we encounter.

***Example* 1.8.** Suppose ϕ is the formula

$$\forall z \forall u(z \in a \wedge u \in z \to u \in c) \wedge \forall y(y \in x \to \exists z(z \in a \wedge y \in z)) \leftrightarrow \forall a(a = a).$$

Then $\phi[c/a]$ is

$$\forall z \forall u(z \in c \wedge u \in z \to u \in c) \wedge \forall y(y \in x \to \exists z(z \in c \wedge y \in z)) \leftrightarrow \forall a(a = a).$$

Note that the bound occurrence of a in $\forall a(a = a)$ does not change. Also note that c is a free variable in ϕ before substitution, but this does not prevent us from substituting it for a.

1.5 Rules of Inference for Quantifiers

Let us recall that we write $\phi(x)$ to emphasize that x is a free variable in the formula ϕ. Also, if y is not a bound variable in ϕ, then we write $\phi(y)$ as a shorthand notation for $\phi[y/x]$. Note that this usage of the notation $\phi(\cdot)$ is compatible with the previous one, since in this case, y is a free variable in $\phi[y/x]$. In addition, remember that when x is not a free variable in ϕ, then $\phi[y/x]$ is the same formula as ϕ.

Axiom 1.9 (Rules of inference for quantifiers)**.** *Suppose Γ denotes a collection of several formulas, which can be empty too. Let ϕ, ψ be formulas, and let x, y be variables. Then we have*

(i) *Introduction of $\forall$, or Universal generalization* :

Suppose x is not a free variable in any of the formulas in Γ.
Then we have : *If* $\Gamma \vdash \phi$ *then* $\Gamma \vdash \forall x \phi$.

(ii) *Elimination of $\forall$, or Universal instantiation* :

$$\forall x \phi \vdash \phi,$$
and if y is not a bound variable in ϕ,
then we also have : $\forall x \phi \vdash \phi[y/x]$.

(iii) *Introduction of $\exists$, or Existential generalization* :

$$\phi \vdash \exists x \phi,$$
and if y is not a bound variable in ϕ,
then we also have : $\phi[y/x] \vdash \exists x \phi$.

(iv) *Elimination of $\exists$* :

Suppose x is not a free variable in ψ, nor in any of the formulas in Γ.
Then we have : *If* $\Gamma, \phi \vdash \psi$ *then* $\Gamma, \exists x \phi \vdash \psi$.

Remark. Note that we do not require x to be a free variable in ϕ, although this is the case that we are actually interested in. The reason is that we need the more general version of the axiom in order to be able to show that $\forall x, \exists x$ are redundant in $\forall x\phi, \exists x\phi$, when x is not a free variable in ϕ. See Theorem 1.9.

Notation. We will use shorthand notations for the inference rules, as we did in the last section. For example, the introduction of $\forall$ will be denoted by I$\forall$, and the elimination of $\exists$ will be denoted by E$\exists$.

Let us inspect the above rules more closely. Let us only consider the meaningful case where x is a free variable in ϕ. The I$\forall$, or *universal generalization*, says that if we can deduce $\phi(x)$, then we can also deduce $\forall x\phi(x)$. In other words, if $\phi(x)$ is true for an arbitrary x, then $\phi(x)$ is true for every x, i.e. $\forall x\phi(x)$ is true. This is how universal statements are usually proved in mathematics; we prove the statement for an arbitrary object, and then conclude that the statement must hold for every object. Intuitively, the reason is that when we prove the statement without assuming anything specific about the object, then the same reasoning can work for any other object, hence the statement is true for every object.

Note that the important part of the above reasoning is that we do not assume anything specific about the object. To incorporate this into the rule I$\forall$, we supposed that x is not a free variable in any of the formulas in Γ. In other words, we supposed that we are not stating anything specific about x in the premises. Furthermore, from a formal syntactic viewpoint, if we had allowed x to be a free variable in some of the premises, then we could come to conclusions which are obviously false. For example from the formula $x \in y$, which says that the set y contains some set x, we could deduce $\forall x(x \in y)$, which says that y contains every set!

Remark. Note that in particular, we cannot deduce $\forall x\phi(x)$ from $\phi(x)$, when x is a free variable in ϕ.

The E$\forall$, or *universal instantiation*, says that if $\forall x\phi(x)$ is true, i.e. if $\phi(x)$ is true for every x, then in particular $\phi(y)$ is true. Similarly, the I$\exists$, or *existential generalization*, says that if $\phi(y)$ is true for some y, then $\exists x\phi(x)$ is true. The only restriction is that y cannot be bound in ϕ, so that we can substitute y for x in ϕ. A question that arises is that why did we also include $\phi[y/x]$ in these rules, in addition to $\phi(x)$? The above intuitive explanations of the rules suggest that allowing $\phi[y/x]$ is semantically preferable, but there are also some technical reasons behind this choice. First, these rules allow us to prove results about substitution of variables in formulas.

The second reason is that sometimes we need to only quantify over some occurrences of a variable, not all of them. For example from $y = y$ we can deduce $\exists x(x = y)$. Because we can consider $y = y$ as $x = y$, in which we substituted y for x. But if we stated the rules only for $\phi(x)$, then from $y = y$ we could only deduce that $\exists y(y = y)$. Note that although both conclusions are true, they are

stating different facts. Another point that should be noted is that we can consider ϕ as a special case of $\phi[y/x]$, namely we can consider ϕ as $\phi[x/x]$. Intuitively, it is obvious that if we substitute x with x, the formula ϕ does not change. However, if we want to treat this fact rigorously using the Axiom 1.7, we have to assume that x is not bound in ϕ, which is an unnecessary restriction. Although we can avoid this restriction by stating more axioms about substitution of variables, we prefer to avoid these complications altogether. Therefore we separated the case of ϕ in the rules E$\forall$ and I$\exists$.

Remark. An important assumption implicit in the rule I$\exists$ is that we are tacitly assuming that at least one set exists. Because otherwise we could not deduce that "there is a set x such that $\phi(x)$ holds", i.e. we could not deduce $\exists x\phi(x)$.

Finally, let us consider E$\exists$. It says that if we can deduce ψ from $\phi(x)$, then we can also deduce ψ from $\exists x\phi(x)$. In other words, if we can deduce that ψ holds by knowing that a particular set like x satisfies ϕ, then we can also deduce that ψ holds simply by knowing that there is a set that satisfies ϕ, i.e. by knowing that $\exists x\phi(x)$. Intuitively, in order for this argument to be valid, we cannot assume anything specific about x. Also, we cannot deduce anything specific about x. Therefore we require that x does not appear as a free variable in the conclusion ψ, nor in any of the formulas in the premises Γ. Hence in deducing ψ, the only property of x that we used is that it satisfies ϕ. Thus, just knowing that some set satisfies ϕ must be sufficient to deduce ψ.

In addition, from a formal syntactic viewpoint, if we had allowed x to be a free variable in some of the premises, or in the conclusion, then we could obtain false results. For example we have $x \in y \vdash x \in y$. Hence we could deduce that $\exists x(x \in y) \vdash x \in y$. Thus by I$\forall$ we could get $\exists x(x \in y) \vdash \forall x(x \in y)$, since x is not a free variable in $\exists x(x \in y)$. Therefore from the fact that the set y contains some set x, we could deduce that y contains every set! As another example consider the entailment $y = x, x = z \vdash y = z$, which reflects the transitivity of equality. If we were allowed to apply E$\exists$ here, we could deduce that $y = x, \exists x(x = z) \vdash y = z$. Then by exchange rule, and a valid application of E$\exists$, we could deduce that $\exists x(y = x), \exists x(x = z) \vdash y = z$. But this entailment is obviously false, because just by knowing that there is a set equal to y, and there is a set equal to z, we certainly cannot deduce that y, z are the same set!

Remark. An inference rule related to E$\exists$, which we do not accept as a valid rule, is *existential instantiation*. It says that $\exists x\phi \vdash \phi[y/x]$, provided that y is not a bound variable in ϕ. In other words, it says that if there is a set for which ϕ holds, then ϕ holds for some set y. This rule is closely related to the process of naming; a name, y, is given to the set for which ϕ holds. For this reason, existential instantiation seems to be a natural rule of inference. However, if we want to accept it as a valid rule of inference, we have to impose complicated constraints on when

we can apply the other rules. Otherwise it leads us to wrong conclusions. For example, existential instantiation implies that $\exists x(x \in z) \vdash y \in z$. Hence by I$\forall$ we get $\exists x(x \in z) \vdash \forall y(y \in z)$, which cannot be a valid entailment as we saw before.

Therefore, in order to avoid the complications caused by existential instantiation, we do not include it in our valid rules of inference, and instead we use the elimination of $\exists$. Note that this choice does not limit us in proving theorems. Because if we want to deduce a formula ψ by assuming $\exists x\phi$, we can deduce ψ by assuming $\phi(x)$, and then conclude our desired entailment by applying E$\exists$. In other words, we can give the element which satisfies ϕ the name x, and then use x to deduce ψ. Finally we can use E$\exists$ to conclude that ψ can be deduced from $\exists x\phi$. This is how such proofs are usually carried out in mathematics. When in our assumptions we have a formula which says that an object with certain properties exists, i.e. we have a formula of the form $\exists x\phi$, we assume that some object satisfies that property, and we use that object to prove our desired results. The part of these proofs in which an object is picked is usually phrased like "Let x be an object such that $\phi(x)$ holds.".

Thus in practice there is not much difference between E$\exists$ and existential instantiation. The only difference is that in E$\exists$ we do not say that $\exists x\phi$ implies $\phi(x)$. Rather, we say that if we can deduce anything from $\phi(x)$, then we can also deduce it from $\exists x\phi$. Note that if we accept existential instantiation as a valid rule, then the last sentence can be obtained from it and the cut rule.

***Example* 1.9.** It is easy to see that for any formula ϕ we have

$$\forall x\phi \vdash \exists x\phi.$$

Because by E$\forall$ and I$\exists$ we have $\forall x\phi \vdash \phi \vdash \exists x\phi$. Hence we get the desired by the cut rule.

Theorem 1.8. *Suppose ϕ, ψ are formulas, and $\phi \equiv \psi$. Let x be a variable. Then we have*

(i) $\forall x\phi \equiv \forall x\psi$.
(ii) $\exists x\phi \equiv \exists x\psi$.

Proof. **(i)** By E$\forall$ and equivalence of ϕ, ψ we have $\forall x\phi \vdash \phi \vdash \psi$. Thus by I$\forall$ we get $\forall x\phi \vdash \forall x\psi$, since x is not a free variable in the premises, i.e. in $\forall x\phi$. Similarly we can show that $\forall x\psi \vdash \forall x\phi$. Hence we get the desired.

(ii) By I$\exists$ and equivalence of ϕ, ψ we have $\phi \vdash \psi \vdash \exists x\psi$. Thus by E$\exists$ we get $\exists x\phi \vdash \exists x\psi$, since x is not a free variable in the conclusion, i.e. in $\exists x\psi$. Similarly we can show that $\exists x\psi \vdash \exists x\phi$. Hence we get the desired. ■

As we have seen after Theorem 1.4, if a formula is composed of several other formulas, then we can replace some of the components by equivalent formulas to

obtain a formula equivalent to the original compound formula. The above theorem enables us to do this when the original compound formula also contains quantifiers. As before, we will not prove the general case of this fact here, rather, we demonstrate it by an example. Suppose $\phi \equiv \psi$ and $\tau \equiv \sigma$. Consider the following formula

$$\exists y(\phi \vee \phi_1) \leftrightarrow (\phi_2 \to \exists u \forall z \tau \wedge \phi_3).$$

We claim that it is equivalent to $\exists y(\psi \vee \phi_1) \leftrightarrow (\phi_2 \to \exists u \forall z \sigma \wedge \phi_3)$. To see this note that by the above theorem $\forall z\tau \equiv \forall z\sigma$, and therefore $\exists u \forall z\tau \equiv \exists u \forall z\sigma$. Also note that any formula is equivalent to itself. Therefore by Theorem 1.4 we have

$$\phi \vee \phi_1 \equiv \psi \vee \phi_1, \qquad \exists u \forall z \tau \wedge \phi_3 \equiv \exists u \forall z \sigma \wedge \phi_3.$$

Now if we apply the above theorem again we obtain $\exists y(\phi \vee \phi_1) \equiv \exists y(\psi \vee \phi_1)$. Finally, by Theorem 1.4 we obtain that $\phi_2 \to \exists u \forall z \tau \wedge \phi_3 \equiv \phi_2 \to \exists u \forall z \sigma \wedge \phi_3$, and therefore we get

$$\exists y(\phi \vee \phi_1) \leftrightarrow (\phi_2 \to \exists u \forall z \tau \wedge \phi_3) \equiv \exists y(\psi \vee \phi_1) \leftrightarrow (\phi_2 \to \exists u \forall z \sigma \wedge \phi_3),$$

as desired.

Theorem 1.9. *Suppose ϕ is a formula, and x is a variable which is not a free variable in ϕ. Then we have*

$$\forall x \phi \equiv \phi \equiv \exists x \phi.$$

Remark. This theorem shows that quantifying over variables which do not occur free is redundant.

Proof. We know that $\phi \vdash \phi$. Thus by I$\forall$ we have $\phi \vdash \forall x \phi$, since x is not a free variable in the premises, i.e. in ϕ. On the other hand, by E$\forall$ we have $\forall x \phi \vdash \phi$. Hence we have $\forall x \phi \equiv \phi$.

Similarly, by I$\exists$ we have $\phi \vdash \exists x \phi$. On the other hand since $\phi \vdash \phi$, and x is not a free variable in the conclusion ϕ, we can use E$\exists$ to obtain $\exists x \phi \vdash \phi$. Hence we also have $\phi \equiv \exists x \phi$. ■

Theorem 1.10. *Suppose ϕ, ψ are formulas, and x, y are variables. Then we have*

(i)
$$\neg(\exists x \phi) \equiv \forall x (\neg \phi).$$

(ii)
$$\neg(\forall x \phi) \equiv \exists x (\neg \phi).$$

(iii)
$$\exists x \exists y \phi \equiv \exists y \exists x \phi.$$

(iv)
$$\forall x \forall y \phi \equiv \forall y \forall x \phi.$$

(v)
$$\exists x \forall y \phi \vdash \forall y \exists x \phi.$$

(vi)
$$\exists x(\phi \vee \psi) \equiv \exists x \phi \vee \exists x \psi.$$

(vii)
$$\forall x(\phi \wedge \psi) \equiv \forall x \phi \wedge \forall x \psi.$$

(viii)
$$\exists x(\phi \wedge \psi) \vdash \exists x \phi \wedge \exists x \psi.$$

(ix)
$$\forall x \phi \vee \forall x \psi \vdash \forall x(\phi \vee \psi).$$

(x) *Suppose x is not a free variable in ϕ, then*
$$\exists x(\phi \wedge \psi) \equiv \phi \wedge \exists x \psi.$$

(xi) *Suppose x is not a free variable in ϕ, then*
$$\forall x(\phi \vee \psi) \equiv \phi \vee \forall x \psi.$$

Proof. **(i)** By the weakening rule and E$\forall$ we have $\forall x(\neg\phi), \phi \vdash \neg\phi$. Hence by E$\neg$ and the cut rule we have $\forall x(\neg\phi), \phi \vdash \bot$. Note that x is not a free variable in $\forall x(\neg\phi), \bot$ due to the Axiom 1.4. Now we can apply the E$\exists$ to conclude that $\forall x(\neg\phi), \exists x\phi \vdash \bot$. Finally by I$\neg$ we get

$$\forall x(\neg\phi) \vdash \neg(\exists x\phi).$$

Conversely, by the weakening rule and I$\exists$ we have $\neg(\exists x\phi), \phi \vdash \exists x\phi$. Hence by E$\neg$ and the cut rule we have $\neg(\exists x\phi), \phi \vdash \bot$. Thus by I$\neg$ we get $\neg(\exists x\phi) \vdash \neg\phi$. Now note that x is not a free variable in $\neg(\exists x\phi)$. Therefore by I$\forall$ we have

$$\neg(\exists x\phi) \vdash \forall x(\neg\phi),$$

as desired.

(ii) By the weakening rule and E$\forall$ we have $\forall x\phi, \neg\phi \vdash \phi$. Hence by E$\neg$ and the cut rule we have $\forall x\phi, \neg\phi \vdash \bot$. Note that $\forall x\phi, \bot$ do not have x as a free variable. Hence we can apply the E$\exists$ to conclude that $\forall x\phi, \exists x(\neg\phi) \vdash \bot$. Finally by exchange rule and I$\neg$ we get

$$\exists x(\neg\phi) \vdash \neg(\forall x\phi).$$

Conversely, note that by the previous part and E$\forall$ we have $\neg\exists x(\neg\phi) \vdash \forall x(\neg\neg\phi) \vdash \neg\neg\phi$. But by Theorem 1.6 we have $\neg\neg\phi \vdash \phi$. So by the cut rule we get $\neg\exists x(\neg\phi) \vdash \phi$. Now note that x is not a free variable in $\neg\exists x(\neg\phi)$. Therefore by I$\forall$ we have $\neg\exists x(\neg\phi) \vdash \forall x\phi$. Hence by E$\neg$ we obtain $\neg(\forall x\phi), \neg\exists x(\neg\phi) \vdash \bot$. Thus by RAA we get

$$\neg(\forall x\phi) \vdash \exists x(\neg\phi),$$

as desired.

(iii) By I$\exists$ we have $\phi \vdash \exists x\phi \vdash \exists y\exists x\phi$. Now note that x, y are not free variables in $\exists y\exists x\phi$ due to the Axiom 1.4. Hence by applying E$\exists$ twice we get $\exists y\phi \vdash \exists y\exists x\phi$, and $\exists x\exists y\phi \vdash \exists y\exists x\phi$. Similarly we can show that $\exists y\exists x\phi \vdash \exists x\exists y\phi$.

(iv) By E$\forall$ we have $\forall x\forall y\phi \vdash \forall y\phi \vdash \phi$. Now note that x, y are not free variables in $\forall x\forall y\phi$ due to the Axiom 1.4. Hence by applying I$\forall$ twice we get $\forall x\forall y\phi \vdash \forall x\phi$, and $\forall x\forall y\phi \vdash \forall y\forall x\phi$. Similarly we can show that $\forall y\forall x\phi \vdash \forall x\forall y\phi$.

(v) By E$\forall$ and I$\exists$ we have $\forall y\phi \vdash \phi \vdash \exists x\phi$. Now note that x is not a free variable in $\exists x\phi$. Hence by E$\exists$ we get $\exists x\forall y\phi \vdash \exists x\phi$. In addition, note that y is not a free variable in $\exists x\forall y\phi$. Thus by I$\forall$ we obtain $\exists x\forall y\phi \vdash \forall y\exists x\phi$, as desired.

(vi) By I$\exists$ and I$\vee$ we have $\phi \vdash \exists x\phi \vdash \exists x\phi \vee \exists x\psi$. Similarly we have $\psi \vdash \exists x\psi \vdash \exists x\phi \vee \exists x\psi$. Therefore by E$\vee$ we get $\phi \vee \psi \vdash \exists x\phi \vee \exists x\psi$. Now note that x is not a free variable in $\exists x\phi \vee \exists x\psi$. Thus by E$\exists$ we obtain

$$\exists x(\phi \vee \psi) \vdash \exists x\phi \vee \exists x\psi.$$

Conversely, by I$\vee$ and I$\exists$ we have $\phi \vdash \phi \vee \psi \vdash \exists x(\phi \vee \psi)$. Similarly we have $\psi \vdash \phi \vee \psi \vdash \exists x(\phi \vee \psi)$. Now note that x is not a free variable in $\exists x(\phi \vee \psi)$. Thus by E$\exists$ we obtain $\exists x\phi \vdash \exists x(\phi \vee \psi)$, and $\exists x\psi \vdash \exists x(\phi \vee \psi)$. Hence by E$\vee$ we get

$$\exists x\phi \vee \exists x\psi \vdash \exists x(\phi \vee \psi).$$

(vii) By E$\forall$ and E$\wedge$ we have $\forall x(\phi \wedge \psi) \vdash \phi \wedge \psi \vdash \phi$. Similarly we have $\forall x(\phi \wedge \psi) \vdash \phi \wedge \psi \vdash \psi$. Now note that x is not a free variable in $\forall x(\phi \wedge \psi)$. Thus by I$\forall$ we get $\forall x(\phi \wedge \psi) \vdash \forall x\phi$, and $\forall x(\phi \wedge \psi) \vdash \forall x\psi$. Hence by the cut rule and I$\wedge$ we obtain

$$\forall x(\phi \wedge \psi) \vdash \forall x\phi \wedge \forall x\psi.$$

Conversely, by E$\wedge$ and E$\forall$ we have $\forall x\phi \wedge \forall x\psi \vdash \forall x\phi \vdash \phi$. Similarly we have $\forall x\phi \wedge \forall x\psi \vdash \forall x\psi \vdash \psi$. Thus by the cut rule and I$\wedge$ we get $\forall x\phi \wedge \forall x\psi \vdash \phi \wedge \psi$. Now note that x is not a free variable in $\forall x\phi \wedge \forall x\psi$. Therefore by I$\forall$ we get

$$\forall x\phi \wedge \forall x\psi \vdash \forall x(\phi \wedge \psi),$$

as desired.

(viii) By E$\wedge$ and I$\exists$ we have $\phi\wedge\psi \vdash \psi \vdash \exists x\psi$. Similarly we have $\phi\wedge\psi \vdash \phi \vdash \exists x\phi$. Thus by the cut rule and I$\wedge$ we get $\phi \wedge \psi \vdash \exists x\phi \wedge \exists x\psi$. Now note that x is not a free variable in $\exists x\phi \wedge \exists x\psi$. Therefore by E$\exists$ we obtain

$$\exists x(\phi \wedge \psi) \vdash \exists x\phi \wedge \exists x\psi.$$

(ix) By E$\forall$ and I$\vee$ we have $\forall x\psi \vdash \psi \vdash \phi\vee\psi$. Similarly we have $\forall x\phi \vdash \phi \vdash \phi\vee\psi$. Thus by E$\vee$ we get $\forall x\phi \vee \forall x\psi \vdash \phi \vee \psi$. Now note that x is not a free variable in $\forall x\phi \vee \forall x\psi$. Hence by I$\forall$ we obtain

$$\forall x\phi \vee \forall x\psi \vdash \forall x(\phi \vee \psi).$$

(x) By E$\wedge$ and I$\exists$ we have $\phi\wedge\psi \vdash \psi \vdash \exists x\psi$. Similarly, by E$\wedge$ we have $\phi\wedge\psi \vdash \phi$. Thus by the cut rule and I$\wedge$ we get $\phi \wedge \psi \vdash \phi \wedge \exists x\psi$. Now note that x is not a free variable in $\phi \wedge \exists x\psi$. Therefore by E$\exists$ we obtain

$$\exists x(\phi \wedge \psi) \vdash \phi \wedge \exists x\psi.$$

Conversely, by I$\wedge$ and I$\exists$ we have $\phi, \psi \vdash \phi \wedge \psi \vdash \exists x(\phi \wedge \psi)$. Now note that x is not a free variable in ϕ and $\exists x(\phi \wedge \psi)$. Thus by E$\exists$ we get $\phi, \exists x\psi \vdash \exists x(\phi \wedge \psi)$. Hence by E$\wedge$ and the cut rule we obtain

$$\phi \wedge \exists x\psi \vdash \exists x(\phi \wedge \psi).$$

(xi) By the previous part we have

$$\neg\phi \wedge \exists x(\neg\psi) \equiv \exists x(\neg\phi \wedge \neg\psi);$$

because x is not a free variable in $\neg\phi$. Now by Theorem 1.4 we have

$$\neg(\neg\phi \wedge \exists x(\neg\psi)) \equiv \neg(\exists x(\neg\phi \wedge \neg\psi)).$$

Hence by De Morgan's law and part (i) of this theorem we get

$$\neg\neg\phi \vee \neg\exists x(\neg\psi) \equiv \forall x(\neg(\neg\phi \wedge \neg\psi)).$$

If we apply the De Morgan's law and part (i) of this theorem again, and use Theorems 1.4, 1.8, we obtain

$$\neg\neg\phi \vee \forall x(\neg\neg\psi) \equiv \forall x(\neg\neg\phi \vee \neg\neg\psi).$$

Finally, by double negation law, and Theorems 1.4, 1.8 we get

$$\phi \vee \forall x\psi \equiv \forall x(\phi \vee \psi),$$

as desired. ■

Remark. The equivalence $\neg(\exists x\phi) \equiv \forall x(\neg\phi)$ says that if there does not exists an x such that ϕ holds, then for every x, $\neg\phi$ must hold. Similarly, the equivalence $\neg(\forall x\phi) \equiv \exists x(\neg\phi)$ says that if it is not the case that for every x, ϕ holds, then there must exist an x such that $\neg\phi$ holds. In other words, if a universal statement is not valid, then a **counterexample** to it must exist.

Remark. As shown in the above theorem, the order of consecutive quantifiers can be changed, as long as they are of the same type. However, we cannot change the order of quantifiers of different types; because in general, $\forall x\exists y\phi$ does not imply $\exists y\forall x\phi$. For example, the formula $\forall x\exists y(y = x)$ says that for every set x there is a set y which is equal to it. Intuitively, this formula is true, since we can take y to be the same as x. But if we change the order of quantifiers we get the formula $\exists y\forall x(y = x)$, which says that there is a set y which is equal to every set x! And this formula is obviously false in the standard theory of sets.

Remark. In the first two parts of the above axiom, if we replace ϕ by $\neg\phi$, and use the fact that $\neg\neg\phi \equiv \phi$, then by Theorem 1.8 we obtain

$$\neg\exists x(\neg\phi) \equiv \forall x\phi, \qquad \neg\forall x(\neg\phi) \equiv \exists x\phi.$$

Therefore we can express each quantifier in terms of the other one.

Remark. In general, $\exists x\phi \wedge \exists x\psi$ does not imply $\exists x(\phi \wedge \psi)$; and $\forall x(\phi \vee \psi)$ does not imply $\forall x\phi \vee \forall x\psi$. Intuitively, these can be seen as follows. Suppose ϕ says that x has no element, and ψ says that x has some element. Then it is easy to check that both $\exists x\phi \wedge \exists x\psi$ and $\forall x(\phi \vee \psi)$ are true, while both $\exists x(\phi \wedge \psi)$ and $\forall x\phi \vee \forall x\psi$ are false.

Theorem 1.11. *Suppose ϕ is a formula, and x is a free variable in ϕ, which is not bound in ϕ. Let y be a variable which is not a free variable nor a bound variable in ϕ. Then we have*

$$\forall x\phi(x) \equiv \forall y\phi(y), \qquad \exists x\phi(x) \equiv \exists y\phi(y).$$

Remark. This theorem provides a tool for substituting bound variables in some cases. Note that we assume that x is not bound in ϕ, hence x is only bounded by one quantifier in the above formulas. However, this special case is sufficient for most of our purposes. Also note that we can apply this theorem repeatedly to change several bound variables in a formula.

Proof. By E$\forall$ we have $\forall x\phi(x) \vdash \phi(y)$. Hence by I$\forall$ we get $\forall x\phi(x) \vdash \forall y\phi(y)$, since y is not a free variable in $\forall x\phi(x)$, because we assumed that y is not a free variable in ϕ. On the other hand, by E$\forall$ we have $\forall y\phi(y) \vdash \phi(y)[x/y]$. Here $\phi(y)[x/y]$ is the result of substituting x back in ϕ, after we had substituted it with y. But by Axiom 1.7, and our assumptions about x, y, we can conclude that $\phi(y)[x/y]$ is the same

formula as $\phi(x)$. Also, note that x is neither a free variable nor a bound variable in $\phi(y)$ and $\forall y\phi(y)$, due to the Axioms 1.4, 1.6, and the fact that x is not bound in ϕ. Therefore by I$\forall$ we have $\forall y\phi(y) \vdash \forall x\phi(x)$. Hence we get $\forall x\phi(x) \equiv \forall y\phi(y)$, as desired.

Now consider the second equivalence. By I$\exists$ we have $\phi(y) \vdash \exists x\phi(x)$. Hence by E$\exists$ we get $\exists y\phi(y) \vdash \exists x\phi(x)$, since y is not a free variable in $\exists x\phi(x)$, because we assumed that it is not a free variable in ϕ. On the other hand, by I$\exists$ we have $\phi(y)[x/y] \vdash \exists y\phi(y)$. Again, we know that $\phi(y)[x/y]$ is the same formula as $\phi(x)$. Therefore by E$\exists$ we have $\exists x\phi(x) \vdash \exists y\phi(y)$, since similarly to the last paragraph, we can see that x is not a free variable in $\exists y\phi(y)$. Hence we get $\exists x\phi(x) \equiv \exists y\phi(y)$, as desired. ■

Finally, let us introduce a helpful shorthand notation for quantification.

Notation. Let ϕ be a formula that has z as a free variable. Then

$$\begin{array}{lll} \forall z \in x\ \phi(z) & \text{is a shorthand notation for} & \forall z(z \in x \to \phi(z)), \\ \exists z \in x\ \phi(z) & \text{is a shorthand notation for} & \exists z(z \in x \wedge \phi(z)). \end{array}$$

Remark. Let ϕ be a formula that has z as a free variable. If x does not have any element, i.e. if $\vdash \forall z(z \notin x)$, then we have

$$\vdash \forall z \in x\ \phi(z).$$

In other words, if x has no element, then $\phi(z)$ is true for every z in x. In this case we say that $\forall z \in x\ \phi(z)$ is **vacuously true**. To show this note that by E$\forall$ we have $\vdash z \notin x$. Hence by I$\wedge$ and the cut rule we get $z \in x \vdash (z \in x) \wedge (z \notin x)$. Therefore by EFQ rule (Theorem 1.1) we obtain

$$z \in x \vdash (z \in x) \wedge (z \notin x) \vdash \phi(z).$$

Thus by I$\to$, and then by I$\forall$ we get $\vdash \forall z(z \in x \to \phi(z))$, as desired.

The following technical result will be needed in later chapters.

Theorem 1.12. *Suppose ϕ, ψ are formulas, and x, z are variables. Suppose z is not a free variable in ψ, but it is a free variable in ϕ. Then we have*

(i)
$$\big(\forall z \in x\ \phi(z)\big) \vee \psi \equiv \forall z \in x\ \big(\phi(z) \vee \psi\big),$$

(ii)
$$\big(\exists z \in x\ \phi(z)\big) \wedge \psi \equiv \exists z \in x\ \big(\phi(z) \wedge \psi\big).$$

If in addition we assume that x is nonempty, i.e. if $\vdash \exists z(z \in x)$, then we have

(iii)
$$\big(\forall z \in x\ \phi(z)\big) \wedge \psi \equiv \forall z \in x\ \big(\phi(z) \wedge \psi\big),$$

(iv)

$$\big(\exists z \in x\ \phi(z)\big) \vee \psi \equiv \exists z \in x\ \big(\phi(z) \vee \psi\big).$$

Proof. We will use Theorems 1.4, 1.8 repeatedly in the following parts without explicit citation.

(i) We have

$$\begin{aligned}
\big(\forall z \in x\ \phi(z)\big) \vee \psi &\equiv \forall z(z \in x \to \phi(z)) \vee \psi && \text{(definition)}\\
&\equiv \forall z\big((z \in x \to \phi(z)) \vee \psi\big) && \text{(Theorem 1.10)}\\
&\equiv \forall z\big((\neg(z \in x) \vee \phi(z)) \vee \psi\big) && \text{(law of material implication)}\\
&\equiv \forall z\big(\neg(z \in x) \vee (\phi(z) \vee \psi)\big) && \text{(associativity of } \vee)\\
&\equiv \forall z\big(z \in x \to (\phi(z) \vee \psi)\big) && \text{(law of material implication)}\\
&\equiv \forall z \in x\ \big(\phi(z) \vee \psi\big). && \text{(definition)}
\end{aligned}$$

(ii) We have

$$\begin{aligned}
\big(\exists z \in x\ \phi(z)\big) \wedge \psi &\equiv \exists z(z \in x \wedge \phi(z)) \wedge \psi && \text{(definition)}\\
&\equiv \exists z\big((z \in x \wedge \phi(z)) \wedge \psi\big) && \text{(Theorem 1.10)}\\
&\equiv \exists z\big(z \in x \wedge (\phi(z) \wedge \psi)\big) && \text{(associativity of } \wedge)\\
&\equiv \exists z \in x\ \big(\phi(z) \wedge \psi\big). && \text{(definition)}
\end{aligned}$$

(iii) We have

$$\begin{aligned}
\big(\forall z \in x\ \phi(z)\big) \wedge \psi &\equiv \forall z(z \in x \to \phi(z)) \wedge \psi && \text{(definition)}\\
&\equiv \forall z\big((z \in x \to \phi(z)) \wedge \psi\big) && \text{(Theorem 1.10)}\\
&\equiv \forall z\big((\neg(z \in x) \vee \phi(z)) \wedge \psi\big) && \text{(law of material implication)}\\
&\equiv \forall z\big((z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi)\big). && \text{(distributivity of } \wedge \text{ over } \vee)
\end{aligned}$$

Now by E$\wedge$ and I$\vee$ we have

$$z \notin x \wedge \psi \vdash z \notin x \vdash z \notin x \vee (\phi(z) \wedge \psi).$$

By I$\vee$ we also have $\phi(z) \wedge \psi \vdash z \notin x \vee (\phi(z) \wedge \psi)$. Thus by E$\forall$ and E$\vee$ we get

$$\begin{aligned}
\forall z\big((z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi)\big) &\vdash (z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi)\\
&\vdash z \notin x \vee (\phi(z) \wedge \psi).
\end{aligned}$$

Therefore by I$\forall$ we obtain

$$\forall z\big((z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi)\big) \vdash \forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big).$$

On the other hand, by I$\wedge$ and EFQ rule (Theorem 1.1) we have

$$z \notin x, z \in x \vdash z \notin x \wedge z \in x \vdash \psi.$$

Also, by E$\wedge$ and the weakening rule we have $\phi(z) \wedge \psi, z \in x \vdash \psi$. Hence by E$\vee$ we get $z \notin x \vee (\phi(z) \wedge \psi), z \in x \vdash \psi$. Now by E$\forall$ and the cut rule we obtain

$$\forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big), z \in x \vdash \psi.$$

Thus by E$\exists$ we get $\forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big), \exists z(z \in x) \vdash \psi$, since z is not a free variable in ψ. However, we assumed that $\vdash \exists z(z \in x)$. Therefore by the cut rule we obtain $\forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big) \vdash \psi$. Now by I$\wedge$, the cut rule, and I$\vee$ we have

$$\forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big), z \notin x \vdash z \notin x \wedge \psi \vdash (z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi).$$

Also, by I$\vee$ and the weakening rule we have

$$\forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big), \phi(z) \wedge \psi \vdash (z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi).$$

Hence by E$\vee$ we get

$$\forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big), z \notin x \vee (\phi(z) \wedge \psi) \vdash (z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi).$$

Thus by E$\forall$ and the cut and contraction rules we obtain

$$\forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big) \vdash (z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi).$$

Finally, by I$\forall$ we get

$$\forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big) \vdash \forall z\big((z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi)\big).$$

So we have shown that $\forall z\big((z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi)\big) \equiv \forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big)$. Using this equivalence and the equivalences at the beginning of the proof of part (ii), we obtain

$$\begin{aligned}
\big(\forall z \in x\ \phi(z)\big) \wedge \psi &\equiv \forall z\big((z \notin x \wedge \psi) \vee (\phi(z) \wedge \psi)\big) \\
&\equiv \forall z\big(z \notin x \vee (\phi(z) \wedge \psi)\big) \\
&\equiv \forall z\big(z \in x \to (\phi(z) \wedge \psi)\big) \qquad \text{(law of material implication)} \\
&\equiv \forall z \in x\ \big(\phi(z) \wedge \psi\big), \qquad \text{(definition)}
\end{aligned}$$

as desired.

(iv) By definition we have $\big(\exists z \in x\ \phi(z)\big) \vee \psi \equiv \exists z(z \in x \wedge \phi(z)) \vee \psi$. Now by I$\vee$ we have $z \in x \vdash z \in x \vee \psi$. Hence by E$\wedge$, the cut rule, and I$\wedge$ we have

$$z \in x \wedge (\phi(z) \vee \psi) \vdash (z \in x \vee \psi) \wedge (\phi(z) \vee \psi).$$

Thus by I$\exists$ we get $z \in x \wedge (\phi(z) \vee \psi) \vdash \exists z\big((z \in x \vee \psi) \wedge (\phi(z) \vee \psi)\big)$. Therefore

$$\begin{aligned}\exists z\big(z \in x \wedge (\phi(z) \vee \psi)\big) &\vdash \exists z\big((z \in x \vee \psi) \wedge (\phi(z) \vee \psi)\big) && \text{(by E}\exists\text{)}\\ &\vdash \exists z\big((z \in x \wedge \phi(z)) \vee \psi\big) && \text{(distributivity of } \vee \text{ over } \wedge\text{)}\\ &\vdash \exists z(z \in x \wedge \phi(z)) \vee \psi. && \text{(Theorem 1.10)}\end{aligned}$$

On the other hand, by I$\vee$ we have $\psi \vdash \phi(z) \vee \psi$. Hence by I$\wedge$ and the cut rule we get

$$z \in x, \psi \vdash z \in x \wedge (\phi(z) \vee \psi).$$

Thus by I$\exists$ we obtain $z \in x, \psi \vdash \exists z\big(z \in x \wedge (\phi(z) \vee \psi)\big)$. Therefore by E$\exists$ we have

$$\exists z(z \in x), \psi \vdash \exists z\big(z \in x \wedge (\phi(z) \vee \psi)\big),$$

since z is not a free variable in ψ. However, we assumed that $\vdash \exists z(z \in x)$. Therefore by the cut rule we get

$$\psi \vdash \exists z\big(z \in x \wedge (\phi(z) \vee \psi)\big). \qquad (*)$$

In addition, by I$\vee$ we have $\phi(z) \vdash \phi(z) \vee \psi$. Hence by E$\wedge$, the cut rule, and I$\wedge$ we have

$$z \in x \wedge \phi(z) \vdash z \in x \wedge (\phi(z) \vee \psi).$$

Thus by I$\exists$ we get $z \in x \wedge \phi(z) \vdash \exists z\big(z \in x \wedge (\phi(z) \vee \psi)\big)$. Therefore by E$\exists$ we obtain

$$\exists z\big(z \in x \wedge \phi(z)\big) \vdash \exists z\big(z \in x \wedge (\phi(z) \vee \psi)\big).$$

So by applying E$\vee$ to the above entailment and $(*)$ we get

$$\exists z\big(z \in x \wedge \phi(z)\big) \vee \psi \vdash \exists z\big(z \in x \wedge (\phi(z) \vee \psi)\big).$$

Hence we have shown that $\exists z\big(z \in x \wedge \phi(z)\big) \vee \psi \equiv \exists z\big(z \in x \wedge (\phi(z) \vee \psi)\big)$, which in shorthand notation can be written as

$$\big(\exists z \in x\ \phi(z)\big) \vee \psi \equiv \exists z \in x\ \big(\phi(z) \vee \psi\big),$$

as desired. ■

1.6 Rules of Inference for Equality

Our last set of inference rules are the rules related to the equality relation.

Axiom 1.10 (Rules of inference for equality)**.** *For every variables like x, y, z we have*

(i) *Reflexivity* :

$$\vdash x = x.$$

(ii) *Symmetry* :

$$x = y \vdash y = x.$$

(iii) *Transitivity* :

$$x = y,\ y = z \vdash x = z.$$

(iv)

$$x = y,\ y \in z \vdash x \in z.$$

(v)

$$x = y,\ z \in y \vdash z \in x.$$

Remark. The above axiom is an axiom schema, so x, y, z can be any variables. Note that in particular, either two of the x, y, z, or all three of them, can be the same variable. Hence for example we also have

$$x = y,\ y \in y \vdash x \in x.$$

Because by part (iv) we have $x = y,\ y \in y \vdash x \in y$. Then by part (v) we have $x = y,\ x \in y \vdash x \in x$. Therefore, by the cut rule we have $x = y,\ y \in y,\ x = y \vdash x \in x$. Finally we get the desired by the exchange and contraction rules.

The first three parts of the above axiom state the elementary properties of equality. The reflexivity of $=$ means that any set is equal to itself. The transitivity of $=$ means that if two sets are equal to a third set, then they are equal. The symmetry of $=$ means that equality is a reciprocal relation, namely, if x equals y then y equals x too. Note that since we can change the variables in the above axiom by any other variables, we also have

$$y = x \vdash x = y.$$

Hence $x = y$ is actually equivalent to $y = x$, i.e. $x = y \equiv y = x$.

Remark. We can also state the above axiom by using quantifiers. For example by I$\forall$ we have $\vdash \forall x(x = x)$. Note that the premises do not contain x as a free variable, so the application of I$\forall$ is justified. Similarly, by E$\wedge$, the cut rule, I$\to$ and I$\forall$ we get

$$\vdash \forall x \forall y \forall z(x = y \wedge y = z \to x = z),$$
$$\vdash \forall x \forall y \forall z(x = y \wedge y \in z \to x \in z),$$
$$\vdash \forall x \forall y \forall z(x = y \wedge z \in y \to z \in x).$$

Also, by I$\to$ and I$\forall$ we have $\vdash \forall x \forall y(x = y \to y = x)$. In fact, since $x = y$ and $y = x$ are equivalent, we can show that $\vdash \forall x \forall y(x = y \leftrightarrow y = x)$.

The last two parts of the above axiom say that if $x = y$ then we can replace y by x in the atomic formulas $y \in z$ and $z \in y$. Note that we can also interpret the transitivity of equality as above, namely if $x = y$ then we can replace y by x in the atomic formula $y = z$. Furthermore, by using the symmetry of equality, we can also show that

$$x = y,\ z = y \vdash z = x.$$

Because by symmetry of equality we have $z = y \vdash y = z$. Now if we apply the cut rule to this entailment and the entailment given in the transitivity of equality, we obtain $z = y,\ x = y \vdash z = x$. And we get the desired by the exchange rule.

Hence if $x = y$ then we can replace y by x in the atomic formula $z = y$. In addition, similarly to the above remark, we can show that if $x = y$ then we can replace y by x in the atomic formula $y = y$. Therefore if $x = y$ then we can replace y by x in any atomic formula that has y as a free variable. If we combine this fact with Axiom 1.1, and use the rules of inference, we get the following: Suppose ϕ is a formula, and y is a free variable in ϕ. Let x be a variable which is not a bound variable in ϕ. Then we have

$$x = y,\ \phi(y) \vdash \phi(x).$$

In other words, if $x = y$ then we can replace the free occurrences of y by x in any formula that has y as a free variable.

Intuitively, this fact is obvious. Because if $x = y$ then x, y are denoting the same object. In other words, x, y are names for the same object. Thus, every formula which states a property about the object whose name is y, can be rephrased to state that property about the object whose name is x. And these different phrasings are equivalent, since they state the same property about the same object. However, we do not have the tools to prove this fact here; and if we need it, we have to accept it as an axiom. But since in practice we can show that it holds for any particular formula ϕ, and we will not use its general form, we do not state this fact as an axiom. The following example illustrates how we can check this fact for a particular formula.

Remark. There is a more general version of the rule $x = y,\ \phi(y) \vdash \phi(x)$, namely

$$x = y,\ \phi[y/z] \vdash \phi[x/z],$$

provided that x, y are not bound variables in ϕ. This more general version can be used when we need to only substitute some occurrences of y with x, not all of them.

Remark. Note that since $x = y$ and $y = x$ are equivalent, we can replace $x = y$ by $y = x$ in the premises of any entailment, due to Theorem 1.2.

***Example* 1.10.** Let ϕ be the formula $\forall z(z = y \to \neg(z \in y))$. Then we have $x = y, z = y \vdash z = x$. On the other hand, we have $y = x, z \in x \vdash z \in y$. Hence by the weakening rule and E$\neg$ we have

$$x = y, \neg(z \in y), z \in x \vdash \bot.$$

Note that in the premises, we replaced $y = x$ by $x = y$, by using Theorem 1.2. Therefore by I$\neg$ we get

$$x = y, \neg(z \in y) \vdash \neg(z \in x). \qquad (*)$$

Now note that $x = y, z = x \vdash z = y$. So by the cut rule and E$\to$ we have

$$x = y, z = y \to \neg(z \in y), z = x \vdash \neg(z \in y). \qquad (**)$$

Thus if we apply the cut rule to $(*), (**)$ we get

$$x = y, z = y \to \neg(z \in y), z = x \vdash \neg(z \in x).$$

Then by I$\to$ we obtain $x = y, z = y \to \neg(z \in y) \vdash z = x \to \neg(z \in x)$. Now by E$\forall$ and the cut rule we get

$$x = y, \forall z(z = y \to \neg(z \in y)) \vdash z = x \to \neg(z \in x).$$

Finally, since z is not free in the premises of the above entailment, we can apply I$\forall$ to obtain

$$x = y, \forall z(z = y \to \neg(z \in y)) \vdash \forall z(z = x \to \neg(z \in x)).$$

Hence we have shown that $x = y$, $\phi(y) \vdash \phi(x)$, as desired.

Remark. Note that a crucial reason that we can carry out the proof in the above example is the symmetry of equality, i.e. the fact that we can replace $x = y$ by $y = x$ whenever we need.

1.7 The Notion of Proof

We have informally used the notion of proof so far. We considered it as a convincing argument toward establishing a theorem, in which we used the axioms, and some elementary inference rules of meta-logic. For example in the proof of Theorem 1.11, when we showed that $\forall x\phi(x) \vdash \forall y\phi(y)$ due to I$\forall$, we had to first check that $\forall x\phi(x) \vdash \phi(y)$, and y is not a bound variable in $\forall x\phi(x)$. Then we could use I$\forall$ and conclude the desired result. But I$\forall$ is a conditional statement in meta-language. So we concluded the consequent of the conditional statement I$\forall$ by checking that its antecedent holds. Thus we have actually used the modus ponens rule in meta-logic.

This is also the case when we have used the other rules which are expressed as conditional statements in meta-language, like I$\to$ or E$\vee$.

Also note that the antecedent of the rule I$\forall$ consists of two parts. Thus in the above example we had to both check that $\forall x\phi(x) \vdash \phi(y)$, and that y is not a bound variable in $\forall x\phi(x)$. Of course we did this by separately checking each statement. Therefore when we concluded that the antecedent of I$\forall$ holds, we were actually using the rule of introduction of "and" in meta-logic. We can continue this exploration and collect all types of meta-logical arguments we presented in this chapter, but that will take us far from our main subject. So we do not pursue this any further. However, let us recall that as we discussed in the introduction to this chapter, we can accept all the results of this chapter as axioms; and we can consider those reasonings in meta-logic as mere convincing rationale for our choice of axioms.

The proofs mentioned in the above paragraphs are proofs in meta-logic. Let us now turn to the notion of proof inside logic. Conceptually, the proofs inside logic and meta-logic are not different. But we can make the notion of proof precise, when we work in logic.

Primitive Notion 1.6. A **proof** is a primitive notion, which intuitively, is a finite sequence of entailments. We represent a proof by writing the representations of its entailments successively, and we separate the entailments by " ;". We assume that we are able to recognize proofs from their representations, and we can distinguish between them through their representations. We also assume that we can recognize the entailments in a proof from the representation of the proof.

Remark. A proof is also called a **derivation** or a **formal proof**.

Axiom 1.11. *Suppose $\Gamma, \Delta, \Lambda_0, \Lambda_1, \Lambda_2$ denote collections of several formulas, which can be empty too. Let ϕ, ψ, τ be formulas. Then we have*

(i) *Let $\Gamma \vdash \psi$ be an entailment. Then $\Gamma \vdash \psi$ is also a proof.*
(ii) *If $\mathscr{D}$ is a proof, and $\Gamma \vdash \psi$ is an entailment, then $\mathscr{D}; \Gamma \vdash \psi$ is also a proof.*
(iii) *If $\mathscr{D}$ is a proof that contains the entailment $\Gamma \vdash \psi$, then $\mathscr{D}; \Gamma, \Delta \vdash \psi$ is also a proof.*
(iv) *If $\mathscr{D}$ is a proof that contains the entailment $\Gamma, \Delta, \Delta \vdash \psi$, then $\mathscr{D}; \Gamma, \Delta \vdash \psi$ is also a proof.*
(v) *If $\mathscr{D}$ is a proof that contains the entailment $\Lambda_0, \Gamma, \Lambda_1, \Delta, \Lambda_2 \vdash \psi$, then $\mathscr{D}; \Lambda_0, \Delta, \Lambda_1, \Gamma, \Lambda_2 \vdash \psi$ is also a proof.*
(vi) *If $\mathscr{D}$ is a proof that contains the entailments $\Gamma \vdash \phi$ and $\Delta, \phi \vdash \psi$, then $\mathscr{D}; \Gamma, \Delta \vdash \psi$ is also a proof.*
(vii) *If $\mathscr{D}$ is a proof that contains the entailment $\Gamma, \phi \vdash \psi$, then $\mathscr{D}; \Gamma \vdash \phi \to \psi$ is also a proof.*

(viii) *If $\mathscr{D}$ is a proof that contains the entailments $\Gamma, \phi \vdash \tau$ and $\Gamma, \psi \vdash \tau$, then $\mathscr{D}; \Gamma, \phi \vee \psi \vdash \tau$ is also a proof.*

(ix) *If $\mathscr{D}$ is a proof that contains the entailments $\Gamma, \phi \vdash \psi$ and $\Gamma, \psi \vdash \phi$, then $\mathscr{D}; \Gamma \vdash \phi \leftrightarrow \psi$ is also a proof.*

(x) *If $\mathscr{D}$ is a proof that contains the entailment $\Gamma, \phi \vdash \bot$, then $\mathscr{D}; \Gamma \vdash \neg\phi$ is also a proof.*

(xi) *If $\mathscr{D}$ is a proof that contains the entailment $\Gamma, \neg\phi \vdash \bot$, then $\mathscr{D}; \Gamma \vdash \phi$ is also a proof.*

(xii) *If $\mathscr{D}$ is a proof that contains the entailment $\Gamma \vdash \phi$, and x is not a free variable in any of the formulas in Γ, then $\mathscr{D}; \Gamma \vdash \forall x\phi$ is also a proof.*

(xiii) *If $\mathscr{D}$ is a proof that contains the entailment $\Gamma, \phi \vdash \psi$, and x is not a free variable in ψ, nor in any of the formulas in Γ, then $\mathscr{D}; \Gamma, \exists x\phi \vdash \psi$ is also a proof.*

Remark. We assume that every proof is constructed after several applications of the above rules, and there is no other way to construct a proof.

As it is evident from the above axiom, the proofs are constructed by using the rules of inference. Each part of the axiom, except the first two, corresponds to a rule of inference. Note that the entailment added in each part to the proof $\mathscr{D}$ is a valid entailment, due to the corresponding rule of inference. The first part of the axiom allows us to start a proof. The second part allows us to add a separately proven entailment to a proof. These two parts also incorporate the rules of inference which are not explicitly stated in the axiom, like I$\exists$ or E$\rightarrow$; because those rules only assert that certain entailments hold. More importantly, the first two parts allow us to use the axioms of equality, or the axioms of set theory (which will be added later) inside a proof.

Remark. When $\Gamma \vdash \psi$ is the last entailment that appears in a proof $\mathscr{D}$, we say that $\mathscr{D}$ is a proof of the entailment $\Gamma \vdash \psi$.

The virtue of a proof of an entailment, is that it encapsulates all the data needed to ensure that the entailment holds. However, it is not easy to understand a formal proof if there are no explanations around it expressed in the meta-language. So, mathematicians usually provide the proofs in the informal style expressed in the meta-language, similar to the proofs that we presented in previous sections. We will also adhere to this convention.

Remark. There is another approach to proofs which considers them as finite sequences of formulas. In this approach, we extend a proof by adding formulas that can be deduced from the previous formulas in the proof. So instead of recording the entailments in an argument, we record the formulas used in those entailments. Thus the two approaches to proofs are essentially the same. However, considering proofs as sequences of entailments seems more natural, and is easier to deal with axiomatically. Hence we chose this approach here.

***Example* 1.11.** Let ϕ, ψ be two formulas. The following is a formal proof of the entailment $\phi \to \psi, \neg\psi \vdash \neg\phi$, which is an instance of modus tollens.

$$\phi \to \psi, \phi \vdash \psi;\ \psi, \neg\psi \vdash \bot;\ \neg\psi, \psi \vdash \bot;\ \phi \to \psi, \phi, \neg\psi \vdash \bot;$$
$$\phi \to \psi, \neg\psi, \phi \vdash \bot;\ \phi \to \psi, \neg\psi \vdash \neg\phi.$$

As you can see, it is rather hard to comprehend a formal proof. To overcome this, we will usually write each entailment of a formal proof in a separate line, and we will mark the inference rule which implies that entailment. We will also number each line. With these conventions, the above formal proof can be rewritten as follows:

1	$\phi \to \psi, \phi \vdash \psi;$	(by E$\to$)
2	$\psi, \neg\psi \vdash \bot;$	(by E$\neg$)
3	$\neg\psi, \psi \vdash \bot;$	(by exchange rule)
4	$\phi \to \psi, \phi, \neg\psi \vdash \bot;$	(by cut rule applied to lines 1,3)
5	$\phi \to \psi, \neg\psi, \phi \vdash \bot;$	(by exchange rule)
6	$\phi \to \psi, \neg\psi \vdash \neg\phi.$	(by I$\neg$)

Remark. Note that the above proof is only a proof of the entailment $\phi \to \psi, \neg\psi \vdash \neg\phi$ for the particular instance of the formulas ϕ, ψ which we started with. And although the same proof works if we replace ϕ, ψ with any other particular pair of formulas, the above proof does not imply that the entailment $\phi \to \psi, \neg\psi \vdash \neg\phi$ holds for every pair of formulas ϕ, ψ. Because the logic that we have constructed only applies to the language of set theory, and that language can only talk about sets. So, statements about arbitrary pairs of formulas in the language of set theory lie outside of the language of set theory itself! Thus the logic that we have constructed cannot deal with such statements, and if we want to prove them we have to enter the realm of meta-logic.

As a side note, let us mention that we can repeat the above proof with some particular formulas ϕ, ψ, whenever we need to use the rule modus tollens. Therefore we do not need to accept modus tollens as a general rule of inference, if we prefer to avoid the theorems proved by meta-logical reasonings. The same remark applies to the other theorems proved in this chapter.

***Example* 1.12.** Let ϕ, ψ be two formulas. The following is a formal proof of the entailment $\vdash \exists x(\phi \vee \psi) \leftrightarrow \exists x\phi \vee \exists x\psi$. Informally, we can also say that the following

is a formal proof of the equivalence $\exists x(\phi \vee \psi) \equiv \exists x\phi \vee \exists x\psi$.

1	$\phi \vdash \exists x\phi$;	(by I$\exists$)
2	$\exists x\phi \vdash \exists x\phi \vee \exists x\psi$;	(by I$\vee$)
3	$\phi \vdash \exists x\phi \vee \exists x\psi$;	(by cut rule applied to lines 1,2)
4	$\psi \vdash \exists x\psi$;	(by I$\exists$)
5	$\exists x\psi \vdash \exists x\phi \vee \exists x\psi$;	(by I$\vee$)
6	$\psi \vdash \exists x\phi \vee \exists x\psi$;	(by cut rule applied to lines 4,5)
7	$\phi \vee \psi \vdash \exists x\phi \vee \exists x\psi$;	(by E$\vee$ applied to lines 3,6)
8	$\exists x(\phi \vee \psi) \vdash \exists x\phi \vee \exists x\psi$;	(by E$\exists$)
9	$\phi \vdash \phi \vee \psi$;	(by I$\vee$)
10	$\phi \vee \psi \vdash \exists x(\phi \vee \psi)$;	(by I$\exists$)
11	$\phi \vdash \exists x(\phi \vee \psi)$;	(by cut rule applied to lines 9,10)
12	$\exists x\phi \vdash \exists x(\phi \vee \psi)$;	(by E$\exists$)
13	$\psi \vdash \phi \vee \psi$;	(by I$\vee$)
14	$\psi \vdash \exists x(\phi \vee \psi)$;	(by cut rule applied to lines 13,10)
15	$\exists x\psi \vdash \exists x(\phi \vee \psi)$;	(by E$\exists$)
16	$\exists x\phi \vee \exists x\psi \vdash \exists x(\phi \vee \psi)$;	(by E$\vee$ applied to lines 12,15)
17	$\vdash \exists x(\phi \vee \psi) \leftrightarrow \exists x\phi \vee \exists x\psi$.	(by I$\leftrightarrow$ applied to lines 8,16)

Chapter 2

Sets

2.1 Axioms of Extensionality and Separation

Primitive Notion 2.1. The notion of **Set** is a primitive notion, which intuitively denotes a collection of some objects. We denote sets by variables in the language of set theory. We assume that everything that we talk about in this language is a set, and there are no other types of objects in the universe of sets, i.e. in our domain of discourse. The notion of **equality of sets** is also a primitive notion. For two sets like x, y, we denote their equality by the formula $x = y$.

The first step in axiomatic development of set theory is to ensure that at least one set exists. However, we do not need to state this fact as an axiom, because it will follow from our rules of inference. By I$\exists$ we have $x = x \vdash \exists x(x = x)$. On the other hand, by Axiom 1.10 we have $\vdash x = x$. Thus by the cut rule we get

$$\vdash \exists x(x = x).$$

In other words, there is at least one set which is equal to itself. In particular, there is at least one set.

Primitive Notion 2.2. The **set membership** relation between sets is a primitive notion. It is denoted by the symbol $\in$. So for two sets like x, y, the formula $x \in y$ means that x is an **element** or a **member** of y.

Remark. If $x \in y$, we also say that "x is in y", "x belongs to y", or "y contains x".

Notation. The expression $x, y \in z$ is a shorthand notation for $(x \in z) \wedge (y \in z)$. Similar notations can also be used when there are more than two elements of a set.

Remark. Since we assumed that everything is a set, the elements of a set are also sets, the elements of the elements of a set are also sets, and so forth. We can include some objects in our theory which are not sets, but can be a member of other

sets. These objects are called *atoms*. However, adding atoms to our set theory will not make the theory richer, and will not produce new mathematics; so we do not incorporate them.

Axiom of Extensionality.

$$\vdash \forall x \forall y \big(\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y \big).$$

Informally, the axiom of extensionality says that if two sets have the same elements, they must be equal. In other words, the elements of a set determine the set uniquely.

Notation. Since a set is uniquely determined by its elements, we can use the elements to represent the set. This is the common practice in mathematics, especially when the set has finitely many elements. For example, if the elements of the set x are a, c, z, u, we will write

$$x = \{a, c, z, u\}.$$

And if we want to define a set by specifying its elements, or in other words, if we want to give a name to set specified by its elements, we will use $:=$ instead of $=$. For example we can write

$$y := \{s, r, u\},$$

in order to give the name y to the set $\{s, r, u\}$.

A consequence of the axiom of extensionality is that the elements of a set do not have any order, and there cannot be any duplication of elements in a set. We do not have the tools to even express these facts rigorously. But informally we can argue that if two sets have the same elements except that one of them contains x, and the other one contains x, x, then the two sets are equal; because they both contain exactly the same elements other than x, and they also both contain x. In addition, if for example a set consists of x, y, and another set consists of y, x, then the two sets are equal; because they both contain x, and they both contain y, and they do not contain any other element.

Remark. By repeated applications of Theorem 1.11, we can change the bound variables in our axioms. This will be useful when we employ the axioms inside the proof of some theorem. For example, we can state the axiom of extensionality as follows:

$$\vdash \forall a \forall b \big(\forall w (w \in a \leftrightarrow w \in b) \rightarrow a = b \big).$$

The axiom of extensionality expresses a relation between $\in$ and $=$. There is also another relation between them given by Axiom 1.10.

The axiom of extensionality can also be used to show that two given sets are not equal. We simply have to find an element of one of the sets that does not belong to the other set.

Theorem 2.1 (converse of extensionality)**.** *If two sets are equal, then they will have the same elements, i.e.*

$$\vdash \forall x \forall y \big(x = y \to \forall z (z \in x \leftrightarrow z \in y) \big).$$

Proof. We will prove this result by presenting its formal proof; mainly in order to illustrate the method. We have

1	$x = y, z \in y \vdash z \in x;$	(by axiom 1.10)
2	$x = y \vdash y = x;$	(by axiom 1.10)
3	$y = x, z \in x \vdash z \in y;$	(by axiom 1.10)
4	$z \in x, y = x \vdash z \in y;$	(by exchange rule)
5	$x = y, z \in x \vdash z \in y;$	(by cut rule applied to lines 2,4)
6	$x = y \vdash z \in x \leftrightarrow z \in y;$	(by I$\leftrightarrow$ applied to lines 5,1)
7	$x = y \vdash \forall z (z \in x \leftrightarrow z \in y);$	(by I$\forall$)
8	$\vdash x = y \to \forall z (z \in x \leftrightarrow z \in y);$	(by I$\to$)
9	$\vdash \forall y \big(x = y \to \forall z (z \in x \leftrightarrow z \in y) \big);$	(by I$\forall$)
10	$\vdash \forall x \forall y \big(x = y \to \forall z (z \in x \leftrightarrow z \in y) \big).$	(by I$\forall$)

■

Remark. If we combine the axiom of extensionality and its converse, then for any two sets like x, y we get

$$x = y \equiv \forall z (z \in x \leftrightarrow z \in y).$$

Because by E$\forall$ we get $\vdash \forall z (z \in x \leftrightarrow z \in y) \to x = y$ and $\vdash x = y \to \forall z (z \in x \leftrightarrow z \in y)$. Hence by E$\to$ we obtain $\forall z (z \in x \leftrightarrow z \in y) \vdash x = y$ and $x = y \vdash \forall z (z \in x \leftrightarrow z \in y)$. Thus the two formulas are equivalent.

The next step is to state axioms which allow us to build sets. Historically, the idea was that if we can describe some sets by a formula, then we can gather all those sets and create a set that contains them all. In other words, the following statement was accepted as an axiom (which we will see shortly that it cannot be true):

Axiom schema of Comprehension (False). *Let ϕ be a formula that has z as a free variable. Suppose that x is not a free variable of ϕ. Then we have*

$$\vdash \exists x \forall z \big(z \in x \leftrightarrow \phi(z) \big).$$

However, this axiom cannot be true, because it results in contradictions, as is shown by **Russell's paradox**. Namely, if we take $\phi(z)$ to be the formula $z \notin z$, then

$$\vdash \exists x \forall z (z \in x \leftrightarrow z \notin z).$$

Intuitively, if the set x belongs to itself, then by definition we must have $x \notin x$. And if x does not belong to itself, then again by definition we must have $x \in x$. Let us demonstrate this observation more carefully. By E$\forall$ and I$\exists$ we have

$$\forall z (z \in x \leftrightarrow z \notin z) \vdash x \in x \leftrightarrow x \notin x \vdash \exists x (x \in x \leftrightarrow x \notin x).$$

Hence by E$\exists$ we get

$$\exists x \forall z (z \in x \leftrightarrow z \notin z) \vdash \exists x (x \in x \leftrightarrow x \notin x).$$

Thus by the cut rule we have $\vdash \exists x (x \in x \leftrightarrow x \notin x)$. In other words, there is a set x which belongs to itself if and only if it does not belong to itself! Informally, it is obvious that this statement is contradictory. But to derive a formal contradiction we argue as follows.

To simplify the notation let us denote $x \in x$ by ϕ. Then $x \notin x$ is $\neg\phi$. Now by E$\leftrightarrow$ we have $\phi \leftrightarrow \neg\phi, \phi \vdash \neg\phi$. We also know that $\phi \leftrightarrow \neg\phi, \phi \vdash \phi$, since $\phi \vdash \phi$. Hence by the cut rule and E$\neg$ we get $\phi \leftrightarrow \neg\phi, \phi \vdash \bot$. Thus by I$\neg$ we have $\phi \leftrightarrow \neg\phi \vdash \neg\phi$. Similarly we can show that $\phi \leftrightarrow \neg\phi, \neg\phi \vdash \bot$. Therefore by RAA we get $\phi \leftrightarrow \neg\phi \vdash \phi$. Thus by the cut rule and E$\neg$ we obtain $\phi \leftrightarrow \neg\phi \vdash \bot$. Hence we have shown that

$$x \in x \leftrightarrow x \notin x \vdash \bot.$$

Thus by E$\exists$ we have $\exists x (x \in x \leftrightarrow x \notin x) \vdash \bot$. Therefore by the cut rule we get $\vdash \bot$, i.e. we formally obtained a contradiction by assuming the axiom of comprehension.

So, the naive idea of constructing sets using formulas results in contradiction. Hence we have to modify this naive idea. Historically, extensive studies of set theory showed that the problem with the axiom of comprehension is that it allows us to construct very large sets. It even allows us to construct the set of all sets by considering the formula $z = z$ (we will see later that there is no such set in the current standard theory of sets). The problem with such large sets is that sometimes a self-referencing can happen through them. For example, in our analysis of Russell's paradox, we could substitute z by x. Thus the defining formula of the elements of x could potentially be satisfied by x itself; and as we saw this was problematic.

Hence, in order to avoid the problems created by the axiom of comprehension, we have to somehow restrict the size of the sets constructed by that axiom. The solution was found to be replacing the axiom of comprehension by the following weaker axiom.

Axiom schema of Separation. *Let ϕ be a formula that has z as a free variable. Suppose that x is not a free variable of ϕ. Then we have*

$$\vdash \forall y \exists x \forall z \big(z \in x \leftrightarrow (z \in y) \wedge \phi(z)\big).$$

Informally, this axiom says that we can separate those elements of y which satisfy the property ϕ, and collect them in a set x. The difference with the axiom of comprehension is that this time we only consider those sets z which belong to y. So, intuitively, the size of x cannot be larger than the set y; and therefore we can avoid the problems created by the axiom of comprehension.

Remark. The axiom schema of separation is also called *subset axiom schema*, or *axiom schema of restricted comprehension.*

As a consequence of the axiom of extensionality, the set x in the axiom of separation is unique. Formally this means that

$$\vdash \forall z\big(z \in x_1 \leftrightarrow (z \in y) \wedge \phi(z)\big) \wedge \forall z\big(z \in x_2 \leftrightarrow (z \in y) \wedge \phi(z)\big) \rightarrow x_1 = x_2.$$

Here x_1, x_2 are variables which are not free variables in ϕ. In fact, a more general version of this uniqueness result is true.

Theorem 2.2. *Let ψ be a formula that has z as a free variable. Suppose that x_1, x_2 are not free variables in ψ. Then we have*

$$\vdash \forall z\big(z \in x_1 \leftrightarrow \psi(z)\big) \wedge \forall z\big(z \in x_2 \leftrightarrow \psi(z)\big) \rightarrow x_1 = x_2.$$

Remark. Informally, this theorem says that there is at most one set corresponding to the formula ψ in the axiom of comprehension. However, note that we cannot say anything about the existence of x_1, x_2. Also note that if we take $\psi(z)$ to be the formula $(z \in y) \wedge \phi(z)$, then we obtain the uniqueness result for the axiom of separation. Finally note that if we show the above entailment, then by I$\forall$ we can also show that

$$\vdash \forall x_1 \forall x_2 \big[\forall z\big(z \in x_1 \leftrightarrow \psi(z)\big) \wedge \forall z\big(z \in x_2 \leftrightarrow \psi(z)\big) \rightarrow x_1 = x_2\big].$$

Proof. By Theorem 1.10, E$\forall$, and E$\wedge$ we have

$$\begin{aligned}
&\forall z\big(z \in x_1 \leftrightarrow \psi(z)\big) \wedge \forall z\big(z \in x_2 \leftrightarrow \psi(z)\big) \\
&\qquad \vdash \forall z\big[\big(z \in x_1 \leftrightarrow \psi(z)\big) \wedge \big(z \in x_2 \leftrightarrow \psi(z)\big)\big] \\
&\qquad \vdash \big(z \in x_1 \leftrightarrow \psi(z)\big) \wedge \big(z \in x_2 \leftrightarrow \psi(z)\big) \vdash z \in x_1 \leftrightarrow \psi(z).
\end{aligned}$$

To simplify the notation let us denote the first premise of the above entailments by τ. Then by the cut rule and E$\leftrightarrow$ we have $\tau, z \in x_1 \vdash \psi(z)$. On the other hand, we can show as above that $\tau \vdash z \in x_2 \leftrightarrow \psi(z)$. Hence by the cut rule and

E↔ we get $\tau, \psi(z) \vdash z \in x_2$. Thus by the cut and contraction rules we obtain $\tau, z \in x_1 \vdash z \in x_2$. Similarly we can show that $\tau, z \in x_2 \vdash z \in x_1$. Therefore by I↔ we get $\tau \vdash z \in x_1 \leftrightarrow z \in x_2$. Thus by I∀ and the axiom of extensionality we have

$$\tau \vdash \forall z(z \in x_1 \leftrightarrow z \in x_2) \vdash x_1 = x_2.$$

Hence we get the desired by I→. ■

Remark. An important idea which we used in the above argument is that in order to show that an object with some properties is unique, we assume that two objects have those properties, and then we show that the two objects must be equal.

Notation. Let ϕ be a formula. If we want to say that there is exactly one set x which satisfies ϕ we can write

$$\exists x \phi(x) \wedge \forall x_1 \forall x_2 \big(\phi(x_1) \wedge \phi(x_2) \rightarrow x_1 = x_2\big).$$

The above formula is sometimes denoted by $\exists! x \phi(x)$. The symbol $\exists!$ is called **uniqueness quantifier** or **unique existential quantifier**. There are also other ways to express the existence of a unique object. For example we can write it as

$$\exists x \big(\phi(x) \wedge \forall y(\phi(y) \rightarrow y = x)\big).$$

***Example* 2.1.** As an example note that we have shown the uniqueness of the set x in the axiom of separation. The existence of x is also given by the axiom itself. So we have

$$\vdash \exists! x \forall z \big(z \in x \leftrightarrow (z \in y) \wedge \phi(z)\big).$$

Now we can use I∀ to conclude that

$$\vdash \forall y \exists! x \forall z \big(z \in x \leftrightarrow (z \in y) \wedge \phi(z)\big).$$

Notation. The set x whose existence is guaranteed by the axiom of separation will usually be denoted by

$$x = \{z \in y : \phi(z)\}.$$

Note that this set is uniquely determined by the above theorem.

Remark. The formula ϕ can have free variables other than z too. In addition, note that y can also be a free variable in ϕ. Suppose for example that $z, y, w_1, \ldots, w_n$ are the distinct free variables of ϕ. Then by repeated applications of I∀ we can conclude that

$$\vdash \forall w_1 \ldots \forall w_n \forall y \exists x \forall z \big(z \in x \leftrightarrow (z \in y) \wedge \phi(z, y, w_1, \ldots, w_n)\big).$$

This is how the axiom of separation is expressed in some texts.

Definition 2.1. We say a set x is a **subset** of a set y if every element of x is also an element of y. Formally this means that

$$\forall z(z \in x \to z \in y).$$

In this case we write $x \subset y$. This relation between x, y is called **inclusion**.

Remark. The relation $x \subset y$ is also referred to by "y is a **superset** of x", "y includes x", or "x is included in y". In addition, we may write $y \supset x$ to mean that $x \subset y$.

Remark. Some authors use the notation $x \subseteq y$ for subsets. This notation emphasizes that the subset x can be equal to the set y. However, we will not use this notation throughout these notes.

Remark. Note that $x \subset y$ is actually a shorthand notation for the formula $\forall z(z \in x \to z \in y)$. So whenever $x \subset y$ appears in an expression, we have to replace it by the above formula. The next example illustrates this.

***Example* 2.2.** Two sets x, y are equal if and only if $x \subset y$ and $y \subset x$. In other words

$$\vdash x = y \leftrightarrow (x \subset y) \wedge (y \subset x).$$

Equivalently we can write

$$x = y \equiv (x \subset y) \wedge (y \subset x).$$

First note that the above equivalence actually means

$$x = y \equiv \forall z(z \in x \to z \in y) \wedge \forall z(z \in y \to z \in x).$$

Now by Theorems 1.6, 1.8 and 1.10 we have

$$\begin{aligned}&\forall z(z \in x \to z \in y) \wedge \forall z(z \in y \to z \in x)\\ &\qquad \equiv \forall z\big((z \in x \to z \in y) \wedge (z \in y \to z \in x)\big) \equiv \forall z(z \in x \leftrightarrow z \in y).\end{aligned}$$

But the last formula is equivalent to $x = y$, as shown in the remark after Theorem 2.1. Thus we get the desired by transitivity of $\equiv$.

Remark. Note that in the axiom schema of separation we have $x \subset y$, because every z in x is also in y.

Theorem 2.3. *There exists a set that has no element, i.e.*

$$\vdash \exists x \forall z(z \notin x).$$

Proof. By the axiom of separation we know that

$$\vdash \forall y \exists x \forall z \big(z \in x \leftrightarrow (z \in y) \wedge (z \neq z)\big).$$

Thus by E$\forall$ we have $\vdash \exists x \forall z \big(z \in x \leftrightarrow (z \in y) \wedge (z \neq z)\big)$. On the other hand, by E$\forall$, Theorem 1.6, and E$\wedge$ we have

$$\begin{aligned} \forall z \big(z \in x \leftrightarrow (z \in y) \wedge (z \neq z)\big) &\vdash z \in x \leftrightarrow (z \in y) \wedge (z \neq z) \\ &\vdash z \in x \rightarrow (z \in y) \wedge (z \neq z). \end{aligned}$$

Now by E$\rightarrow$, the cut rule, and E$\wedge$ we get

$$\forall z \big(z \in x \leftrightarrow (z \in y) \wedge (z \neq z)\big), z \in x \vdash (z \in y) \wedge (z \neq z) \vdash z \neq z.$$

However, we know that $\vdash z = z$. So by the weakening rule and E$\neg$ we have

$$\forall z \big(z \in x \leftrightarrow (z \in y) \wedge (z \neq z)\big), z \in x \vdash \bot.$$

Hence by I$\neg$ we get $\forall z \big(z \in x \leftrightarrow (z \in y) \wedge (z \neq z)\big) \vdash z \notin x$. Thus by I$\forall$ and I$\exists$ we obtain

$$\forall z \big(z \in x \leftrightarrow (z \in y) \wedge (z \neq z)\big) \vdash \forall z (z \notin x) \vdash \exists x \forall z (z \notin x).$$

Finally by E$\exists$ we get

$$\vdash \exists x \forall z \big(z \in x \leftrightarrow (z \in y) \wedge (z \neq z)\big) \vdash \exists x \forall z (z \notin x).$$

Hence we must have $\vdash \exists x \forall z (z \notin x)$ due to the cut rule. ■

The set with no element is called the **empty set**. It is denoted by $\emptyset$ or $\{\,\}$. Notice that when we assign a name or a notation to an object, we have to ensure that the object is uniquely determined. The next theorem confirms this for the empty set.

Theorem 2.4. *The empty set is unique, i.e.*

$$\vdash \forall z (z \notin x) \wedge \forall z (z \notin y) \rightarrow x = y.$$

Proof. By Theorem 1.10, E$\forall$, and E$\wedge$ we have

$$\forall z (z \notin x) \wedge \forall z (z \notin y) \vdash \forall z \big((z \notin x) \wedge (z \notin y)\big) \vdash (z \notin x) \wedge (z \notin y) \vdash z \notin x.$$

Hence by I$\wedge$ we get

$$\forall z (z \notin x) \wedge \forall z (z \notin y), z \in x \vdash (z \in x) \wedge (z \notin x).$$

But the contradiction $(z \in x) \wedge (z \notin x)$ can imply anything, so in particular we have

$$\forall z(z \notin x) \wedge \forall z(z \notin y), z \in x \vdash (z \in x) \wedge (z \notin x) \vdash z \in y.$$

Similarly we can show that $\forall z(z \notin x) \wedge \forall z(z \notin y), z \in y \vdash z \in x$. Therefore by I$\leftrightarrow$ we obtain

$$\forall z(z \notin x) \wedge \forall z(z \notin y) \vdash z \in x \leftrightarrow z \in y.$$

Thus by I$\forall$ and the axiom of extensionality we get

$$\forall z(z \notin x) \wedge \forall z(z \notin y) \vdash \forall z(z \in x \leftrightarrow z \in y) \vdash x = y.$$

Hence by I$\rightarrow$ we get the desired. ■

A question that arises is that what exactly $\emptyset$ is. We know that it is a notation for a uniquely determined set in the universe of sets. However, when we use the rules of inference, we have to work with formulas. So we have to give meanings to the expressions involving $\emptyset$ which we will use. Let x be a variable. Then the following expressions will be interpreted as follows:

$x = \emptyset$, $\emptyset = x$	are shorthand notations for	$\forall z(z \notin x)$,
$\emptyset \in x$	is a shorthand notation for	$\exists y\big((y \in x) \wedge \forall z(z \notin y)\big)$.

The expressions $x \in \emptyset$ and $\emptyset \in \emptyset$ can be replaced by any false formula, like a contradiction. And, the expression $\emptyset = \emptyset$ can be replaced by any true formula, like the negation of a contradiction.

Remark. There is another approach to address the above question. In this approach we add the symbol $\emptyset$ to the alphabet of the language. We treat $\emptyset$ similar to the variables, except that we do not quantify over it. Symbols of this type are called **constants**. So, we can extend the language of set theory by adding the constant $\emptyset$ to it. More explicitly, for every variable like x, we add the following expressions to the atomic formulas of the language:

$$x = \emptyset, \quad \emptyset = x, \quad \emptyset = \emptyset, \quad \emptyset \in x, \quad x \in \emptyset, \quad \emptyset \in \emptyset.$$

Then we can use these formulas and other atomic formulas to build more complex formulas. We also have to modify the rules of inference, and add extra rules, in order to be able to deal with the above formulas. For example, we have to accept the following entailments:

$$\vdash \emptyset = \emptyset, \qquad \vdash \forall z(z \notin \emptyset), \qquad \vdash \emptyset \notin \emptyset.$$

We will not provide the details of this approach, since we are not going to follow it. Instead, we consider $\emptyset$ as a shorthand notation, and we interpret the formulas containing $\emptyset$ as we explained before.

Remark. We say a set is **nonempty** if it has at least one element. If x is nonempty we write $x \neq \emptyset$. Note that if we interpret $x \neq \emptyset$ as $\neg(x = \emptyset)$, then it means

$$\neg\big(\forall z(z \notin x)\big) \equiv \exists z(z \in x).$$

So $x \neq \emptyset$ says that x has at least one element; and this is in agreement with our definition.

***Example* 2.3.** Let x be a set. Then we have $\emptyset \subset x$. Informally, the reason is that $\emptyset$ has no element, so all its elements are also elements of x. Formally, we have to show that

$$\vdash \forall z(z \notin y) \to \forall z(z \in y \to z \in x).$$

By I$\wedge$ and the EFQ rule (Theorem 1.1) we have

$$z \notin y, z \in y \vdash z \in y \wedge z \notin y \vdash z \in x.$$

Hence by I$\to$ we get $z \notin y \vdash z \in y \to z \in x$. Thus by E$\forall$ we obtain

$$\forall z(z \notin y) \vdash z \notin y \vdash z \in y \to z \in x.$$

Therefore by I$\forall$ we have $\forall z(z \notin y) \vdash \forall z(z \in y \to z \in x)$. Now we get the desired result by I$\to$.

Theorem 2.5. *There is no set of all sets, i.e.*

$$\vdash \forall y \exists z(z \notin y).$$

Proof. The proof is by contradiction, i.e. we assume the negation of the above formula, aka $\exists y \forall z(z \in y)$, and we derive a contradiction. Then we get the desired result. To derive a contradiction from $\exists y \forall z(z \in y)$, we essentially argue as we did in Russell's paradox. Informally, we consider the subset $x := \{z \in y : z \notin z\}$; and then we arrive at a contradiction by noting that $x \in x$ if and only if $x \notin x$. Notice that this idea works because y contains every set; hence $x \in y$, and therefore it makes sense to ask whether x belongs to x or not. Let us formalize this idea.

By the axiom of separation we have

$$\vdash \forall y \exists x \forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \notin z)\big).$$

Hence by E$\forall$ we get $\vdash \exists x \forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \notin z)\big)$. Now by E$\forall$ and I$\exists$ we have

$$\forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \notin z)\big) \vdash x \in x \leftrightarrow (x \in y) \wedge (x \notin x).$$

Also, by E$\leftrightarrow$ and E$\wedge$ we have

$$\forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \notin z)\big), \forall z(z \in y), x \in x \vdash (x \in y) \wedge (x \notin x) \vdash x \notin x.$$

Note that we have added $\forall z(z \in y)$ to the premises by using the weakening rule. Also note that by E$\forall$ we have $\forall z(z \in y) \vdash x \in y$. In addition, by I$\wedge$ we get $\forall z(z \in y), x \notin x \vdash (x \in y) \wedge (x \notin x)$. Thus by E$\leftrightarrow$ and the cut rule we obtain

$$\forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \notin z)\big), \forall z(z \in y), x \notin x \vdash x \in x.$$

Therefore by I$\leftrightarrow$ and I$\exists$ we get

$$\begin{aligned}\forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \notin z)\big), \forall z(z \in y) &\vdash x \in x \leftrightarrow x \notin x \\ &\vdash \exists x(x \in x \leftrightarrow x \notin x).\end{aligned}$$

Hence by E$\exists$ we obtain

$$\exists x \forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \notin z)\big), \forall z(z \in y) \vdash \exists x(x \in x \leftrightarrow x \notin x).$$

But we know that $\vdash \exists x \forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \notin z)\big)$; so by the cut rule we get

$$\forall z(z \in y) \vdash \exists x(x \in x \leftrightarrow x \notin x).$$

Therefore by E$\exists$ we get

$$\exists y \forall z(z \in y) \vdash \exists x(x \in x \leftrightarrow x \notin x).$$

However, as we have shown in our discussion of Russell's paradox, the formula $\exists x(x \in x \leftrightarrow x \notin x)$ implies contradictions. Thus by I$\neg$, and Theorem 1.10 we get

$$\vdash \neg \exists y \forall z(z \in y) \equiv \forall y \exists z(z \notin y),$$

as desired. ■

2.2 Axioms of Pairing, Union, and Power Set

The axiom schema of separation allows us to construct subsets of a given set. So, informally, if we have a set, then we can construct sets which are smaller than that given set. However, unlike the axiom of comprehension, the axiom of separation does not allow us to enlarge a given set. Hence we need additional axioms to be able to do that. The next three axioms provide us methods for construction of larger sets from some given sets.

Axiom of Pairing.

$$\vdash \forall x \forall y \exists w\big(\forall z(z \in w \leftrightarrow z = x \vee z = y)\big).$$

This axiom says that for any two sets like x, y, there is a set w whose elements are exactly x, y. We denote the set w by

$$\{x, y\}.$$

Of course, we have to check that the set $\{x, y\}$ is uniquely determined.

Theorem 2.6. *For every two sets x, y, the set $\{x, y\}$ is uniquely determined, i.e.*

$$\vdash \forall z(z \in w_1 \leftrightarrow z = x \vee z = y) \wedge \forall z(z \in w_2 \leftrightarrow z = x \vee z = y) \rightarrow w_1 = w_2.$$

Proof. This follows easily from Theorem 2.2 by taking the formula $\psi(z)$ to be $z = x \vee z = y$. ■

If we take y to be the same as x, then the set $\{x, x\}$ only contains x. We denote this set by $\{x\}$. The set $\{x\}$ has exactly one element, and that element is x. A set which has exactly one element is called a **singleton**. Let us rigorously prove the existence and uniqueness of the singleton $\{x\}$ in the following theorem.

Theorem 2.7. *For every set x, there is a unique set whose only element is x, i.e.*

$$\vdash \forall x \exists! w\big(\forall z(z \in w \leftrightarrow z = x)\big).$$

Proof. First note that the above entailment actually means that

$$\vdash \forall x\big[\exists w\big(\forall z(z \in w \leftrightarrow z = x)\big) \wedge \forall w_1 \forall w_2\big(\forall z(z \in w_1 \leftrightarrow z = x) \wedge \forall z(z \in w_2 \leftrightarrow z = x) \rightarrow w_1 = w_2\big)\big].$$

The uniqueness part follows easily from Theorem 2.2 by taking the formula $\psi(z)$ to be $z = x$. For the existence note that by applying E$\forall$ to the axiom of pairing we get

$$\vdash \forall y \exists w\big(\forall z(z \in w \leftrightarrow z = x \vee z = y)\big).$$

If we apply E$\forall$ one more time, and substitute y by x, we obtain

$$\vdash \exists w\big(\forall z(z \in w \leftrightarrow z = x \vee z = x)\big).$$

On the other hand, we know that $z = x \vee z = x \equiv z = x$. Therefore by repeated applications of Theorems 1.4 and 1.8 we get

$$\exists w\big(\forall z(z \in w \leftrightarrow z = x \vee z = x)\big) \equiv \exists w\big(\forall z(z \in w \leftrightarrow z = x)\big).$$

Hence by the cut rule we get $\vdash \exists w\big(\forall z(z \in w \leftrightarrow z = x)\big)$. Now if we combine this entailment with the entailment for uniqueness using I$\wedge$, and then apply I$\forall$, we obtain the desired result. ■

Next, let us state more clearly what are the meanings of the notations $\{x, y\}$ and $\{x\}$ when they appear in a formula. Let z be a variable. Then we have

$$\begin{array}{lll}
z \in \{x, y\} & \text{is a shorthand notation for} & z = x \vee z = y, \\
z \in \{x\} & \text{is a shorthand notation for} & z = x, \\
z = \{x, y\} & \text{is a shorthand notation for} & \forall v(v \in z \leftrightarrow v = x \vee v = y), \\
z = \{x\} & \text{is a shorthand notation for} & \forall v(v \in z \leftrightarrow v = x), \\
\{x, y\} \in z & \text{is a shorthand notation for} & \\
 & & \exists w\big(w \in z \wedge \forall v(v \in w \leftrightarrow v = x \vee v = y)\big), \\
\{x\} \in z & \text{is a shorthand notation for} & \exists w\big(w \in z \wedge \forall v(v \in w \leftrightarrow v = x)\big).
\end{array}$$

Note that equality is still a symmetric relation when it is used in shorthand notations. So $\{x, y\} = z$ and $\{x\} = z$ have the same meaning as $z = \{x, y\}$ and $z = \{x\}$ respectively.

Axiom of Union.

$$\vdash \forall x \exists w \forall z\big(z \in w \leftrightarrow \exists y(y \in x \wedge z \in y)\big).$$

This axiom says that for a set like x, there is a set w whose elements are exactly the elements of the elements of x. So, a set belongs to w if and only if it belongs to at least one element of x. In other words, if we consider x as a family of sets, i.e. as the family of the sets which are elements of x, then w is the **union** of the members of this family of sets. We denote w by

$$\bigcup x, \qquad \text{or} \qquad \bigcup_{y \in x} y.$$

Theorem 2.8. *For every set x, the set $\bigcup x$ is uniquely determined, i.e.*

$$\vdash \forall z\big(z \in w_1 \leftrightarrow \exists y(y \in x \wedge z \in y)\big) \wedge \forall z\big(z \in w_2 \leftrightarrow \exists y(y \in x \wedge z \in y)\big) \rightarrow w_1 = w_2.$$

Proof. This follows easily from Theorem 2.2 by taking the formula $\psi(z)$ to be $\exists y(y \in x \wedge z \in y)$. ■

Since in the axiom of union we did not assume anything about the set x, we are allowed to form the union of any family of sets. In particular, we can form the union of infinitely many sets. However, the case of finitely many sets deserves special attention. For two sets like x, y, we define their **union** to be

$$x \cup y := \bigcup\{x, y\}.$$

The next theorem states the main property of $x \cup y$.

Theorem 2.9. *For every two sets x, y, there is a unique set whose elements are exactly those sets which belong to at least one of x or y; i.e.*

$$\vdash \forall x \forall y \exists! w \big(\forall z (z \in w \leftrightarrow z \in x \vee z \in y)\big).$$

Proof. The uniqueness follows from Theorem 2.2 by taking the formula $\psi(z)$ to be $z \in x \vee z \in y$. For the existence we implement the idea of $x \cup y := \bigcup\{x, y\}$ more rigorously. By applying E$\forall$ to the axiom of pairing we get $\vdash \exists w \big(\forall z (z \in w \leftrightarrow z = x \vee z = y)\big)$. By using Theorem 1.11 we can change the bound variable w to v. Then we have

$$\vdash \exists v \big(\forall z (z \in v \leftrightarrow z = x \vee z = y)\big).$$

Now by applying E$\forall$ to the axiom of union, and substituting the variable x by v, we obtain $\vdash \exists w \forall z \big(z \in w \leftrightarrow \exists y (y \in v \wedge z \in y)\big)$. By Theorems 1.11, 1.4 and 1.8 we can change the bound variable y in this formula to u. Hence we get

$$\vdash \exists w \forall z \big(z \in w \leftrightarrow \exists u (u \in v \wedge z \in u)\big).$$

On the other hand, by E$\forall$ we have

$$\begin{aligned} \tau: &\quad \forall z (z \in v \leftrightarrow z = x \vee z = y) \vdash u \in v \leftrightarrow u = x \vee u = y, \\ \sigma: &\quad \forall z \big(z \in w \leftrightarrow \exists u (u \in v \wedge z \in u)\big) \vdash z \in w \leftrightarrow \exists u (u \in v \wedge z \in u). \end{aligned}$$

To simplify the notation, we have denoted the premises of the above entailments by τ, σ respectively. Note that we have $\vdash \exists v \tau$ and $\vdash \exists w \sigma$.

Now by the Axiom 1.10 we have $z \in u, u = x \vdash z \in x$, and $z \in u, u = y \vdash z \in y$. Hence by constructive dilemma (Theorem 1.1) we get

$$z \in u, u = x \vee u = y \vdash z \in x \vee z \in y.$$

But by E$\leftrightarrow$ we have $\tau, u \in v \vdash u = x \vee u = y$. Thus by the cut rule we get

$$\tau, u \in v, z \in u \vdash z \in x \vee z \in y.$$

Hence by E$\wedge$ and the cut rule we get

$$\tau, u \in v \wedge z \in u \vdash z \in x \vee z \in y.$$

Thus by E$\exists$ we obtain $\tau, \exists u (u \in v \wedge z \in u) \vdash z \in x \vee z \in y$. Note that u is not a free variable in τ, so we were allowed to use E$\exists$. In addition, by E$\leftrightarrow$ and the cut rule we have $\sigma, z \in w \vdash \exists u (u \in v \wedge z \in u)$. Therefore by the cut rule we get

$$\tau, \sigma, z \in w \vdash z \in x \vee z \in y. \qquad (*)$$

On the other hand, by E$\forall$ we have

$$\forall z(z \in v \leftrightarrow z = x \vee z = y) \vdash x \in v \leftrightarrow x = x \vee x = y.$$

Note that the premise of the above entailment is τ. Also, by reflexivity of equality and I$\vee$ we have $\vdash x = x \vdash x = x \vee x = y$. Hence by the cut rule and E$\leftrightarrow$ we get $\tau \vdash x \in v$. Thus by I$\wedge$ and I$\exists$ we have $\tau, z \in x \vdash x \in v \wedge z \in x \vdash \exists u(u \in v \wedge z \in u)$. Similarly we can show that $\tau, z \in y \vdash \exists u(u \in v \wedge z \in u)$. Hence by E$\vee$ we get

$$\tau, z \in x \vee z \in y \vdash \exists u(u \in v \wedge z \in u).$$

Now by E$\leftrightarrow$ and the cut rule we obtain

$$\tau, \sigma, z \in x \vee z \in y \vdash z \in w. \tag{$**$}$$

Therefore by applying I$\leftrightarrow$ to $(*), (**)$ we obtain $\tau, \sigma \vdash z \in w \leftrightarrow z \in x \vee z \in y$. Hence by I$\forall$ and I$\exists$ we get

$$\tau, \sigma \vdash \forall z(z \in w \leftrightarrow z \in x \vee z \in y) \vdash \exists w \forall z(z \in w \leftrightarrow z \in x \vee z \in y).$$

Note that z is not a free variable in τ, σ, so we were allowed to use I$\forall$. Then by E$\exists$ we get $\tau, \exists w \sigma \vdash \exists w \forall z(z \in w \leftrightarrow z \in x \vee z \in y)$. Note that w is not a free variable in τ, so we were allowed to use E$\exists$. Now we know that $\vdash \exists w \sigma$. Hence by the cut rule we get

$$\tau \vdash \exists w \forall z(z \in w \leftrightarrow z \in x \vee z \in y).$$

Finally by E$\exists$ we get $\vdash \exists v \tau \vdash \exists w \forall z(z \in w \leftrightarrow z \in x \vee z \in y)$. So we get the desired by the cut rule. ■

Let x be a nonempty set. Then there is a set w whose elements are exactly those sets which belong to every element of x. In other words, if we consider x as a family of sets, i.e. as the family of the sets which are elements of x, then w is the **intersection** of the members of this family of sets. We denote w by

$$\bigcap x, \qquad \text{or} \qquad \bigcap_{y \in x} y.$$

The next theorem shows that the intersection of a nonempty family of sets exists and is unique.

Theorem 2.10. *For every nonempty set x, there is a unique set whose elements are exactly those sets which belong to every element of x; i.e.*

$$\vdash \forall x \big[\exists u(u \in x) \rightarrow \exists! w \forall z \big(z \in w \leftrightarrow \forall y(y \in x \rightarrow z \in y)\big)\big].$$

Proof. The uniqueness follows from Theorem 2.2 by taking the formula $\psi(z)$ to be $\forall y(y \in x \to z \in y)$. The idea for proving the existence is to define the set w by the axiom of separation using the formula

$$(z \in u) \wedge \forall y(y \in x \to z \in y),$$

where u is some element of x (which we know exists, since x is nonempty). Now note that if $\forall y(y \in x \to z \in y)$ holds then $z \in u$ holds too, since $u \in x$. So we can drop $z \in u$ from the defining formula of w, and get the desired. To implement this idea rigorously we argue as follows.

First, as a practice in using the rules of inference, let us show how to rigorously change the variables in a formula. The axiom of separation implies that

$$\vdash \forall y \exists x \forall z \big(z \in x \leftrightarrow (z \in y) \wedge \forall y_0(y_0 \in x_0 \to z \in y_0)\big).$$

By applying E$\forall$ to the above formula, and substituting u for y, we get

$$\vdash \exists x \forall z \big(z \in x \leftrightarrow (z \in u) \wedge \forall y_0(y_0 \in x_0 \to z \in y_0)\big).$$

Then by Theorems 1.11, 1.4, and 1.8 we can change the bound variables to get

$$\vdash \exists w \forall z \big(z \in w \leftrightarrow (z \in u) \wedge \forall y(y \in x_0 \to z \in y)\big).$$

Finally by I$\forall$ we can add $\forall x_0$ to the formula; and then by E$\forall$ we can drop $\forall x_0$, and substitute x_0 by x to get

$$\vdash \exists w \forall z \big(z \in w \leftrightarrow (z \in u) \wedge \forall y(y \in x \to z \in y)\big). \qquad (\star)$$

And the variables in this formula are compatible with what we want to derive.

Let us denote $\forall z\big(z \in w \leftrightarrow (z \in u) \wedge \forall y(y \in x \to z \in y)\big)$ by τ. Note that we have shown that $\vdash \exists w \tau$. Now by E$\forall$ we have

$$\tau \vdash z \in w \leftrightarrow (z \in u) \wedge \forall y(y \in x \to z \in y). \qquad (*)$$

On the other hand, by E$\forall$ we have $\forall y(y \in x \to z \in y) \vdash u \in x \to z \in u$. Therefore by E$\to$ we get

$$\forall y(y \in x \to z \in y), u \in x \vdash z \in u.$$

So by I$\wedge$ we have $\forall y(y \in x \to z \in y), u \in x \vdash (z \in u) \wedge \forall y(y \in x \to z \in y)$. Thus by E$\leftrightarrow$ applied to $(*)$ we get

$$\tau, u \in x, \forall y(y \in x \to z \in y) \vdash z \in w,$$
$$\tau, u \in x, z \in w \vdash \forall y(y \in x \to z \in y).$$

Hence by I$\leftrightarrow$ we obtain $\tau, u \in x \vdash z \in w \leftrightarrow \forall y(y \in x \to z \in y)$. Then by I$\forall$ and I$\exists$ we get

$$\tau, u \in x \vdash \forall z\big(z \in w \leftrightarrow \forall y(y \in x \to z \in y)\big)$$
$$\vdash \exists w \forall z\big(z \in w \leftrightarrow \forall y(y \in x \to z \in y)\big).$$

Now by E$\exists$ we get $\exists w \tau, u \in x \vdash \exists w \forall z\big(z \in w \leftrightarrow \forall y(y \in x \to z \in y)\big)$. But we know that $\vdash \exists w \tau$. So we have $u \in x \vdash \exists w \forall z\big(z \in w \leftrightarrow \forall y(y \in x \to z \in y)\big)$. Therefore by E$\exists$ we obtain

$$\exists u(u \in x) \vdash \exists w \forall z\big(z \in w \leftrightarrow \forall y(y \in x \to z \in y)\big).$$

Hence we get the desired result by applying I$\to$, and then I$\forall$. ■

Remark. Note that if we allow x to be empty in the above theorem, then the formula $\forall y(y \in x \to z \in y)$ becomes true for every z, because $y \in x$ is false for every y. Therefore every set z must belong to the intersection of an empty family of sets! Thus to avoid such complications, we do not define the intersection of the empty set.

Remark. Also note that the formula $(\star)$ in the above proof is just a statement of the axiom of separation (with the initial $\forall$ dropped) using different variables than what we used earlier. However, we wanted to show how to start from the exact statement of an axiom, and change the variables into other variables. Although in the sequel, we usually state the axioms using the variables which we want to use, and rarely repeat these kinds of reasonings.

Note that similarly to the case of unions, we are allowed to form the intersection of any nonempty family of sets. In particular, we can form the intersection of infinitely many sets. But, the intersection of finitely many sets deserves special attention too. For two sets like x, y, we denote their **intersection** by

$$x \cap y.$$

The next theorem proves the main property of $x \cap y$, together with its existence and uniqueness.

Theorem 2.11. *For every two sets x, y, there is a unique set whose elements are exactly those sets which belong to both x and y; i.e.*

$$\vdash \forall x \forall y \exists! w\big(\forall z(z \in w \leftrightarrow z \in x \wedge z \in y)\big).$$

Proof. The uniqueness follows from Theorem 2.2 by taking the formula $\psi(z)$ to be $z \in x \wedge z \in y$. To prove the existence, note that by the axiom of separation we have $\vdash \forall y \exists x \forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \in x_0)\big)$. Then by I$\forall$ we get

$$\vdash \forall x_0 \forall y \exists x \forall z\big(z \in x \leftrightarrow (z \in y) \wedge (z \in x_0)\big).$$

Now by repeated applications of Theorem 1.11 we can change the bound variables to obtain

$$\vdash \forall y \forall x \exists w \forall z \big(z \in w \leftrightarrow (z \in x) \wedge (z \in y)\big).$$

Finally, we can switch $\forall y$ and $\forall x$ by Theorem 1.10 to get the desired result. ■

Remark. We say the two sets x, y are **disjoint** if their intersection is empty, i.e. if $x \cap y = \emptyset$. When x, y are disjoint, we sometimes denote their union by $x \sqcup y$.

Similarly, we say that the elements of the set x are **pairwise disjoint** or **mutually disjoint** if for every distinct $u, z \in x$ we have $u \cap z = \emptyset$; in other words, if we have

$$\forall u \forall z \big[\big((u \in x \wedge z \in x) \wedge u \neq z\big) \rightarrow u \cap z = \emptyset\big].$$

In this case, we sometimes denote $\bigcup x$ by $\bigsqcup x$.

Axiom of Power Set.

$$\vdash \forall x \exists w \big(\forall z (z \in w \leftrightarrow z \subset x)\big).$$

This axiom says that for a set like x, there is a set w whose elements are exactly the subsets of x. The set w is called the **power set** of x, and is denoted by

$$\mathcal{P}(x).$$

Note that we can replace $z \subset x$ by its interpretation; so a more formal statement of the axiom is

$$\vdash \forall x \exists w \forall z \big(z \in w \leftrightarrow \forall v (v \in z \rightarrow v \in x)\big).$$

Remark. The axiom of power set tacitly implies that all the subsets of x are also sets. Note that by the axiom schema of separation, every subset of x that can be described by a formula is a set; but, that axiom does not imply that every collection of some elements of x is itself a set. However, at least informally, the axiom of power set implies that any collection of some elements of x, i.e. any subset of x, is itself a set; and the collection of all subsets of x is a set too. But, we cannot state this fact formally in the language of set theory, because if we can talk about a collection of some elements of x in this language, then that collection is already assumed to be a set!

Theorem 2.12. *For every set x, the set $\mathcal{P}(x)$ is uniquely determined, i.e.*

$$\vdash \forall z \big(z \in w_1 \leftrightarrow z \subset x\big) \wedge \forall z \big(z \in w_2 \leftrightarrow z \subset x\big) \rightarrow w_1 = w_2.$$

Proof. This follows easily from Theorem 2.2 by taking the formula $\psi(z)$ to be $z \subset x$. ■

Remark. Since for every set like x we have $\emptyset \subset x$, we get $\emptyset \in \mathcal{P}(x)$. So in particular $\mathcal{P}(x) \neq \emptyset$.

Next, let us state more clearly what are the meanings of some of the notations, when they appear in a formula. We will only state the more useful cases; the reader should be able to figure out the rest of them. Let x, y, z be variables. Note that when we consider $\bigcap x$, we assume that x is nonempty. We have

$z \in \bigcup x$	is a shorthand notation for	$\exists y(y \in x \wedge z \in y)$,
$z \in x \cup y$	is a shorthand notation for	$z \in x \vee z \in y$,
$z \in \bigcap x$	is a shorthand notation for	$\forall y(y \in x \to z \in y)$,
$z \in x \cap y$	is a shorthand notation for	$z \in x \wedge z \in y$,
$z \in \mathcal{P}(x)$	is a shorthand notation for	$z \subset x$,
$x \cap y \neq \emptyset$	is a shorthand notation for	$\exists z(z \in x \wedge z \in y)$,
$x \cap y = \emptyset$	is a shorthand notation for	$\neg\exists z(z \in x \wedge z \in y)$.

So far we have introduced four axioms and one axiom schema. There are three more axioms and one more axiom schema in the standard theory of sets. The **axiom of infinity** states that an infinite set exists. We will study this axiom in depth in Chapter 4. Of course, we have to make the notion of infinite set precise, before we can even state this axiom. The next axiom is the **axiom of choice**. It states that if we have a set x whose elements are all nonempty sets, then we can choose an element from all those nonempty sets in x. In other words, there is a set whose intersection with any element of x is a singleton. This axiom has a special place among the axioms. We will study the axiom of choice and its consequences in Chapter 7.

The next two axioms are not used in the development of most of mathematics outside set theory. So we only briefly talk about them here.

Axiom of Regularity.

$$\vdash \forall x\big(x \neq \emptyset \to \exists z \in x\,(z \cap x = \emptyset)\big).$$

This axiom says that any nonempty set x has an element z whose intersection with x is empty. As a consequence, a set that belongs to itself cannot exist. Because if $x \in x$ then the set $\{x\}$ does not satisfy the axiom of regularity, since we have $\{x\} \cap x = x \neq \emptyset$. This axiom also implies that a cycle of membership cannot exist. For example, we cannot have $x \in y$ and $y \in x$; because otherwise the set $\{x, y\}$ would not satisfy the axiom of regularity. Although the axiom of regularity excludes some undesirable objects from the universe of sets, it is only used inside set theory, and is irrelevant to the other parts of mathematics.

Axiom schema of Replacement. *Let ϕ be a formula that has u, z as free variables. Suppose that x is not a free variable of ϕ. Then we have*

$$\vdash \forall y\big[\forall u \in y\ \exists! z \phi(u,z) \to \exists x \forall z\big(z \in x \leftrightarrow \exists u \in y\ \phi(u,z)\big)\big].$$

This axiom says that if we can map every element u of the set y to a uniquely determined z, which satisfies $\phi(u,z)$, then we can collect all those z's and form a set x. In other words, if we replace each u by the corresponding z, the resulting collection is still a set. Informally we can also say that if we have a function whose domain is the set y, then the image of the function is also a set, which we called x. The axiom of replacement is needed if we want to construct very large sets, which are rarely used outside set theory. It is also needed for the construction of some specific sets, called *ordinals*, which are of great importance in set theory.

The axiomatic theory of sets that we developed in this chapter is called **Zermelo–Fraenkel** set theory. It is abbreviated by **ZFC**, where C stands for the axiom of choice. Currently, ZFC is regarded as the standard theory of sets by most mathematicians. The theory which results when we exclude the axiom of choice from the axioms of ZFC is denoted by **ZF**. Note that both ZF and ZFC have infinitely many axioms, because the axiom schemata of separation and replacement are actually collections of infinitely many axioms, one for each formula ϕ.

2.3 Algebra of Sets

Notation. To make the formulas more readable, from now on we will also use upper case letters to denote sets.

Our goal in this section is to find various relations between the operations on sets, like union and intersection. We also study the properties of the inclusion relation. The next theorem, which is a generalization of Theorem 2.2, is a useful tool in this regard. A similar result about the inclusion will be stated afterwards.

Theorem 2.13. *Let ϕ, ψ be two formulas that have z as a free variable. Suppose that $\phi \equiv \psi$. Then we have*

$$\vdash \forall z\big(z \in x \leftrightarrow \phi(z)\big) \wedge \forall z\big(z \in y \leftrightarrow \psi(z)\big) \to x = y.$$

Remark. Informally, this theorem says that if two sets are defined by equivalent formulas, then they must be equal. However, note that we do not claim that such sets x, y actually exist.

Proof. By Theorem 1.10, E$\forall$, and E$\wedge$ we have

$$\begin{aligned}
&\forall z\big(z \in x \leftrightarrow \phi(z)\big) \wedge \forall z\big(z \in y \leftrightarrow \psi(z)\big) \\
&\qquad \vdash \forall z\big[\big(z \in x \leftrightarrow \phi(z)\big) \wedge \big(z \in y \leftrightarrow \psi(z)\big)\big] \\
&\qquad \vdash \big(z \in x \leftrightarrow \phi(z)\big) \wedge \big(z \in y \leftrightarrow \psi(z)\big) \vdash z \in x \leftrightarrow \phi(z).
\end{aligned}$$

To simplify the notation let us denote the first premise of the above entailments by τ. Then by the cut rule, E$\leftrightarrow$, and the equivalence of ϕ, ψ we get

$$\tau, z \in x \vdash \phi(z) \vdash \psi(z).$$

On the other hand, we can show as above that $\tau \vdash z \in y \leftrightarrow \psi(z)$. Hence by the cut rule and E$\leftrightarrow$ we get $\tau, \psi(z) \vdash z \in y$. Thus by the cut and contraction rules we obtain $\tau, z \in x \vdash z \in y$. Similarly we can show that $\tau, z \in y \vdash z \in x$. Therefore by I$\leftrightarrow$ we get $\tau \vdash z \in x \leftrightarrow z \in y$. Thus by I$\forall$ and the axiom of extensionality we have

$$\tau \vdash \forall z(z \in x \leftrightarrow z \in y) \vdash x = y.$$

Hence we get the desired by I$\rightarrow$. ■

Theorem 2.14. *Let ϕ, ψ be two formulas that have z as a free variable. Suppose*

$$\vdash \forall z\big(z \in x \leftrightarrow \phi(z)\big), \qquad \vdash \forall z\big(z \in y \leftrightarrow \psi(z)\big).$$

Then we have

$$x \subset y \qquad \textit{if and only if} \qquad \phi \vdash \psi.$$

Remark. Remember that $x \subset y$ is just a shorthand notation for the formula $\forall z(z \in x \rightarrow z \in y)$.

Remark. Informally, this theorem says that if two sets x, y are defined by formulas ϕ, ψ respectively, then $\phi \vdash \psi$ is equivalent to $x \subset y$. However, note that we do not claim that such sets x, y actually exist.

Proof. Suppose $\phi \vdash \psi$. By E$\forall$ we have $\vdash \forall z\big(z \in x \leftrightarrow \phi(z)\big) \vdash z \in x \leftrightarrow \phi(z)$. Also, by E$\leftrightarrow$ we get

$$z \in x, z \in x \leftrightarrow \phi(z) \vdash \phi(z) \vdash \psi(z).$$

Then by the cut rule we obtain $z \in x \vdash \psi(z)$. On the other hand, by E$\forall$ we have

$$\vdash \forall z\big(z \in y \leftrightarrow \psi(z)\big) \vdash z \in y \leftrightarrow \psi(z).$$

Hence by the cut rule and E$\leftrightarrow$ we get $\psi(z) \vdash z \in y$. Thus by the cut rule we obtain $z \in x \vdash z \in y$. Therefore by I$\rightarrow$ we get $\vdash z \in x \rightarrow z \in y$. Thus by I$\forall$ we have

$$\vdash \forall z(z \in x \rightarrow z \in y),$$

as desired. Conversely, suppose $x \subset y$, i.e. $\vdash \forall z(z \in x \rightarrow z \in y)$. Then by E$\forall$ we get

$$\vdash \forall z\big(z \in x \rightarrow z \in y\big) \vdash z \in x \rightarrow z \in y.$$

Hence by the cut rule and E$\rightarrow$ we have $z \in x \vdash z \in y$. Similarly to the above, we can also show that $z \in y \vdash \psi$, and $\phi \vdash z \in x$. Therefore by the cut rule we get $\phi \vdash \psi$, as desired. ■

The following technical result will be needed later in the section.

Theorem 2.15. *Let A, B be two sets. Let ϕ be a formula that has x, a as free variables, and does not contain b as a free variable. Suppose*

$$\vdash \forall a \in A\ \exists b \forall x(x \in b \leftrightarrow \phi(a, x)).$$

Also suppose that

$$b \in B \equiv \exists a \in A\ \forall x(x \in b \leftrightarrow \phi(a, x)).$$

Then we have

(i)
$$x \in \bigcap B \equiv \forall a \in A\ \phi(a, x),$$

(ii)
$$x \in \bigcup B \equiv \exists a \in A\ \phi(a, x).$$

Remark. Informally, the theorem says that if the elements of B are constructed from the elements of A by using the formula ϕ, then we can express $\bigcap B, \bigcup B$ in terms of A. Note that the first assumption of the theorem means that it is possible to construct a set from every element of A by using the formula ϕ. As an example, suppose $\phi(a, x)$ is $x \in a \vee x \in c$. In this case we will have $b = a \cup c$. Then the theorem implies that $x \in \bigcap_{b \in B} b$ if and only if for every $a \in A$ we have $x \in a \vee x \in c$, i.e. $x \in a \cup c$. We can express this by saying that $x \in \bigcap_{a \in A} a \cup c$. Hence we can write $\bigcap_{b \in B} b = \bigcap_{a \in A} a \cup c$. Similarly we have $\bigcup_{b \in B} b = \bigcup_{a \in A} a \cup c$.

Proof. To simplify the notation, let us denote the formula $\forall x(x \in b \leftrightarrow \phi(a, x))$ by $\psi(a, b)$. Then the first assumption of the theorem can be written as

$$\vdash \forall a(a \in A \to \exists b\, \psi(a, b)). \tag{$*$}$$

Also, we have

$$b \in B \equiv \exists a(a \in A \wedge \psi(a, b)). \tag{$**$}$$

(i) Note that $x \in \bigcap B$ means $\forall b \in B\ x \in b$, or more explicitly

$$\forall b(b \in B \to x \in b).$$

We want to show that this formula is equivalent to the formula

$$\forall a(a \in A \to \phi(a, x)).$$

Now by E$\wedge$ and E$\to$ we have

$$a \in A \to \phi(a, x),\ a \in A \wedge \psi(a, b) \vdash \phi(a, x).$$

We also have $a \in A \wedge \psi(a,b) \vdash \psi(a,b) \vdash x \in b \leftrightarrow \phi(a,x)$, due to E$\wedge$ and E$\forall$. Hence by the cut rule and E$\leftrightarrow$ we get

$$a \in A \to \phi(a,x),\ a \in A \wedge \psi(a,b) \vdash x \in b.$$

Thus by E$\forall$ we have

$$\forall a(a \in A \to \phi(a,x)),\ a \in A \wedge \psi(a,b) \vdash x \in b.$$

Now we can apply E$\exists$ to obtain

$$\forall a(a \in A \to \phi(a,x)),\ \exists a(a \in A \wedge \psi(a,b)) \vdash x \in b.$$

Hence by $(**)$ we get $\forall a(a \in A \to \phi(a,x)),\ b \in B \vdash x \in b$. So by I$\to$ we have

$$\forall a(a \in A \to \phi(a,x)) \vdash b \in B \to x \in b,$$

and by I$\forall$ we obtain $\forall a(a \in A \to \phi(a,x)) \vdash \forall b(b \in B \to x \in b)$, since b is not a free variable in ϕ.

On the other hand, by I$\wedge$, I$\exists$, and $(**)$ we obtain

$$a \in A, \psi(a,b) \vdash \exists a(a \in A \wedge \psi(a,b)) \vdash b \in B.$$

Hence by the cut rule, E$\forall$, and E$\to$ we get

$$a \in A, \psi(a,b), \forall b(b \in B \to x \in b) \vdash x \in b.$$

Also, by E$\forall$ we have $\psi(a,b) \vdash x \in b \leftrightarrow \phi(a,x)$. Hence by E$\leftrightarrow$ we get

$$a \in A, \psi(a,b), \forall b(b \in B \to x \in b) \vdash \phi(a,x).$$

So by E$\exists$ we obtain

$$a \in A, \exists b\, \psi(a,b), \forall b(b \in B \to x \in b) \vdash \phi(a,x),$$

since b is not a free variable in ϕ. However, by applying E$\forall$ and E$\to$ to $(*)$ we get $a \in A \vdash \exists b\, \psi(a,b)$. Therefore by the cut rule we get

$$a \in A, \forall b(b \in B \to x \in b) \vdash \phi(a,x).$$

Thus by I$\to$ we have $\forall b(b \in B \to x \in b) \vdash a \in A \to \phi(a,x)$. Hence by I$\forall$ we obtain

$$\forall b(b \in B \to x \in b) \vdash \forall a(a \in A \to \phi(a,x)),$$

as desired.

(ii) Note that $x \in \bigcup B$ means $\exists b \in B\ x \in b$, or more explicitly

$$\exists b(b \in B \wedge x \in b).$$

We want to show that this formula is equivalent to the formula

$$\exists a(a \in A \wedge \phi(a, x)).$$

Now by E$\wedge$ we have

$$a \in A \wedge \psi(a, x) \vdash a \in A.$$

We also have $a \in A \wedge \psi(a, b) \vdash \psi(a, b) \vdash x \in b \leftrightarrow \phi(a, x)$, due to E$\wedge$ and E$\forall$. Hence by the cut rule and E$\leftrightarrow$ we get

$$a \in A \wedge \psi(a, x), x \in b \vdash \phi(a, x).$$

Thus by the cut rule, I$\wedge$, and I$\exists$ we obtain

$$a \in A \wedge \psi(a, x), x \in b \vdash a \in A \wedge \phi(a, x) \vdash \exists a(a \in A \wedge \phi(a, x)).$$

Therefore by E$\exists$ we get

$$\exists a(a \in A \wedge \psi(a, x)), x \in b \vdash \exists a(a \in A \wedge \phi(a, x)).$$

Hence by $(**)$ and E$\wedge$ we have $b \in B \wedge x \in b \vdash \exists a(a \in A \wedge \phi(a, x))$. So by E$\exists$ we get

$$\exists b(b \in B \wedge x \in b) \vdash \exists a(a \in A \wedge \phi(a, x)),$$

since b is not a free variable in ϕ.

On the other hand, by I$\wedge$, I$\exists$, and $(**)$ we obtain

$$a \in A, \psi(a, b) \vdash \exists a(a \in A \wedge \psi(a, b)) \vdash b \in B.$$

We also have $\psi(a, b) \vdash x \in b \leftrightarrow \phi(a, x)$, due to E$\forall$. Hence by the cut rule, E$\wedge$, and E$\leftrightarrow$ we get

$$a \in A \wedge \phi(a, x), \psi(a, x) \vdash x \in b.$$

Thus by I$\wedge$ and I$\exists$ we obtain

$$a \in A \wedge \phi(a, x), \psi(a, x) \vdash b \in B \wedge x \in b \vdash \exists b(b \in B \wedge x \in b).$$

So by E$\exists$ we get

$$a \in A \wedge \phi(a, x), \exists b\, \psi(a, x) \vdash \exists b(b \in B \wedge x \in b),$$

since b is not a free variable in ϕ. However, by applying E$\forall$ and E$\rightarrow$ to $(*)$ we get $a \in A \vdash \exists b\, \psi(a, b)$. Therefore by the cut rule we get

$$a \in A \wedge \phi(a, x) \vdash \exists b(b \in B \wedge x \in b).$$

Thus by E$\exists$ we obtain $\exists a(a \in A \wedge \phi(a, x)) \vdash \exists b(b \in B \wedge x \in b)$, as desired. ■

Notation. From now on, we will also use

$$\iff$$

to denote the equivalence of formulas, or other expressions which are not technically formulas, but can be rewritten as formulas. We may also use

$$\implies$$

to denote the entailment relation between formulas or such expressions. Hence $\phi \iff \psi$ means $\phi \equiv \psi$, or equivalently $\vdash \phi \leftrightarrow \psi$; and $\phi \implies \psi$ means $\phi \vdash \psi$, or equivalently $\vdash \phi \to \psi$.

For two sets like A, B, the set $B - A$, which is called the **difference** of A, B, is the set of all elements of B that do not belong to A. In other words

$$B - A := \{x \in B : x \notin A\}.$$

The next theorem proves the main property of $B - A$, together with its existence and uniqueness.

Remark. Another notation for $B - A$ is $B \setminus A$.

Theorem 2.16. *For every two sets A, B, there is a unique set whose elements are exactly those sets which belong to B but not to A; i.e.*

$$\vdash \forall A \forall B \exists! C \big(\forall x (x \in C \leftrightarrow x \in B \wedge x \notin A)\big).$$

Proof. The uniqueness follows from Theorem 2.2 by taking the formula ψ to be $x \in B \wedge x \notin A$. The existence follows from the axiom of separation by taking the formula ϕ to be $x \notin A$. Note that we have to change the variables in that axiom into the variables in this theorem. ■

Note that for two sets A, B we have two differences, $B - A$ and $A - B$. The union of these two differences is called the **symmetric difference** of A, B, and is denoted by $A \Delta B$. In other words

$$A \Delta B := (A - B) \cup (B - A).$$

Theorem 2.17. *Suppose A, B, C are sets. Then we have*

(i) *Commutativity* :

$$A \cap B = B \cap A, \qquad \textit{and} \qquad A \cup B = B \cup A.$$

(ii) *Associativity* :

$$(A \cap B) \cap C = A \cap (B \cap C), \qquad \textit{and} \qquad (A \cup B) \cup C = A \cup (B \cup C).$$

(iii) *Distributivity* :

$$A\cup(B\cap C) = (A\cup B)\cap(A\cup C), \qquad \textit{and} \qquad A\cap(B\cup C) = (A\cap B)\cup(A\cap C).$$

(iv) *Idempotency* :

$$A \cap A = A, \qquad \textit{and} \qquad A \cup A = A.$$

(v) *Absorption* :

$$A \cap (A \cup B) = A, \qquad \textit{and} \qquad A \cup (A \cap B) = A.$$

(vi) *De Morgan's laws* :

$$C-(A\cap B) = (C-A)\cup(C-B), \qquad \textit{and} \qquad C-(A\cup B) = (C-A)\cap(C-B).$$

Proof. These equalities follow from Theorem 2.13, since the sets in each equation are defined by equivalent formulas. In the following parts we will only state the formulas that define each set; the equivalence of the corresponding formulas is an easy consequence of Theorem 1.6.

(i) $A \cap B$ is defined by the formula $x \in A \wedge x \in B$, and $B \cap A$ is defined by $x \in B \wedge x \in A$. Similarly, $A \cup B$ is defined by the formula $x \in A \vee x \in B$, and $B \cup A$ is defined by $x \in B \vee x \in A$.

(ii) $(A \cap B) \cap C$ is defined by the formula $(x \in A \wedge x \in B) \wedge x \in C$, and $A \cap (B \cap C)$ is defined by $x \in A \wedge (x \in B \wedge x \in C)$. Similarly, $(A \cup B) \cup C$ is defined by the formula $(x \in A \vee x \in B) \vee x \in C$, and $A \cup (B \cup C)$ is defined by $x \in A \vee (x \in B \vee x \in C)$.

(iii) $A \cup (B \cap C)$ is defined by the formula $x \in A \vee (x \in B \wedge x \in C)$, and $(A \cup B) \cap (A \cup C)$ is defined by $(x \in A \vee x \in B) \wedge (x \in A \vee x \in C)$. Similarly, $A\cap(B\cup C)$ is defined by the formula $x \in A\wedge(x \in B\vee x \in C)$, and $(A\cap B)\cup(A\cap C)$ is defined by $(x \in A \wedge x \in B) \vee (x \in A \wedge x \in C)$.

(iv) $A \cap A$ is defined by the formula $x \in A \wedge x \in A$, which is equivalent to the formula $x \in A$. Similarly, $A \cup A$ is defined by the formula $x \in A \vee x \in A$, which is also equivalent to the formula $x \in A$. However, by the axiom of extensionality the formula $x \in A$ defines the set A. Hence $A \cap A$ and $A \cup A$ are equal to A.

(v) $A \cap (A \cup B)$ is defined by the formula $x \in A \wedge (x \in A \vee x \in B)$, which is equivalent to the formula $x \in A$. Similarly, $A \cup (A \cap B)$ is defined by the formula $x \in A \vee (x \in A \wedge x \in B)$, which is also equivalent to the formula $x \in A$. Finally note that by the axiom of extensionality the formula $x \in A$ defines the set A.

(vi) $C - (A \cap B)$ is defined by the formula $x \in C \wedge x \notin A \cap B$, which is just a shorthand notation for the formula $x \in C \wedge \neg(x \in A \cap B)$. Now we have

$$\begin{aligned} x \in C \wedge \neg(x \in A \cap B) &\equiv x \in C \wedge \neg(x \in A \wedge x \in B) \\ &\equiv x \in C \wedge \big(\neg(x \in A) \vee \neg(x \in B)\big) \\ &\equiv x \in C \wedge (x \notin A \vee x \notin B) \\ &\equiv (x \in C \wedge x \notin A) \vee (x \in C \wedge x \notin B). \end{aligned}$$

The last formula in the above expression defines the set $(C - A) \cup (C - B)$, as desired. Similarly, $C - (A \cup B)$ is defined by the formula $x \in C \wedge x \notin A \cup B$, which is just a shorthand notation for the formula $x \in C \wedge \neg(x \in A \cup B)$. Now we have

$$\begin{aligned} x \in C \wedge \neg(x \in A \cup B) &\equiv x \in C \wedge \neg(x \in A \vee x \in B) \\ &\equiv x \in C \wedge \big(\neg(x \in A) \wedge \neg(x \in B)\big) \\ &\equiv (x \in C \wedge x \in C) \wedge (x \notin A \wedge x \notin B) \\ &\equiv \big((x \in C \wedge x \in C) \wedge x \notin A\big) \wedge x \notin B \\ &\equiv \big(x \in C \wedge (x \in C \wedge x \notin A)\big) \wedge x \notin B \\ &\equiv \big(x \in C \wedge (x \notin A \wedge x \in C)\big) \wedge x \notin B \\ &\equiv (x \in C \wedge x \notin A) \wedge (x \in C \wedge x \notin B). \end{aligned}$$

The last formula in the above expression defines the set $(C - A) \cap (C - B)$, as desired. ■

Remark. Consider the set $(A \cap B) \cap C$. We know that it equals $A \cap (B \cap C)$. Thus we can drop the parentheses, and denote this set simply by $A \cap B \cap C$. Similarly, we can write $A \cup B \cup C$ to denote the set $(A \cup B) \cup C$. The associativity also implies that if several sets are all connected by union or intersection, then the arrangement of parentheses between them does not alter the resulting set. In other words, all the possible arrangements of parentheses result in the same set. So for example, $D \cup ((A \cup B) \cup C)$ is equal to $(D \cup A) \cup (B \cup C)$. We can denote the set denoted by these expressions simply by $D \cup A \cup B \cup C$. Similar abbreviated notations can be used when we have more sets. However, we do not have the tools now to state the general version of this fact precisely, and to prove it rigorously. See Sections 4.5 and 5.6 for the details of the general version.

We can also state a version of some parts of the above theorem for the unions and intersections of arbitrary families of sets.

Theorem 2.18. *Suppose $\mathcal{A}, \mathcal{B}, C$ are sets, and $\mathcal{A}$ is nonempty. Then we have*

(i) *Distributivity* :

$$\Big(\bigcap_{A \in \mathcal{A}} A\Big) \cup C = \bigcap_{A \in \mathcal{A}} (A \cup C), \qquad \textit{and} \qquad \Big(\bigcup_{B \in \mathcal{B}} B\Big) \cap C = \bigcup_{B \in \mathcal{B}} (B \cap C).$$

(ii)

$$\Big(\bigcap_{A \in \mathcal{A}} A\Big) \cap C = \bigcap_{A \in \mathcal{A}} (A \cap C), \qquad \textit{and} \qquad \Big(\bigcup_{A \in \mathcal{A}} A\Big) \cup C = \bigcup_{A \in \mathcal{A}} (A \cup C).$$

(iii) *De Morgan's laws* :

$$C - \Big(\bigcap_{A \in \mathcal{A}} A\Big) = \bigcup_{A \in \mathcal{A}} (C - A), \qquad \textit{and} \qquad C - \Big(\bigcup_{A \in \mathcal{A}} A\Big) = \bigcap_{A \in \mathcal{A}} (C - A).$$

Remark. Note that part (ii) of this theorem is not a case of associativity, rather, it states the distributivity of $\cap$ over $\cap$, and of $\cup$ over $\cup$.

Remark. The set $\bigcap_{A\in\mathcal{A}}(A\cup C)$ is the intersection of all the elements of a set $\mathcal{S}$ whose elements are of the form $A\cup C$ for some $A\in\mathcal{A}$. More precisely, let

$$\mathcal{S} := \{S\in\mathcal{P}(C\cup\bigcup\mathcal{A}) : \exists A\in\mathcal{A}\ (S = A\cup C)\}.$$

Then by the axiom of separation we have

$$S\in\mathcal{S} \equiv \big(S\in\mathcal{P}(C\cup\bigcup\mathcal{A})\big)\wedge\big(\exists A\in\mathcal{A}\ (S = A\cup C)\big).$$

However, we can easily show that $\exists A\in\mathcal{A}\ (S = A\cup C)\vdash S\in\mathcal{P}(C\cup\bigcup\mathcal{A})$ (see Theorem 2.19). Hence we have

$$S\in\mathcal{S}\equiv\exists A\in\mathcal{A}\ (S = A\cup C).$$

Thus by Theorem 2.15 we get $\bigcap\mathcal{S} = \bigcap_{A\in\mathcal{A}}(A\cup C)$. Similar remarks apply to the other similar expressions in the theorem.

Proof. These equalities follow from Theorem 2.13, since the sets in each equation are defined by equivalent formulas.

(i) $\big(\bigcap_{A\in\mathcal{A}}A\big)\cup C$ is defined by the formula $(\forall A\in\mathcal{A}\ x\in A)\vee x\in C$, which by Theorem 1.12 is equivalent to

$$\forall A\in\mathcal{A}\ (x\in A\vee x\in C).$$

However, this last formula is equivalent to $\forall A\in\mathcal{A}\ (x\in A\cup C)$, which defines the set $\bigcap_{A\in\mathcal{A}}(A\cup C)$ by Theorem 2.15.

Similarly, $\big(\bigcup_{B\in\mathcal{B}}B\big)\cap C$ is defined by the formula $(\exists B\in\mathcal{B}\ x\in B)\wedge x\in C$, which by Theorem 1.12 is equivalent to

$$\exists B\in\mathcal{B}\ (x\in B\wedge x\in C).$$

But this last formula is equivalent to $\exists B\in\mathcal{B}\ (x\in B\cap C)$, which defines the set $\bigcup_{B\in\mathcal{B}}(B\cap C)$ by Theorem 2.15.

(ii) $\big(\bigcap_{A\in\mathcal{A}}A\big)\cap C$ is defined by the formula $(\forall A\in\mathcal{A}\ x\in A)\wedge x\in C$, which by Theorem 1.12 is equivalent to

$$\forall A\in\mathcal{A}\ (x\in A\wedge x\in C).$$

However, this last formula is equivalent to $\forall A\in\mathcal{A}\ (x\in A\cap C)$, which defines the set $\bigcap_{A\in\mathcal{A}}(A\cap C)$ by Theorem 2.15.

Similarly, $\big(\bigcup_{A\in\mathcal{A}}A\big)\cup C$ is defined by the formula $(\exists A\in\mathcal{A}\ x\in A)\vee x\in C$, which by Theorem 1.12 is equivalent to

$$\exists A\in\mathcal{A}\ (x\in A\vee x\in C).$$

But this last formula is equivalent to $\exists A \in \mathcal{A}\ (x \in A \cup C)$, which defines the set $\bigcup_{A\in\mathcal{A}}(A \cup C)$ by Theorem 2.15.

(iii) $C - \left(\bigcap_{A\in\mathcal{A}} A\right)$ is defined by the formula $x \in C \wedge \neg(\forall A \in \mathcal{A}\ x \in A)$. Now we have

$$\begin{aligned} x \in C \wedge \neg(\forall A \in \mathcal{A}\ x \in A) &\equiv x \in C \wedge \neg\big(\forall A(A \in \mathcal{A} \to x \in A)\big) \\ &\equiv x \in C \wedge \big(\exists A \neg(A \in \mathcal{A} \to x \in A)\big) \\ &\equiv x \in C \wedge \big(\exists A(A \in \mathcal{A} \wedge x \notin A)\big) \\ &\equiv x \in C \wedge (\exists A \in \mathcal{A}\ x \notin A) \\ &\equiv \exists A \in \mathcal{A}\ (x \in C \wedge x \notin A) \qquad \text{(Theorem 1.12)} \\ &\equiv \exists A \in \mathcal{A}\ (x \in C - A). \end{aligned}$$

The last formula in the above expression defines the set $\bigcup_{A\in\mathcal{A}}(C - A)$ by Theorem 2.15, as desired.

Similarly, $C - \left(\bigcup_{A\in\mathcal{A}} A\right)$ is defined by the formula $x \in C \wedge \neg(\exists A \in \mathcal{A}\ x \in A)$. Now we have

$$\begin{aligned} x \in C \wedge \neg(\exists A \in \mathcal{A}\ x \in A) &\equiv x \in C \wedge \neg\big(\exists A(A \in \mathcal{A} \wedge x \in A)\big) \\ &\equiv x \in C \wedge \big(\forall A \neg(A \in \mathcal{A} \wedge x \in A)\big) \\ &\equiv x \in C \wedge \big(\forall A(A \notin \mathcal{A} \vee x \notin A)\big) \\ &\equiv x \in C \wedge \big(\forall A(A \in \mathcal{A} \to x \notin A)\big) \\ &\equiv x \in C \wedge (\forall A \in \mathcal{A}\ x \notin A) \\ &\equiv \forall A \in \mathcal{A}\ (x \in C \wedge x \notin A) \qquad \text{(Theorem 1.12)} \\ &\equiv \forall A \in \mathcal{A}\ (x \in C - A). \end{aligned}$$

The last formula in the above expression defines the set $\bigcap_{A\in\mathcal{A}}(C - A)$ by Theorem 2.15, as desired. ■

Remark. Note that the fact that $\mathcal{A}$ is nonempty is needed when we deal with its intersection, or the intersection of other families of sets constructed from $\mathcal{A}$. In addition, the nonemptiness of $\mathcal{A}$ is needed in part (ii) and the second equality of part (iii) so that we can apply Theorem 1.12. An interesting counterexample is when $\mathcal{A}$ is empty in the second equality of (ii); in this case the left hand side of the equality is equal to C, while its right hand side is equal to $\emptyset$.

***Exercise* 2.1.** Show that for any two sets A, B we have

$$\begin{aligned} A - B &= A - (A \cap B), \\ A\Delta B &= B\Delta A, \\ A\Delta B &= (A \cup B) - (A \cap B). \end{aligned}$$

***Exercise* 2.2.** Suppose $\mathcal{A}, \mathcal{B}, \mathcal{A}', \mathcal{B}'$ are sets, and $\mathcal{A}, \mathcal{A}'$ are nonempty. Show that we have

$$\bigcup(\mathcal{B} \cup \mathcal{B}') = \big(\bigcup\mathcal{B}\big) \cup \big(\bigcup\mathcal{B}'\big), \qquad \text{and} \qquad \bigcap(\mathcal{A} \cup \mathcal{A}') = \big(\bigcap\mathcal{A}\big) \cap \big(\bigcap\mathcal{A}'\big).$$

Theorem 2.19. *Suppose A, B, C are sets. Then we have*

(i) *Reflexivity* : $A \subset A$.
(ii) *Antisymmetry* : $A \subset B$ *and* $B \subset A$ *if and only if* $A = B$.
(iii) *Transitivity* : *If* $A \subset B$ *and* $B \subset C$ *then* $A \subset C$.
(iv)
$$\emptyset \subset A, \qquad A \cap B \subset A, B, \qquad A, B \subset A \cup B.$$
(v) *If* $A \subset C$ *and* $B \subset C$ *then* $A \cup B \subset C$.
(vi) *If* $C \subset A$ *and* $C \subset B$ *then* $C \subset A \cap B$.
(vii)
$$A \subset B \iff A \cap B = A \iff A \cup B = B \iff A - B = \emptyset.$$
(viii) *If* $A \subset B$ *then* $\bigcup A \subset \bigcup B$.
(ix) *If* $A \subset B$, *and* A *is nonempty, then* $\bigcap A \supset \bigcap B$.
(x) *For every* $b \in B$ *we have* $b \subset \bigcup B$.
(xi) *If* A *is nonempty, then for every* $a \in A$ *we have* $\bigcap A \subset a$.
(xii) $\bigcup\mathcal{P}(B) = B$; *so if* $A \subset \mathcal{P}(B)$ *then* $\bigcup A \subset B$. *In addition, if* A *is nonempty we have* $\bigcap A \subset B$.

Proof. These relations follow from Theorem 2.14. Because the subset in each relation is defined by a formula which implies the formula defining the superset in the relation.

(i) It is obvious that $x \in A \vdash x \in A$. And by the axiom of extensionality the formula $x \in A$ defines the set A.

(ii) This has been shown in Example 2.2.

(iii) If $x \in A \vdash x \in B$ and $x \in B \vdash x \in C$, then $x \in A \vdash x \in C$ by the cut rule.

(iv) $\emptyset \subset A$ has been proved in Example 2.3. Next, $A \cap B$ is defined by the formula $x \in A \wedge x \in B$, and by E$\wedge$ we have

$$x \in A \wedge x \in B \vdash x \in A, \qquad x \in A \wedge x \in B \vdash x \in B.$$

Finally, $A \cup B$ is defined by the formula $x \in A \vee x \in B$, and by I$\vee$ we have

$$x \in A \vdash x \in A \vee x \in B, \qquad x \in B \vdash x \in A \vee x \in B.$$

(v) If $x \in A \vdash x \in C$ and $x \in B \vdash x \in C$, then $x \in A \vee x \in B \vdash x \in C$ by E$\vee$.

(vi) If $x \in C \vdash x \in A$ and $x \in C \vdash x \in B$, then $x \in C \vdash x \in A \wedge x \in B$ by I$\wedge$ and the cut rule.

(vii) We will use the previous parts in the proof of this part without explicit citation. If $A \subset B$ then we have $A \subset A \cap B$, since $A \subset A$. On the other hand we have $A \cap B \subset A$. Hence we get $A \cap B = A$. Conversely, if $A \cap B = A$ then we have $A \subset B$, since we know that $A \cap B \subset B$.

Next, note that if $A \subset B$ then we have $A \cup B \subset B$, since $B \subset B$. On the other hand we have $B \subset A \cup B$. Hence we get $A \cup B = B$. Conversely, if $A \cup B = B$ then we have $A \subset B$, since we know that $A \subset A \cup B$.

Finally, note that if $A \subset B$ then we have $x \in A \vdash x \in B$. On the other hand, $A - B$ is defined by the formula $x \in A \wedge x \notin B$. However, by E$\wedge$ we have

$$x \in A \wedge x \notin B \vdash x \notin B, \qquad x \in A \wedge x \notin B \vdash x \in A \vdash x \in B.$$

Thus by E$\neg$ we get $x \in A \wedge x \notin B \vdash \bot$. In other words, $x \in A - B$ implies a contradiction. Hence by I$\neg$ we get $\vdash x \notin A - B$. Therefore by I$\forall$ we obtain

$$\vdash \forall x(x \notin A - B),$$

i.e. $A - B$ is the empty set.

Conversely, suppose $A - B = \emptyset$. Then we have $\vdash \forall x(x \notin A - B)$. Thus by E$\forall$ we get $\vdash x \notin A - B$. So by weakening rule we also have $x \in A, x \notin B \vdash x \notin A - B$. Now by I$\wedge$ we have

$$x \in A, x \notin B \vdash x \in A \wedge x \notin B \vdash x \in A - B.$$

Hence by E$\neg$ we obtain $x \in A, x \notin B \vdash \bot$. Therefore by RAA we get $x \in A \vdash x \in B$, as desired.

(viii) We know that $x \in A \vdash x \in B$. Now by I$\wedge$ and I$\exists$ we have

$$x \in A \wedge z \in x \vdash x \in B \wedge z \in x \vdash \exists x(x \in B \wedge z \in x).$$

Hence by E$\exists$ we get $\exists x(x \in A \wedge z \in x) \vdash \exists x(x \in B \wedge z \in x)$. Note that these formulas define $\bigcup A, \bigcup B$ respectively.

(ix) We know that $x \in A \vdash x \in B$. Note that B is nonempty since A is nonempty. So $\bigcap A, \bigcap B$ are defined. Now by E$\forall$ we have $\forall x(x \in B \to z \in x) \vdash x \in B \to z \in x$. Hence by the cut rule and E$\to$ we get

$$\forall x(x \in B \to z \in x), x \in A \vdash z \in x.$$

Thus by I$\to$ we have $\forall x(x \in B \to z \in x) \vdash x \in A \to z \in x$. Therefore by I$\forall$ we obtain

$$\forall x(x \in B \to z \in x) \vdash \forall x(x \in A \to z \in x),$$

as desired. Note that the above formulas define $\bigcap B, \bigcap A$ respectively.

(x) By I$\wedge$ and I$\exists$ we have $b \in B, z \in b \vdash b \in B \wedge z \in b \vdash \exists x(x \in B \wedge z \in x)$. However we assumed that $\vdash b \in B$. Thus by the cut rule we get

$$z \in b \vdash \exists x(x \in B \wedge z \in x),$$

as desired.

(xi) By E$\forall$ we have $\forall x(x \in A \to z \in x) \vdash a \in A \to z \in a$. Hence by E$\to$ we get

$$\forall x(x \in A \to z \in x), a \in A \vdash z \in a.$$

But we assumed that $\vdash a \in A$. Hence by the cut rule we obtain

$$\forall x(x \in A \to z \in x) \vdash z \in a,$$

as desired.

(xii) We know that $B \subset B$, so $B \in \mathcal{P}(B)$. Hence we have $B \subset \bigcup \mathcal{P}(B)$. On the other hand, $z \in \bigcup \mathcal{P}(B)$ means that $\exists x(x \in \mathcal{P}(B) \wedge z \in x)$; which is equivalent to $\exists x(x \subset B \wedge z \in x)$. However, $x \subset B$ implies that $\vdash z \in x \to z \in B$. Thus by E$\wedge$ and E$\to$ we get $x \subset B \wedge z \in x \vdash z \in B$. Hence by E$\exists$ we obtain

$$\exists x(x \subset B \wedge z \in x) \vdash z \in B.$$

In other words, we have shown that $\bigcup \mathcal{P}(B) \subset B$. So we get the desired equality. Now, if $A \subset \mathcal{P}(B)$ then we have $\bigcup A \subset \bigcup \mathcal{P}(B) = B$. In addition, when A is nonempty, for some $a \in A$ we have $\bigcap A \subset a \subset B$, since the elements of A are subsets of B. ■

It frequently happens in mathematics that we only consider sets which are subsets of a given fixed set. This fixed set plays the role of universe in our study of that subject. Suppose M is the set that we consider as the universe. Then for every $A \subset M$ we define the **complement** of A to be

$$A^c := M - A.$$

Note that the notion of complement depends on the universe that we considered.

Theorem 2.20. *Suppose M is a set, and $A, B \subset M$. Also suppose that $X \subset \mathcal{P}(M)$ is nonempty. Then we have*

(i) *De Morgan's laws* :

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

(ii) *De Morgan's laws* :

$$\Big(\bigcap_{C \in X} C\Big)^c = \bigcup_{C \in X} C^c, \qquad \text{and} \qquad \Big(\bigcup_{C \in X} C\Big)^c = \bigcap_{C \in X} C^c.$$

(iii)
$$(A^c)^c = A.$$

(iv)
$$B - A = B \cap A^c.$$

(v)
$$B - A = A^c - B^c.$$

(vi)
$$A \subset B \iff B^c \subset A^c.$$

(vii)
$$\begin{aligned} M^c = \emptyset, &\quad \textit{and} \quad \emptyset^c = M, \\ A \cup A^c = M, &\quad \textit{and} \quad A \cap A^c = \emptyset, \\ A \cap M = A, &\quad \textit{and} \quad A \cup \emptyset = A, \\ A \cup M = M, &\quad \textit{and} \quad A \cap \emptyset = \emptyset. \end{aligned}$$

Proof. **(i)** This easily follows from the previous statement of De Morgan's laws.

(ii) This easily follows from the previous statement of De Morgan's laws for families of sets. Note that the elements of X are subsets of M. Also, $\bigcap X$ and $\bigcup X$ are subsets of M by the previous theorem. Hence we are allowed to form their complements.

(iii) We know that $(A^c)^c = M - (M - A)$. Hence $x \in (A^c)^c$ means that

$$\begin{aligned} x \in M \wedge x \notin M - A &\equiv x \in M \wedge \neg(x \in M \wedge x \notin A) \\ &\equiv x \in M \wedge (x \notin M \vee x \in A) \\ &\equiv (x \in M \wedge x \notin M) \vee (x \in M \wedge x \in A) \\ &\equiv \bot \vee (x \in M \wedge x \in A) \equiv x \in M \wedge x \in A. \end{aligned}$$

However, we know that $x \in A \vdash x \in M$, since $A \subset M$. Hence by I$\wedge$ we have $x \in A \vdash x \in M \wedge x \in A$. Also, by E$\wedge$ we have $x \in M \wedge x \in A \vdash x \in A$. Thus

$$x \in M \wedge x \in A \equiv x \in A.$$

So $x \in (A^c)^c \equiv x \in A$. Therefore $(A^c)^c = A$ by Theorem 2.13.

(iv) $B \cap A^c$ is defined by the formula

$$x \in B \wedge x \in A^c \equiv x \in B \wedge (x \in M \wedge x \notin A) \equiv (x \in B \wedge x \in M) \wedge x \notin A.$$

However, as we have shown in the proof of the previous part, we have

$$x \in B \wedge x \in M \equiv x \in B,$$

since $B \subset M$. Therefore $x \in B \wedge x \in A^c \equiv x \in B \wedge x \notin A$; and this last formula defines the set $B - A$. Hence by Theorem 2.13 we have $B \cap A^c = B - A$, as desired.

(v) We have

$$B - A = B \cap A^c = A^c \cap B = A^c \cap (B^c)^c = A^c - B^c.$$

(vi) By Theorem 2.19 we know that $A \subset B$ is equivalent to $B - A = \emptyset$. Similarly, $B^c \subset A^c$ is equivalent to $A^c - B^c = \emptyset$. However, we know that $B - A = A^c - B^c$. Thus $A \subset B$ is equivalent to $B^c \subset A^c$.

(vii) We have $M^c = M - M = \emptyset$, since $M \subset M$. Hence we get

$$\emptyset^c = (M^c)^c = M.$$

Now note that $x \in M, x \notin A \vdash x \in M \wedge x \notin A$, and $x \in M \wedge x \notin A \equiv x \in A^c$. Hence by I$\rightarrow$ we have $x \in M \vdash x \notin A \rightarrow x \in A^c$. But by the law of material implication we have

$$x \notin A \rightarrow x \in A^c \equiv x \in A \vee x \in A^c \equiv x \in A \cup A^c.$$

Thus $x \in M \vdash x \in A \cup A^c$. Therefore $M \subset A \cup A^c$. On the other hand, $A \cup A^c \subset M$, since $A, A^c \subset M$. So $A \cup A^c = M$.

Next we have $x \in A \cap A^c \vdash x \in A \wedge (x \in M \wedge x \notin A)$. But

$$x \in A \wedge (x \in M \wedge x \notin A) \equiv (x \in A \wedge x \notin A) \wedge x \in M \equiv \bot,$$

since the contradiction $x \in A \wedge x \notin A$ is equivalent to $\bot$, and the conjunction of $\bot$ with any formula is equivalent to $\bot$. Thus $x \in A \cap A^c \vdash \bot$. Hence by I$\neg$ we get $\vdash x \notin A \cap A^c$. Therefore by I$\forall$ we obtain

$$\vdash \forall x (x \notin A \cap A^c),$$

i.e. $A \cap A^c$ is the empty set, as desired.

Finally, note that the other equalities of this part follow from Theorem 2.19, since we have $\emptyset \subset A \subset M$. Note that we can also interpret $x \in \emptyset$ to be the false formula $x \neq x$, and compute $A \cup \emptyset, A \cap \emptyset$ directly. ■

Remark. As a consequence of the above theorem, for every set M we have

$$M - \emptyset = \emptyset^c = M.$$

We can also interpret $x \in \emptyset$ to be the false formula $x \neq x$, and compute $M - \emptyset$ directly.

Chapter 3

Relations and Functions

3.1 Ordered Pairs

Let a, b be two sets. The axiom of pairing allows us to form the set $\{a, b\}$. So we can form a pair whose components are a, b. But this notion of pair is inadequate for many applications, because the sets a, b have no order in the pair $\{a, b\}$. Remember that by the axiom of extensionality we have

$$\{a, b\} = \{b, a\},$$

since both of these sets contain exactly a, b. The following elegant definition of pair, due to Kuratowski, captures the idea of order.

Definition 3.1. Let a, b be two sets. The **ordered pair** (a, b) is

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

The set a, b are called the **components** of the ordered pair (a, b).

We have to use the axiom of pairing to construct $\{a, b\}$, $\{a\}$, and $\{\{a\}, \{a, b\}\}$. Note that the ordered pair (a, b) is a set that contains the unordered pair $\{a, b\}$, and also contains the first component $\{a\}$, so that it can record the order of the pair a, b. The next theorem makes this intuition precise.

Theorem 3.1. *Let a, b, c, d be sets. Then we have*

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

Proof. If $a = c$ and $b = d$, then by the axiom of extensionality we have $\{a\} = \{c\}$, and $\{a, b\} = \{c, d\}$. Hence we also have $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$, as desired. Conversely, suppose that $(a, b) = (c, d)$. Then we have $\{a\} \in \{\{c\}, \{c, d\}\}$. There

are two possible cases. If $\{a\} = \{c, d\}$, then $\{c, d\}$ is a singleton. Thus we must have $c = d$. Therefore $\{c, d\} = \{c\}$. We also have $a \in \{c\}$; so $a = c$. In addition, $(c, d) = \{\{c\}\}$ is also a singleton. Hence (a, b) must be a singleton too. Thus we must have $\{a, b\} = \{a\}$. Therefore $b \in \{a\}$, and we obtain that $b = a$. So we have shown that $a = c$, and $b = a = c = d$.

Now let us consider the other case. Suppose we have $\{a\} = \{c\}$. Then we get $a = c$. On the other hand, we also have $\{a, b\} \in \{\{c\}, \{c, d\}\}$. If $\{a, b\} = \{c\}$ then we must have $b = c = a$. Thus we can repeat the above argument by exchanging the roles of (a, b) and (c, d), to conclude that $d = c = b$. So let us suppose that $\{a, b\} = \{c, d\}$. Then we have $b \in \{c, d\}$. Again, if $b = c = a$, then we can conclude that $d = c = b$. Hence there is only one other possibility left, namely we must have $b = d$, which is the desired. ■

Note that the exact structure of the set (a, b) is not important; the important thing is that (a, b) satisfies the characteristic property expressed in the above theorem. Namely, we need (a, b) to somehow tells us that a comes first, and b comes next; and it is not important how this information is stored in (a, b).

Remark. Note that unlike unordered pairs, the ordered pairs (a, b) and (b, a) are different, unless $a = b$. If $a = b$ then we have

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

Theorem 3.2. *Let A, B be two sets. Then for every $a \in A$ and $b \in B$ we have*

$$(a, b) \in \mathcal{P}(\mathcal{P}(A \cup B)).$$

Proof. We have $a, b \in A \cup B$. Hence $\{a\}, \{a, b\} \subset A \cup B$. Thus $\{a\}, \{a, b\} \in \mathcal{P}(A \cup B)$. Therefore we have $\{\{a\}, \{a, b\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$, as desired. ■

Definition 3.2. Let A, B be two sets. Then the **Cartesian product** $A \times B$ of A, B is the set of all ordered pairs (a, b), where $a \in A$ and $b \in B$. In other words

$$A \times B := \{x \in \mathcal{P}(\mathcal{P}(A \cup B)) : \exists a \in A \text{ and } \exists b \in B \text{ such that } x = (a, b)\}.$$

Remark. The expression defining $A \times B$ can be written more formally as

$$\exists a \exists b \big(a \in A \wedge b \in B \wedge (x = (a, b))\big).$$

The existence of $A \times B$ is guaranteed by the axioms of union, power set, and separation. Note that the property $x = (a, b)$ can be expressed by a formula; so the application of the axiom of separation is legitimate. In addition, note that $A \times B$ will contain every ordered pair (a, b), where $a \in A$ and $b \in B$, due to the above theorem.

Notation. The set $A \times A$ is also denoted by A^2.

***Example* 3.1.** For every set A we have

$$A \times \emptyset = \emptyset = \emptyset \times A.$$

Because if $x \in A \times \emptyset$ then we get $\exists a \in A$ and $\exists b \in \emptyset$ such that $x = (a, b)$, which is a contradiction. Hence for every x we must have $x \notin A \times \emptyset$. Thus $A \times \emptyset = \emptyset$. The other equality can be proved similarly.

Theorem 3.3. *Suppose $A, B, C, D, \mathcal{X}$ are sets. Then we have*

(i) $A \times (B \cup D) = (A \times B) \cup (A \times D)$.
(ii) $A \times (B \cap D) = (A \times B) \cap (A \times D)$.
(iii) $A \times (B - D) = (A \times B) - (A \times D)$.
(iv) $A \times \big(\bigcup_{X \in \mathcal{X}} X\big) = \bigcup_{X \in \mathcal{X}} (A \times X)$.
(v) $A \times \big(\bigcap_{X \in \mathcal{X}} X\big) = \bigcap_{X \in \mathcal{X}} (A \times X)$, *provided that $\mathcal{X}$ is nonempty.*
(vi) *If $B \subset D$ then $A \times B \subset A \times D$.*
(vii) $(A \cap C) \times (B \cap D) = (A \times B) \cap (C \times D)$.

Remark. Similar identities to (i)-(vi) hold when we multiply A from right.

Remark. The set $\bigcup_{X \in \mathcal{X}} (A \times X)$ is the union of all the elements of a set $\mathcal{S}$ whose elements are of the form $A \times X$ for some $X \in \mathcal{X}$. More precisely, let

$$\mathcal{S} := \{S \in \mathcal{P}(A \times \bigcup \mathcal{X}) : \exists X \in \mathcal{X}\ (S = A \times X)\}.$$

Then by the axiom of separation we have

$$S \in \mathcal{S} \iff \big(S \in \mathcal{P}(A \times \bigcup \mathcal{X})\big) \wedge \big(\exists X \in \mathcal{X}\ (S = A \times X)\big).$$

However, we can easily show that $\exists X \in \mathcal{X}\ (S = A \times X) \implies S \in \mathcal{P}(A \times \bigcup \mathcal{X})$. Hence we have

$$S \in \mathcal{S} \iff \exists X \in \mathcal{X}\ (S = A \times X).$$

Thus by Theorem 2.15 we get $\bigcup \mathcal{S} = \bigcup_{X \in \mathcal{X}} (A \times X)$. Similar remarks apply to $\bigcap_{X \in \mathcal{X}} (A \times X)$.

Proof. To simplify the notation, we will denote $x = (a, b)$ by x_a^b.

(i) We have

$$\begin{aligned}
x \in A \times (B \cup D) &\iff \exists a \exists b \big(a \in A \wedge (b \in B \cup D) \wedge x_a^b\big) \\
&\iff \exists a \exists b \big(a \in A \wedge (b \in B \vee b \in D) \wedge x_a^b\big) \\
&\iff \exists a \exists b \big((a \in A \wedge b \in B \wedge x_a^b) \vee (a \in A \wedge b \in D \wedge x_a^b)\big) \\
&\iff \exists a \exists b \big(a \in A \wedge b \in B \wedge x_a^b\big) \vee \exists a \exists b \big(a \in A \wedge b \in D \wedge x_a^b\big) \\
&\iff \big(x \in A \times B\big) \vee \big(x \in A \times D\big) \\
&\iff x \in (A \times B) \cup (A \times D).
\end{aligned}$$

(ii) This follows from part (vii) by taking $C = A$, and noting that $A \cap A = A$.
(iii) We have

$$\begin{aligned} x \in (A \times B) - (A \times D) &\iff (x \in A \times B) \wedge (x \notin A \times D) \\ &\iff \exists a \exists b (a \in A \wedge b \in B \wedge x_a^b) \\ &\qquad \wedge \forall a \forall b (a \notin A \vee b \notin D \vee \neg x_a^b) \\ &\iff \exists a \exists b (a \in A \wedge b \in B \wedge x_a^b) \\ &\qquad \wedge \forall a \forall b (x_a^b \to a \notin A \vee b \notin D). \end{aligned}$$

We also have

$$\begin{aligned} x \in A \times (B - D) &\iff \exists a \exists b \big(a \in A \wedge (b \in B - D) \wedge x_a^b\big) \\ &\iff \exists a \exists b \big(a \in A \wedge (b \in B \wedge b \notin D) \wedge x_a^b\big) \\ &\iff \exists a \exists b \big((a \in A \wedge b \in B) \wedge x_a^b \wedge b \notin D\big). \end{aligned}$$

Now note that

$$a \in A \wedge b \in B \wedge x_a^b,\ b \notin D \implies \exists a \exists b \big((a \in A \wedge b \in B) \wedge x_a^b \wedge b \notin D\big).$$

Also, if we replace $b \notin D$ with $a \notin A$, or $\neg x_a^b$, we will have the same conclusion; because we can derive any formula from a contradiction. Hence by E$\vee$ we get

$$\begin{aligned} a \in A \wedge b \in B \wedge x_a^b,\ a \notin A \vee b \notin D \vee \neg x_a^b \\ \implies \exists a \exists b \big((a \in A \wedge b \in B) \wedge x_a^b \wedge b \notin D\big). \end{aligned}$$

Thus by applying E$\forall$, and then applying E$\exists$ and E$\wedge$, we obtain

$$\begin{aligned} \exists a \exists b (a \in A \wedge b \in B \wedge x_a^b) \wedge \forall a \forall b (a \notin A \vee b \notin D \vee \neg x_a^b) \\ \implies \exists a \exists b \big((a \in A \wedge b \in B) \wedge x_a^b \wedge b \notin D\big). \end{aligned}$$

Next, let us prove the reverse of the above implication. First note that

$$(a \in A \wedge b \in B) \wedge x_a^b \wedge b \notin D \implies \exists a \exists b \big(a \in A \wedge b \in B \wedge x_a^b\big).$$

Thus by E$\exists$ we get

$$\exists a \exists b \big((a \in A \wedge b \in B) \wedge x_a^b \wedge b \notin D\big) \implies \exists a \exists b \big(a \in A \wedge b \in B \wedge x_a^b\big).$$

Before we continue the formal proof, let us pause here to explain the idea of the proof intuitively. So far we have shown that $(A \times B) - (A \times D) \subset A \times (B - D)$. Also, in this paragraph, we have shown that if $x \in A \times (B - D)$ then $x \in A \times B$. We only need to show that $x \notin A \times D$. To this end, we need to prove that x is not equal

to any ordered pair in $A \times D$, i.e. we have to show that if $x = (c, d)$ for some c, d, then either $c \notin A$ or $d \in D$. However, we know that $x \in A \times (B - D)$; so $x = (a, b)$ for some $a \in A$ and $b \in B - D$. Therefore if $x = (c, d)$ then $(c, d) = (a, b)$, and we must have $c = a \in A$ and $d = b \in B - D$. In particular, we can conclude that $d \notin D$, as desired. Note that the above argument can be carried out because the components of an ordered pair are uniquely determined by it, namely the ordered pair x uniquely determines its components a, b, and we cannot represent x as a different ordered pair (c, d). Let us rigorously implement this idea in a more formal way.

Now note that $(c \in A \wedge d \in B) \wedge x_c^d \wedge d \notin D,\ x_a^b$ imply $x_a^b \wedge x_c^d$. But x_a^b means $x = (a, b)$, and x_c^d means $x = (c, d)$. Hence $x_a^b \wedge x_c^d$ implies that $c = a$ and $d = b$. Thus in any formula we can substitute c with a, and d with b. Therefore we get

$$(c \in A \wedge d \in B) \wedge x_c^d \wedge d \notin D,\ x_a^b \implies b \notin D \implies a \notin A \vee b \notin D.$$

Hence by I$\to$ and E$\exists$ we obtain

$$\exists c \exists d\big((c \in A \wedge d \in B) \wedge x_c^d \wedge d \notin D\big) \implies x_a^b \to a \notin A \vee b \notin D.$$

But by Theorem 1.11 we can change the bound variables c, d to a, b, and conclude that

$$\exists a \exists b\big((a \in A \wedge b \in B) \wedge x_a^b \wedge b \notin D\big) \implies x_a^b \to a \notin A \vee b \notin D.$$

Finally, by I$\forall$ we get

$$\exists a \exists b\big((a \in A \wedge b \in B) \wedge x_a^b \wedge b \notin D\big) \implies \forall a \forall b(x_a^b \to a \notin A \vee b \notin D).$$

At the end, we get the desired result by I$\wedge$. Note that from a syntactical pint of view, using the variables c, d on one side of the entailment, and a, b on the other side, allows us to apply the rules E$\exists$ and I$\forall$, and conclude the desired result.

(iv) We have

$$\begin{aligned}
x \in A \times \big(\textstyle\bigcup_{X \in \mathcal{X}} X\big) &\iff \exists a \exists b\big(a \in A \wedge b \in \textstyle\bigcup_{X \in \mathcal{X}} X \wedge x_a^b\big) \\
&\iff \exists a \exists b\big(a \in A \wedge (\exists X \in \mathcal{X}\ b \in X) \wedge x_a^b\big) \\
&\iff \exists a \exists b\big(\exists X \in \mathcal{X}\ (a \in A \wedge b \in X \wedge x_a^b)\big) && \text{(Thm. 1.12)} \\
&\iff \exists a \exists b \exists X\big(X \in \mathcal{X} \wedge (a \in A \wedge b \in X \wedge x_a^b)\big) \\
&\iff \exists X(X \in \mathcal{X} \wedge \exists a \exists b(a \in A \wedge b \in X \wedge x_a^b)\big) && \text{(Thm. 1.10)} \\
&\iff \exists X \in \mathcal{X}\ \big(x \in A \times X\big) \\
&\iff x \in \textstyle\bigcup_{X \in \mathcal{X}} (A \times X).
\end{aligned}$$

(v) We have

$$\begin{aligned}
x \in A \times \big(\textstyle\bigcap_{X\in\mathcal{X}} X\big) &\iff \exists a \exists b \big(a \in A \wedge b \in \textstyle\bigcap_{X\in\mathcal{X}} X \wedge x_a^b\big) \\
&\iff \exists a \exists b \big(a \in A \wedge (\forall X \in \mathcal{X}\ b \in X) \wedge x_a^b\big) \\
&\iff \exists a \exists b \big(\forall X \in \mathcal{X}\ (a \in A \wedge b \in X \wedge x_a^b)\big) \qquad \text{(Thm. 1.12)} \\
&\iff \exists a \exists b \forall X \big(X \in \mathcal{X} \to (a \in A \wedge b \in X \wedge x_a^b)\big) \\
&\iff \exists a \exists b \forall X \big(X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b)\big) \\
&\implies \forall X \big(X \notin \mathcal{X} \vee \exists a \exists b (a \in A \wedge b \in X \wedge x_a^b)\big) \quad \text{(Thm. 1.10)} \\
&\iff \forall X \big(X \in \mathcal{X} \to \exists a \exists b (a \in A \wedge b \in X \wedge x_a^b)\big) \\
&\iff \forall X \in \mathcal{X}\ \big(x \in A \times X\big) \\
&\iff x \in \textstyle\bigcap_{X\in\mathcal{X}} (A \times X).
\end{aligned}$$

Note that in order to define its intersection, and to apply Theorem 1.12, we need $\mathcal{X}$ to be nonempty.

Next, let us show that $x \in \bigcap_{X\in\mathcal{X}} (A \times X)$ implies that $x \in A \times \big(\bigcap_{X\in\mathcal{X}} X\big)$. We need to show that the reverse of the one-way implication in the above expression holds. We will use the same idea as in the proof of part (iii). First note that $c \in A \wedge d \in X \wedge x_c^d,\ x_a^b$ imply $x_a^b \wedge x_c^d$. But x_a^b means $x = (a, b)$, and x_c^d means $x = (c, d)$. Hence $x_a^b \wedge x_c^d$ implies that $c = a$ and $d = b$. Thus in any formula we can substitute c with a, and d with b. Therefore we get

$$\begin{aligned}
c \in A \wedge d \in X \wedge x_c^d,\ x_a^b &\implies a \in A \wedge b \in X \wedge x_a^b \\
&\implies X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b).
\end{aligned}$$

Hence by E$\exists$ we obtain

$$\exists c \exists d (c \in A \wedge d \in X \wedge x_c^d),\ x_a^b \implies X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b).$$

But by Theorem 1.11 we can change the bound variables c, d to a, b, and conclude that

$$\exists a \exists b (a \in A \wedge b \in X \wedge x_a^b),\ x_a^b \implies X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b).$$

In addition, we know that $X \notin \mathcal{X} \implies X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b)$. Thus by E$\vee$ and E$\forall$ we get

$$\begin{aligned}
\forall X \big(X \notin \mathcal{X} \vee \exists a \exists b (a \in A \wedge b \in X \wedge x_a^b)\big),\ x_a^b \\
\implies X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b).
\end{aligned}$$

Hence by I$\forall$ and I$\exists$ we obtain

$$\begin{aligned}
\forall X \big(X \notin \mathcal{X} \vee \exists a \exists b (a \in A \wedge b \in X \wedge x_a^b)\big),\ x_a^b \\
\implies \exists a \exists b \forall X \big(X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b)\big).
\end{aligned}$$

Now by E$\exists$ we have

$$\forall X\big(X \notin \mathcal{X} \vee \exists a\exists b(a \in A \wedge b \in X \wedge x_a^b)\big),\ \exists a\exists b\, x_a^b \\ \implies \exists a\exists b\forall X\big(X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b)\big).$$

However, if we assume $x \in \bigcap_{X\in\mathcal{X}}(A \times X)$, then $\exists a\exists b\, x_a^b$ must be true. Therefore we can conclude that

$$\forall X\big(X \notin \mathcal{X} \vee \exists a\exists b(a \in A \wedge b \in X \wedge x_a^b)\big) \\ \implies \exists a\exists b\forall X\big(X \notin \mathcal{X} \vee (a \in A \wedge b \in X \wedge x_a^b)\big),$$

as desired.

(vi) We know that $b \in B \implies b \in D$. Thus we have

$$a \in A \wedge b \in B \wedge x_a^b \implies a \in A \wedge b \in D \wedge x_a^b \implies \exists a\exists b(a \in A \wedge b \in D \wedge x_a^b).$$

Hence by E$\exists$ we get $\exists a\exists b(a \in A \wedge b \in B \wedge x_a^b) \implies \exists a\exists b(a \in A \wedge b \in D \wedge x_a^b)$; which means that $x \in A \times B \implies x \in A \times D$, as desired.

(vii) We have

$$\begin{aligned} x \in (A \cap C) \times (B \cap D) &\iff \exists a\exists b\big((a \in A \cap C) \wedge (b \in B \cap D) \wedge x_a^b\big) \\ &\iff \exists a\exists b\big((a \in A \wedge a \in C) \wedge (b \in B \wedge b \in D) \wedge (x_a^b \wedge x_a^b)\big) \\ &\iff \exists a\exists b\big((a \in A \wedge b \in B \wedge x_a^b) \wedge (a \in C \wedge b \in D \wedge x_a^b)\big) \\ &\implies \exists a\exists b\big(a \in A \wedge b \in B \wedge x_a^b\big) \wedge \exists a\exists b\big(a \in C \wedge b \in D \wedge x_a^b\big) \\ &\iff \big(x \in A \times B\big) \wedge \big(x \in C \times D\big) \\ &\iff x \in (A \times B) \cap (C \times D). \end{aligned}$$

On the other hand, similarly to the proof of part (iii), $a \in A \wedge b \in B \wedge x_a^b$ and $c \in C \wedge d \in D \wedge x_c^d$ imply $x_a^b \wedge x_c^d$. But x_a^b means $x = (a, b)$, and x_c^d means $x = (c, d)$. Hence $x_a^b \wedge x_c^d$ implies that $c = a$ and $d = b$. Thus in any formula we can substitute c with a, and d with b. Therefore we get

$$a \in A \wedge b \in B \wedge x_a^b,\ c \in C \wedge d \in D \wedge x_c^d \\ \implies (a \in A \wedge b \in B \wedge x_a^b) \wedge (a \in C \wedge b \in D \wedge x_a^b) \\ \implies \exists a\exists b\big((a \in A \wedge b \in B \wedge x_a^b) \wedge (a \in C \wedge b \in D \wedge x_a^b)\big).$$

Hence by E$\exists$ we obtain

$$\exists a\exists b\big(a \in A \wedge b \in B \wedge x_a^b\big),\ \exists c\exists d\big(c \in C \wedge d \in D \wedge x_c^d\big) \\ \implies \exists a\exists b\big((a \in A \wedge b \in B \wedge x_a^b) \wedge (a \in C \wedge b \in D \wedge x_a^b)\big).$$

Now by Theorem 1.11 we can change the bound variables c, d to a, b, and apply E$\wedge$, to conclude that

$$\exists a \exists b \big(a \in A \wedge b \in B \wedge x_a^b\big) \wedge \exists a \exists b \big(a \in C \wedge b \in D \wedge x_a^b\big)$$
$$\implies \exists a \exists b \big((a \in A \wedge b \in B \wedge x_a^b) \wedge (a \in C \wedge b \in D \wedge x_a^b)\big).$$

as desired. ■

3.2 Equivalence Relations

Remark. In the rest of these notes, we usually do not state every inference rule that we use inside a proof. Rather, we present the proofs in an informal way, similarly to the other texts on mathematics. In principle, it is always possible to convert an informal proof into a formal one; although the corresponding formal proof might be lengthy and hard to decipher. Occasionally we may provide the details of a formal proof in order to illustrate various points.

Suppose some elements of the set A have a particular relationship with some of the elements of the set B. For example some $a \in A$ are subsets of some $b \in B$. As we have seen, we can easily write a formula that says that a is a subset of b. However, sometimes the relationship between a, b cannot be described easily. Thus we need another way to describe a relation, instead of finding a formula that specifies it. An easy way to store the information of a relation is to keep all the pairs (a, b) where $a \in A$ has the desired relationship with $b \in B$. With this motivation in mind, we define a relation as follows.

Definition 3.3. Let A, B be two sets. A **relation** between A, B is a subset R of $A \times B$. When $B = A$ we say that R is a relation on A.

Notation. Suppose $R \subset A \times B$ is a relation, and $(a, b) \in R$. Then we write

$$aRb.$$

And when $(a, b) \in A \times B - R$ we write

$$a \not R b.$$

Definition 3.4. Let $R \subset A \times B$ be a relation. The **domain** of R is the set of all $a \in A$ such that $(a, b) \in R$ for some $b \in B$. In other words, the domain of R is the set

$$\{a \in A : \exists b \in B \ (a, b) \in R\}.$$

Note that by the next theorem we can characterize the domain of R solely in terms of R as follows:

$$\{a \in \bigcup(\bigcup R) : \exists b \ (a, b) \in R\}.$$

Theorem 3.4. *Suppose a, b, C are sets, and $(a, b) \in C$. Then*

$$a, b \in \bigcup(\bigcup C).$$

Proof. We have $\{a, b\} \in \{\{a\}, \{a, b\}\} \in C$. Thus $\{a, b\} \in \bigcup C$. Hence we get $a, b \in \bigcup(\bigcup C)$, as desired. ■

Definition 3.5. Let $R \subset A \times A$ be a relation on A. Then

(i) R is **reflexive** if for every $a \in A$ we have aRa.
(ii) R is **irreflexive** if for every $a \in A$ we have $a \not R a$.
(iii) R is **symmetric** if for every $a, b \in A$, aRb implies that bRa.
(iv) R is **antisymmetric** if for every $a, b \in A$, aRb and bRa imply that $a = b$.
(v) R is **transitive** if for every $a, b, c \in A$, aRb and bRc imply that aRc.

In the rest of this section, we study an important type of relations which abundantly appear throughout mathematics.

Definition 3.6. Let $R \subset X \times X$ be a relation on a set X. We say that R is an **equivalence relation** if it is reflexive, symmetric, and transitive; i.e. if for every $x, y, z \in X$ we have

(i) xRx.
(ii) If xRy then yRx.
(iii) If xRy and yRz then xRz.

Definition 3.7. Suppose R is an equivalence relation on the set X. Let $a \in X$. The **equivalence class** of a is the set

$$[a] := \{x \in X : aRx\}.$$

The set of all equivalence classes corresponding to R, i.e. the set

$$X/R := \{[x] : x \in X\},$$

is called the **quotient set** of X by R.

Remark. If we want to emphasize that an equivalence class $[a]$ corresponds to a relation R, we denote it by $[a]_R$. Also, note that $\{[x] : x \in X\}$ is a shorthand notation for the set

$$\{z \in \mathcal{P}(X) : \exists x \in X \ z = [x]\}.$$

***Example* 3.2.** Equality is an equivalence relation on any set, i.e. if for $x, y \in X$ we define xRy to mean $x = y$, then R is an equivalence relation. Note that in this case, for any $a \in X$ we have $[a] = \{a\}$.

Theorem 3.5. *Suppose R is an equivalence relation on the set X. Let $a, b \in X$. Then we have*

(i) $a \in [a]$.
(ii) aRb *if and only if* $[a] = [b]$.
(iii) $a \not R b$ *if and only if* $[a] \cap [b] = \emptyset$.

Proof. **(i)** By reflexivity of R we have aRa. Thus we have $a \in [a]$.

(ii) If $[a] = [b]$ then $b \in [a]$, since $b \in [b]$. Hence we have aRb. Conversely, suppose aRb. Then for any $x \in [b]$ we have bRx. Thus by transitivity of R we get aRx, i.e. $x \in [a]$. So $[b] \subset [a]$. In addition, by symmetry of R we also have bRa. Therefore by repeating the above argument we can show that $[b] \subset [a]$ too. Hence we have $[a] = [b]$, as desired.

(iii) Suppose $[a] \cap [b] = \emptyset$. Then we must have $b \notin [a]$, since $b \in [b]$. Thus we have $a \not R b$. Conversely, suppose $a \not R b$. Now suppose to the contrary that $[a] \cap [b] \neq \emptyset$. Let $x \in [a] \cap [b]$. Then we have aRx and bRx, since $x \in [a]$ and $x \in [b]$. Now by symmetry of R we get xRb. Hence by transitivity of R we obtain aRb, which results in a contradiction. Therefore we must have $[a] \cap [b] = \emptyset$, as desired. ■

Definition 3.8. A **partition** of a set X is a set $\mathcal{A} \subset \mathcal{P}(X)$ such that
(i) $X = \bigcup \mathcal{A}$.
(ii) The elements of $\mathcal{A}$ are nonempty, i.e. for every $A \in \mathcal{A}$ we have $A \neq \emptyset$.
(iii) The elements of $\mathcal{A}$ are pairwise disjoint sets, i.e. for every $A, B \in \mathcal{A}$ where $A \neq B$ we have $A \cap B = \emptyset$.

Theorem 3.6. *Suppose R is an equivalence relation on the set X. Then the quotient set X/R is a partition of X.*

Proof. First note that for every $[a] \in X/R$ we have $[a] \neq \emptyset$, since $a \in [a]$. Also if for $[a], [b] \in X/R$ we have $[a] \neq [b]$, then by the last theorem we must have $a \not R b$; and consequently we get $[a] \cap [b] = \emptyset$. Finally, note that $\bigcup X/R \subset X$, since $X/R \subset \mathcal{P}(X)$. On the other hand, for every $a \in X$ we have $a \in [a] \subset \bigcup X/R$. Therefore $X \subset \bigcup X/R$. Hence we have $X = \bigcup X/R$, as desired. ■

The converse of the above theorem is also true, as is shown in the next theorem.

Theorem 3.7. *Suppose $\mathcal{A}$ is a partition of a set X. Then there is an equivalence relation R on X such that $\mathcal{A} = X/R$.*

Proof. Let $R := \{(x, y) \in X^2 : \exists A \in \mathcal{A}$ such that $x, y \in A\}$. We claim that R is an equivalence relation, and $\mathcal{A} = X/R$. Let $a, b, c \in X$. Then we have $a \in X = \bigcup \mathcal{A}$; so there is $A \in \mathcal{A}$ such that $a \in A$. Hence we have aRa, since $a, a \in A$. Next suppose aRb. Then there is $A \in \mathcal{A}$ such that $a, b \in A$. Therefore we also have $b, a \in A$; so bRa too. Finally, suppose aRb and bRc. Then there are $A, B \in \mathcal{A}$ such that $a, b \in A$ and $b, c \in B$. Thus $b \in A \cap B$. So A, B are not disjoint. Therefore we must have $A = B$. Consequently we get $a, c \in A$, which implies aRc. Hence we have shown that R is an equivalence relation.

Now let us show that $\mathcal{A} = X/R$. Let $A \in \mathcal{A}$. Then A is nonempty; so there is $a \in A$. We claim that $A = [a]$. First note that by definition of R, if $b \in A$ then aRb, i.e. $b \in [a]$. Hence $A \subset [a]$. Conversely, suppose $b \in [a]$, i.e. aRb. Then there is $B \in \mathcal{A}$ such that $a, b \in B$. But $a \in A \cap B$. So A, B are not disjoint. Therefore we must have $A = B$. Hence $b \in A$. Thus $[a] \subset A$, and therefore $A = [a] \in X/R$. So $\mathcal{A} \subset X/R$. Next, consider $[a] \in X/R$. As we have shown in the first paragraph, there is $A \in \mathcal{A}$ such that $a \in A$. Then as we have shown above, we must have $A = [a]$. Hence $[a] \in \mathcal{A}$. Thus $X/R \subset \mathcal{A}$. Therefore $\mathcal{A} = X/R$, as desired. ■

3.3 Functions

Functions are ubiquitous in mathematics. One may even say that functions are the main object of study in mathematics. Intuitively, a function f is an entity that to each element x of a set X assigns a uniquely determined object $f(x)$. The following definition captures this intuitive idea nicely.

Definition 3.9. Let X, Y be two sets. A **function**, or a **map**, from X to Y is a relation $f \subset X \times Y$ such that for every $x \in X$ there exists a unique $y \in Y$ such that $(x, y) \in f$. When f is a function from X to Y we write

$$f : X \to Y.$$

If $(x, y) \in f$, then we denote y by $f(x)$, and we say that f **maps** x to y. We may also write $f : x \mapsto y$. When $y = f(x)$ we say that y is the **image** of x under f, or y is the **value** of f at x.

Remark. Note that a relation $f \subset X \times Y$ is a function from X to Y if and only if the following two conditions are satisfied:

(i) The **domain** of f is X, i.e. for every $x \in X$ there is at least one $y \in Y$ such that $(x, y) \in f$.

(ii) For every $x \in X$ there is at most one $y \in Y$ such that $(x, y) \in f$. In other words

$$\text{if } (x, y_1) \in f \text{ and } (x, y_2) \in f \text{ then } y_1 = y_2.$$

Another way that the second condition is expressed in some texts is that if $x_1 = x_2$ then $f(x_1) = f(x_2)$. However, this is somehow misleading, due to its resemblance to the basic properties of equality.

Definition 3.10. Let $f : X \to Y$ be a function. The set Y is called the **codomain** of f. The **image** of f, denoted by $f(X)$, is the set of all $y \in Y$ such that $f(x) = y$ for some $x \in X$. In other words, the image of f is the set

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

Remark. Another term related to functions, which is used in some texts, is the "*range*" of the function. Depending on the text, it can mean the codomain of the function, or the image of the function. However, in more recent texts, it usually means the image of the function. In these notes, we will not use it though, to avoid any possible confusion.

***Example* 3.3.** Let X be a set. Then the function that maps every $x \in X$ to itself is called the **identity map** of X, and is denoted by id_X. Thus we have

$$\mathrm{id}_X := \{z \in X \times X : z = (x, x) \text{ for some } x \in X\}.$$

We also have $\mathrm{id}_X(x) = x$ for every $x \in X$.

***Example* 3.4.** Let X, Y be two sets, and let $c \in Y$ be a fixed element. Then the function that maps every $x \in X$ to c is called a **constant** function. In other words, a constant function is of the form

$$\{z \in X \times Y : z = (x, c) \text{ for some } x \in X\}.$$

We sometimes write $f \equiv c$ to denote that a function f is constant and has value c.

***Example* 3.5.** Let X, Y be two sets, and suppose that $X \subset Y$. Then the function that maps every $x \in X$ to $x \in Y$ is called the **inclusion map**. In other words, the inclusion map is of the form

$$\{z \in X \times Y : z = (x, x) \text{ for some } x \in X\}.$$

Note that as a set the inclusion map is the same as id_X; however we regard the inclusion map as a subset of $X \times Y$, while we regard id_X as a subset of $X \times X$. Hence the difference lies in our viewpoint.

Remark. Similarly to the above example, when $f : X \to Y$ is a function, and $Y \subset Z$, we have

$$f \subset X \times Y \subset X \times Z.$$

So we can also view f as a function from X to Z. In particular, when $f(X) \subset Z$ we can view f as a function from X to Z, because we have

$$f \subset X \times f(X) \subset X \times Z.$$

Notation. Let X, Y, Z be sets, and let $f : X \times Y \to Z$ be a function. Then for $a = (x, y) \in X \times Y$ we usually denote $f(a)$ by $f(x, y)$.

***Example* 3.6.** Let X, Y be two sets. Let

$$\begin{aligned}\pi_1 &:= \{(z, x) \in (X \times Y) \times X : z = (x, y) \text{ for some } y \in Y\} \\ &= \{w \in (X \times Y) \times X : \exists x \in X \text{ and } \exists y \in Y \text{ such that } w = ((x, y), x)\}.\end{aligned}$$

Then π_1 is a function, since by definition, for every $z \in X \times Y$ there is $x \in X$ such that $z = (x, y)$. On the other hand, the element x is uniquely determined. Because if we also have $z = (x', y')$ then we must have $x = x'$ and $y = y'$. The function $\pi_1 : X \times Y \to X$ is called the **projection** on the first component. It satisfies

$$\pi_1(x, y) = x,$$

for every $x \in X$ and $y \in Y$. We can similarly define the projection on the second component $\pi_2 : X \times Y \to Y$ which satisfies

$$\pi_2(x, y) = y.$$

The next theorem provides a useful tool for checking whether two function are equal or not.

Theorem 3.8. *Suppose that f, g are two functions. Then $f = g$ if and only if they have the same domain X, and for every $x \in X$ we have $f(x) = g(x)$.*

Proof. Suppose the domains of the functions f, g are X, W, respectively. First assume $f = g$. Let $x \in X$. Then $(x, f(x)) \in f$. So $(x, f(x)) \in g$. Therefore x must be in the domain of g, i.e. $x \in W$. Thus $X \subset W$, because x is an arbitrary element of X. We can similarly show that $W \subset X$; so $W = X$, and f, g have the same domain. In addition, we know that $(x, g(x)) \in g$. Hence we must have $f(x) = g(x)$, since g is a function.

Conversely, suppose f, g have the same domain X, and for every $x \in X$ we have $f(x) = g(x)$. Let $(x, y) \in f$. Then we have $y = f(x) = g(x)$. Also, since x is an element of the domain of g, there is z such that $(x, z) \in g$. However, we must have $z = g(x)$; so $z = y$. Therefore $(x, y) \in g$. Hence $f \subset g$, since (x, y) is an arbitrary element of f. We can similarly show that $g \subset f$. Thus $f = g$, as desired. ■

Theorem 3.9. *Let X be a set. Suppose $\phi(x, y)$ is a formula such that for every $x \in X$ there is a unique $y \in Y$ for which $\phi(x, y)$ holds, i.e.*

$$\vdash \forall x \in X \, \exists! y \in Y \ \phi(x, y).$$

Then there is a unique function $f : X \to Y$, such that for every $x \in X$ and $y \in Y$ we have

$$f(x) = y \qquad \Longleftrightarrow \qquad \phi(x, y).$$

Remark. Informally, this theorem says that if for each element of a set X we can specify a uniquely determined value assigned to that element, then there is a unique function that takes those values.

Remark. We can weaken the assumption of this theorem and assume that for every $x \in X$ there is a unique y for which $\phi(x, y)$ holds, i.e.

$$\vdash \forall x \in X \, \exists! y \; \phi(x, y).$$

But then we need the axiom of replacement to prove that the codomain of f is a set. However, the stronger assumption is usually satisfied in the applications, and we can get along without using the axiom of replacement.

Proof. Let

$$f := \{z \in X \times Y : \exists x \in X \; \exists y \in Y \; z = (x, y) \text{ and } \phi(x, y)\}.$$

Then by our assumption, for every $x \in X$ there is $y \in Y$ such that $\phi(x, y)$; so $(x, y) \in f$. On the other hand, if $(x, y_1), (x, y_2) \in f$ then by definition we must have $y_1, y_2 \in Y$. In addition, both $\phi(x, y_1)$ and $\phi(x, y_2)$ must hold. Hence by our assumption we get $y_1 = y_2$. Therefore f is a function. Furthermore, by definition of f we have

$$f(x) = y \iff (x, y) \in f \iff \phi(x, y),$$

as desired. Finally, suppose $f, g : X \to Y$ are two functions that satisfy the above equivalence. Then for every $x \in X$ we have

$$y = f(x) \implies \phi(x, y) \implies y = g(x).$$

Thus $f(x) = y = g(x)$. Hence by Theorem 3.8 we get $f = g$. Therefore there is a unique function with our desired property. ∎

Note that by definition a function $f : X \to Y$ is a subset of $X \times Y$. Thus we have $f \in \mathcal{P}(X \times Y)$. In other words, every function from X to Y is an element of $\mathcal{P}(X \times Y)$.

Definition 3.11. Let X, Y be two sets. Then the set of all functions from X to Y is

$$Y^X := \{f \in \mathcal{P}(X \times Y) : f \text{ is a function from } X \text{ to } Y\}.$$

The existence of Y^X is guaranteed by the axioms of power set and separation. Note that the property "f is a function from X to Y" can be expressed by a formula; so the application of the axiom of separation is legitimate.

Definition 3.12. Let $f : X \to Y$ be a function. Suppose $A \subset X$ and $B \subset Y$. The **image** of A under f, denoted by $f(A)$, is the set of all $y \in Y$ such that $f(a) = y$ for some $a \in A$. In other words, the image of A under f is the set

$$f(A) := \{y \in Y : \exists a \in A \; y = f(a)\}.$$

The **inverse image** or the **preimage** of B under f, denoted by $f^{-1}(B)$, is the set of all $x \in X$ such that $f(x) \in B$. In other words, the inverse image of B under f is the set

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

Remark. More formally, the image and inverse image can be expressed as

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

Remark. Note that our notation for the image of a subset of X is similar to our notation for the image of an element of X; however, it should always be clear from the context which one is intended. Also note that our notation for the image of f is compatible with our notation for the image of subsets, since the image of f is the same as the image of X under f.

Notation. Suppose $f : X \to Y$ is a function, and ϕ is a formula. Then we use the following notation

$$\{f(x) : \phi(x)\}$$

to denote the set $\{y \in Y : \exists x \in X \; (y = f(x)) \wedge \phi(x)\}$.

For example, for $A \subset X$ we have

$$f(A) = \{f(a) : a \in A\}.$$

Because we have $f(A) = \{y \in Y : \exists a \in X \; (y = f(a)) \wedge (a \in A)\}$, since $a \in X \wedge a \in A$ is equivalent to $a \in A$.

Definition 3.13. Let $f : X \to Y$ be a function, and suppose $A \subset X$. The **restriction** of f to A is the relation

$$f \cap (A \times Y) \subset A \times Y.$$

We denote the restriction of f to A by $f|_A$.

Theorem 3.10. *Suppose $f : X \to Y$ is a function, and $A \subset X$. Then $f|_A$ is a function from A to Y, i.e.*

$$f|_A : A \to Y.$$

In addition, for every $a \in A$ we have $f|_A(a) = f(a)$.

Proof. Let $a \in A$. Then we have $(a, f(a)) \in f$, and $(a, f(a)) \in A \times Y$. Hence $(a, f(a)) \in f|_A$. Thus the domain of $f|_A$ is A. Now suppose we have $(a, y_1), (a, y_2) \in f|_A$. Then we also have $(a, y_1), (a, y_2) \in f$. Hence $y_1 = y_2$, since f is a function. Thus $f|_A$ is a function too. Finally, note that $f|_A(a) = f(a)$, since we know that $(a, f(a)) \in f|_A$. ■

Theorem 3.11. *Suppose $f : X \to Y$ is a function, and $A \subset B \subset X$. Then we have*

$$(f|_B)|_A = f|_A.$$

Proof. We have $A \times Y \subset B \times Y$. Hence we get

$$(f|_B)|_A = f|_B \cap (A \times Y) = \big(f \cap (B \times Y)\big) \cap (A \times Y) = f \cap (A \times Y) = f|_A,$$

as desired. ■

Definition 3.14. Suppose f, g are two functions. If f equals the restriction of g to some set, then we say that g is an **extension** of f.

Theorem 3.12. *Let $f : X \to Y$ be a function. Suppose that $A, A_1, A_2 \subset X$, and $B, B_1, B_2 \subset Y$. Also suppose that $\mathcal{C}, \mathcal{C}' \subset \mathcal{P}(X)$, $\mathcal{D}, \mathcal{D}' \subset \mathcal{P}(Y)$, and $\mathcal{C}', \mathcal{D}'$ are nonempty. Then we have*

(i)

$$f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2), \qquad \text{and} \qquad f\Big(\bigcap_{C \in \mathcal{C}'} C\Big) \subset \bigcap_{C \in \mathcal{C}'} f(C).$$

(ii)

$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2), \qquad \text{and} \qquad f\Big(\bigcup_{C \in \mathcal{C}} C\Big) = \bigcup_{C \in \mathcal{C}} f(C).$$

(iii)

$$f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2), \qquad \text{and} \qquad f^{-1}\Big(\bigcap_{D \in \mathcal{D}'} D\Big) = \bigcap_{D \in \mathcal{D}'} f^{-1}(D).$$

(iv)

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2), \qquad \text{and} \qquad f^{-1}\Big(\bigcup_{D \in \mathcal{D}} D\Big) = \bigcup_{D \in \mathcal{D}} f^{-1}(D).$$

(v)

$$f^{-1}(B_1 - B_2) = f^{-1}(B_1) - f^{-1}(B_2), \qquad \text{and} \qquad f^{-1}(B^c) = \big(f^{-1}(B)\big)^c.$$

(vi) *If* $A_1 \subset A_2$ *and* $B_1 \subset B_2$ *then*

$$f(A_1) \subset f(A_2), \qquad \text{and} \qquad f^{-1}(B_1) \subset f^{-1}(B_2).$$

(vii)
$$(f|_A)^{-1}(B) = A \cap f^{-1}(B).$$

(viii)
$$f(f^{-1}(B)) \subset B, \qquad \text{and} \qquad A \subset f^{-1}(f(A)).$$

Remark. Note that in part (v) the complements are with respect to Y, X. So $B^c = Y - B$, and $\big(f^{-1}(B)\big)^c = X - f^{-1}(B)$.

Remark. The set $\bigcap_{C\in\mathcal{C}'} f(C)$ is the intersection of all the elements of a set $\mathcal{S}$ whose elements are of the form $f(C)$ for some $C \in \mathcal{C}'$. More precisely, let

$$\mathcal{S} := \{S \in \mathcal{P}(Y) : \exists C \in \mathcal{C}'\ S = f(C)\}.$$

Then by the axiom of separation we have

$$S \in \mathcal{S} \iff \big(S \in \mathcal{P}(Y)\big) \wedge \big(\exists C \in \mathcal{C}'\ S = f(C)\big).$$

However, we can easily show that $\exists C \in \mathcal{C}'\ S = f(C) \implies S \in \mathcal{P}(Y)$. Hence we have

$$S \in \mathcal{S} \iff \exists C \in \mathcal{C}'\ S = f(C).$$

Thus by Theorem 2.15 we get $\bigcap \mathcal{S} = \bigcap_{C\in\mathcal{C}'} f(C)$. Similar remarks apply to the other similar expressions in the theorem.

Proof. We will provide a more formal proof for this theorem; since in this way we can see more clearly when we employ the fact that f is a function. In the following, we will use Theorem 2.13 to show the equality of two sets, and Theorem 2.14 to show the inclusion of one set in another. Also, to manipulate quantified formulas, we will use Theorems 1.10 and 1.12 frequently, without explicit citation.

(i) We have

$$\begin{aligned}
y \in f(A_1 \cap A_2) &\iff \exists a \in A_1 \cap A_2\ (a,y) \in f \\
&\iff \exists a\big(a \in A_1 \wedge a \in A_2 \wedge (a,y) \in f \wedge (a,y) \in f\big) \\
&\iff \exists a\big((a \in A_1 \wedge (a,y) \in f) \wedge (a \in A_2 \wedge (a,y) \in f)\big) \\
&\implies \exists a\big(a \in A_1 \wedge (a,y) \in f\big) \wedge \exists a\big(a \in A_2 \wedge (a,y) \in f\big) \\
&\iff y \in f(A_1) \wedge y \in f(A_2) \iff y \in f(A_1) \cap f(A_2).
\end{aligned}$$

Next, we have

$$\begin{aligned}
y \in f\big(\textstyle\bigcap_{C\in\mathcal{C}'} C\big) &\iff \exists x \in \textstyle\bigcap_{C\in\mathcal{C}'} C\ \ (x,y) \in f \\
&\iff \exists x\big((\forall C \in \mathcal{C}'\ x \in C) \wedge (x,y) \in f\big) \\
&\iff \exists x\big(\forall C \in \mathcal{C}'\ (x \in C \wedge (x,y) \in f)\big)
\end{aligned}$$

$$
\begin{aligned}
&\iff \exists x \forall C\big(C \in \mathcal{C}' \to (x \in C \wedge (x,y) \in f)\big)\\
&\implies \forall C \exists x\big(C \in \mathcal{C}' \to (x \in C \wedge (x,y) \in f)\big)\\
&\iff \forall C \exists x\big(C \notin \mathcal{C}' \vee (x \in C \wedge (x,y) \in f)\big)\\
&\iff \forall C\big(C \notin \mathcal{C}' \vee \exists x(x \in C \wedge (x,y) \in f)\big)\\
&\iff \forall C\big(C \in \mathcal{C}' \to \exists x(x \in C \wedge (x,y) \in f)\big)\\
&\iff \forall C \in \mathcal{C}'\ \exists x(x \in C \wedge (x,y) \in f)\\
&\iff \forall C \in \mathcal{C}'\ y \in f(C) \iff y \in \textstyle\bigcap_{C\in\mathcal{C}'} f(C).
\end{aligned}
$$

(ii) We have

$$
\begin{aligned}
y \in f(A_1 \cup A_2) &\iff \exists a \in A_1 \cup A_2\ (a,y) \in f\\
&\iff \exists a\big((a \in A_1 \vee a \in A_2) \wedge (a,y) \in f\big)\\
&\iff \exists a\big((a \in A_1 \wedge (a,y) \in f) \vee (a \in A_2 \wedge (a,y) \in f)\big)\\
&\iff \exists a\big(a \in A_1 \wedge (a,y) \in f\big) \vee \exists a\big(a \in A_2 \wedge (a,y) \in f\big)\\
&\iff y \in f(A_1) \vee y \in f(A_2) \iff y \in f(A_1) \cup f(A_2).
\end{aligned}
$$

We also have

$$
\begin{aligned}
y \in f\big(\textstyle\bigcup_{C\in\mathcal{C}} C\big) &\iff \exists x \in \textstyle\bigcup_{C\in\mathcal{C}} C\ \ (x,y) \in f\\
&\iff \exists x\big((\exists C \in \mathcal{C}\ x \in C) \wedge (x,y) \in f\big)\\
&\iff \exists x\big(\exists C \in \mathcal{C}\ (x \in C \wedge (x,y) \in f)\big)\\
&\iff \exists x \exists C\big(C \in \mathcal{C} \wedge (x \in C \wedge (x,y) \in f)\big)\\
&\iff \exists C \exists x\big(C \in \mathcal{C} \wedge (x \in C \wedge (x,y) \in f)\big)\\
&\iff \exists C\big(C \in \mathcal{C} \wedge \exists x(x \in C \wedge (x,y) \in f)\big)\\
&\iff \exists C \in \mathcal{C}\ \exists x(x \in C \wedge (x,y) \in f)\\
&\iff \exists C \in \mathcal{C}\ y \in f(C) \iff y \in \textstyle\bigcup_{C\in\mathcal{C}} f(C).
\end{aligned}
$$

(iii) We have

$$
\begin{aligned}
x \in f^{-1}(B_1 \cap B_2) &\iff \exists b \in B_1 \cap B_2\ (x,b) \in f\\
&\iff \exists b\big(b \in B_1 \wedge b \in B_2 \wedge (x,b) \in f \wedge (x,b) \in f\big)\\
&\iff \exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge (b \in B_2 \wedge (x,b) \in f)\big)\\
&\implies \exists b\big(b \in B_1 \wedge (x,b) \in f\big) \wedge \exists b\big(b \in B_2 \wedge (x,b) \in f\big)\\
&\iff x \in f^{-1}(B_1) \wedge x \in f^{-1}(B_2)\\
&\iff x \in f^{-1}(B_1) \cap f^{-1}(B_2).
\end{aligned}
$$

Before we continue the formal proof, let us explain the idea of the proof intuitively. So far we have shown that $f^{-1}(B_1 \cap B_2) \subset f^{-1}(B_1) \cap f^{-1}(B_2)$. We need

to show that the reverse inclusion holds too. To this end, we need to prove that if $(x,b) \in f$ for some $b \in B_1$, and $(x,c) \in f$ for some $c \in B_2$, then $(x,d) \in f$ for some $d \in B_1 \cap B_2$. However, we know that the value of the function f at x is uniquely determined by x, i.e. we must have $b = f(x) = c$. Hence b also belongs to B_2, and we get $b \in B_1 \cap B_2$. Thus we can set $d = b$, and obtain the desired. Note that this argument works because $f(x)$ is uniquely determined by x, i.e. f is a function. And if f was an arbitrary relation which was not a function, we could not carry out the above argument.

Let us rigorously implement the above idea in a more formal way. First note that $b \in B_1 \wedge (x,b) \in f,\ c \in B_2 \wedge (x,c) \in f$ imply that $b = c$, since f is a function. Thus in any formula we can substitute c with b. Therefore we get

$$\begin{aligned} & b \in B_1 \wedge (x,b) \in f,\ c \in B_2 \wedge (x,c) \in f \\ & \qquad \Longrightarrow\ (b \in B_1 \wedge (x,b) \in f) \wedge (b \in B_2 \wedge (x,b) \in f) \\ & \qquad \Longrightarrow\ \exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge (b \in B_2 \wedge (x,b) \in f)\big). \end{aligned}$$

Hence by E$\exists$ we obtain

$$\begin{aligned} & \exists b\big(b \in B_1 \wedge (x,b) \in f\big),\ \exists c\big(c \in B_2 \wedge (x,c) \in f\big) \\ & \qquad \Longrightarrow\ \exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge (b \in B_2 \wedge (x,b) \in f)\big). \end{aligned}$$

Note that from a syntactical pint of view, using the different variables b, c on the premises of the above entailment allows us to apply the rule E$\exists$, and be able to conclude the desired result below. Now by Theorem 1.11 we can change the bound variable c to b, and apply E$\wedge$, to conclude that

$$\begin{aligned} & \exists b\big(b \in B_1 \wedge (x,b) \in f\big) \wedge \exists b\big(b \in B_2 \wedge (x,b) \in f\big) \\ & \qquad \Longrightarrow\ \exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge (b \in B_2 \wedge (x,b) \in f)\big). \end{aligned}$$

as desired.

Next, we have

$$\begin{aligned} x \in f^{-1}\big(\textstyle\bigcap_{D\in\mathcal{D}'} D\big) &\iff \exists y \in \textstyle\bigcap_{D\in\mathcal{D}'} D\ \ (x,y) \in f \\ &\iff \exists y\big((\forall D \in \mathcal{D}'\ y \in D) \wedge (x,y) \in f\big) \\ &\iff \exists y\big(\forall D \in \mathcal{D}'\ (y \in D \wedge (x,y) \in f)\big) \\ &\Longrightarrow \forall D \in \mathcal{D}'\ \exists y(y \in D \wedge (x,y) \in f) \\ &\iff \forall D \in \mathcal{D}'\ x \in f^{-1}(D) \iff x \in \textstyle\bigcap_{D\in\mathcal{D}'} f^{-1}(D). \end{aligned}$$

Note that we exchanged the two quantifiers $\exists y$ and $\forall D \in \mathcal{D}'$ similarly to the proof of part (i).

Conversely, let us show that $x \in \bigcap_{D\in\mathcal{D}'} f^{-1}(D)$ implies that $x \in f^{-1}\big(\bigcap_{D\in\mathcal{D}'} D\big)$. We need to show that the reverse of the one-way implication in the above expression

holds. We will use the same idea as above. Note that $z \in D \wedge (x,z) \in f,\ (x,y) \in f$ imply that $z = y$, since f is a function. Thus in any formula we can substitute z with y. Therefore we get

$$\begin{aligned} z \in D \wedge (x,z) \in f,\ (x,y) \in f &\implies y \in D \wedge (x,y) \in f \\ &\implies D \notin \mathcal{D}' \vee (y \in D \wedge (x,y) \in f). \end{aligned}$$

Hence by E$\exists$ we obtain

$$\exists z(z \in D \wedge (x,z) \in f),\ (x,y) \in f \implies D \notin \mathcal{D}' \vee (y \in D \wedge (x,y) \in f).$$

But by Theorem 1.11 we can change the bound variable z to y, and conclude that

$$\exists y(y \in D \wedge (x,y) \in f),\ (x,y) \in f \implies D \notin \mathcal{D}' \vee (y \in D \wedge (x,y) \in f).$$

In addition, we know that $D \notin \mathcal{D}' \implies D \notin \mathcal{D}' \vee (y \in D \wedge (x,y) \in f)$. Thus by E$\vee$ and E$\forall$ we get

$$\begin{aligned} \forall D\big(D \notin \mathcal{D}' \vee \exists y(y \in D \wedge (x,y) \in f)\big),\ (x,y) \in f \\ \implies D \notin \mathcal{D}' \vee (y \in D \wedge (x,y) \in f). \end{aligned}$$

Hence by I$\forall$ and I$\exists$ we obtain

$$\begin{aligned} \forall D\big(D \notin \mathcal{D}' \vee \exists y(y \in D \wedge (x,y) \in f)\big),\ (x,y) \in f \\ \implies \exists y \forall D\big(D \notin \mathcal{D}' \vee (y \in D \wedge (x,y) \in f)\big). \end{aligned}$$

Now by E$\exists$ we have

$$\begin{aligned} \forall D\big(D \notin \mathcal{D}' \vee \exists y(y \in D \wedge (x,y) \in f)\big),\ \exists y\big((x,y) \in f\big) \\ \implies \exists y \forall D\big(D \notin \mathcal{D}' \vee (y \in D \wedge (x,y) \in f)\big). \end{aligned}$$

However, $\exists y\big((x,y) \in f\big)$ must be true, since x is in the domain of the function f. Therefore we can conclude that

$$\begin{aligned} \forall D\big(D \notin \mathcal{D}' \vee \exists y(y \in D \wedge (x,y) \in f)\big) \\ \implies \exists y \forall D\big(D \notin \mathcal{D}' \vee (y \in D \wedge (x,y) \in f)\big). \end{aligned}$$

Hence, by the law of material implication we obtain

$$\begin{aligned} \forall D\big(D \in \mathcal{D}' \to \exists y(y \in D \wedge (x,y) \in f)\big) \\ \implies \exists y \forall D\big(D \in \mathcal{D}' \to (y \in D \wedge (x,y) \in f)\big), \end{aligned}$$

as desired.

(iv) We have

$$\begin{aligned} x \in f^{-1}(B_1 \cup B_2) &\iff \exists b \in B_1 \cup B_2 \ (x,b) \in f \\ &\iff \exists b\big((b \in B_1 \vee b \in B_2) \wedge (x,b) \in f\big) \\ &\iff \exists b\big((b \in B_1 \wedge (x,b) \in f) \vee (b \in B_2 \wedge (x,b) \in f)\big) \\ &\iff \exists b\big(b \in B_1 \wedge (x,b) \in f\big) \vee \exists b\big(b \in B_2 \wedge (x,b) \in f\big) \\ &\iff x \in f^{-1}(B_1) \vee x \in f^{-1}(B_2) \\ &\iff x \in f^{-1}(B_1) \cup f^{-1}(B_2). \end{aligned}$$

We also have

$$\begin{aligned} x \in f^{-1}\big(\textstyle\bigcup_{D\in\mathcal{D}} D\big) &\iff \exists y \in \textstyle\bigcup_{D\in\mathcal{D}} D \ (x,y) \in f \\ &\iff \exists y\big((\exists D \in \mathcal{D} \ y \in D) \wedge (x,y) \in f\big) \\ &\iff \exists y\big(\exists D \in \mathcal{D} \ (y \in D \wedge (x,y) \in f)\big) \\ &\iff \exists D \in \mathcal{D} \ \exists y(y \in D \wedge (x,y) \in f) \\ &\iff \exists D \in \mathcal{D} \ x \in f^{-1}(D) \iff x \in \textstyle\bigcup_{D\in\mathcal{D}} f^{-1}(D). \end{aligned}$$

Note that we exchanged the two quantifiers $\exists y$ and $\exists D \in \mathcal{D}$ similarly to the proof of part (ii).

(v) We have

$$\begin{aligned} x \in f^{-1}(B_1) - f^{-1}(B_2) &\iff x \in f^{-1}(B_1) \wedge x \notin f^{-1}(B_2) \\ &\iff \exists b\big(b \in B_1 \wedge (x,b) \in f\big) \\ &\qquad \wedge \forall b\big(b \notin B_2 \vee (x,b) \notin f\big) \\ &\iff \exists b\big(b \in B_1 \wedge (x,b) \in f\big) \\ &\qquad \wedge \forall b\big(b \in B_2 \to (x,b) \notin f\big). \end{aligned}$$

We also have

$$\begin{aligned} x \in f^{-1}(B_1 - B_2) &\iff \exists b \in B_1 - B_2 \ (x,b) \in f \\ &\iff \exists b\big(b \in B_1 \wedge b \notin B_2 \wedge (x,b) \in f\big) \\ &\iff \exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge b \notin B_2\big). \end{aligned}$$

Now note that

$$b \in B_1 \wedge (x,b) \in f, \ b \notin B_2 \implies \exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge b \notin B_2\big).$$

Also, if we replace $b \notin B_2$ with $(x,b) \notin f$, we will have the same conclusion; because we can derive any formula from a contradiction. Hence by E$\vee$ we get

$$b \in B_1 \wedge (x,b) \in f, \ b \notin B_2 \vee (x,b) \notin f \implies \exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge b \notin B_2\big).$$

Thus by applying E$\forall$, and then applying E$\exists$ and E$\wedge$, we obtain

$$\exists b\big(b \in B_1 \wedge (x,b) \in f\big) \wedge \forall b\big(b \notin B_2 \vee (x,b) \notin f\big) \\ \implies \exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge b \notin B_2\big).$$

Next, let us prove the reverse of the above implication. We will use the same idea as in the proof of part (iii). First note that

$$(b \in B_1 \wedge (x,b) \in f) \wedge b \notin B_2 \implies \exists b\big(b \in B_1 \wedge (x,b) \in f\big).$$

Thus by E$\exists$ we get

$$\exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge b \notin B_2\big) \implies \exists b\big(b \in B_1 \wedge (x,b) \in f\big).$$

In addition, note that $(c \in B_1 \wedge (x,c) \in f) \wedge c \notin B_2,\ (x,b) \in f$ imply that $c = b$, since f is a function. Thus in any formula we can substitute c with b. Therefore we get

$$(c \in B_1 \wedge (x,c) \in f) \wedge c \notin B_2,\ (x,b) \in f \implies b \notin B_2.$$

Hence by I$\to$ and E$\exists$ we obtain

$$\exists c\big((c \in B_1 \wedge (x,c) \in f) \wedge c \notin B_2\big) \implies (x,b) \in f \to b \notin B_2.$$

But by Theorem 1.11 we can change the bound variable c to b, and conclude that

$$\exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge b \notin B_2\big) \implies (x,b) \in f \to b \notin B_2.$$

Now by I$\forall$ we get

$$\exists b\big((b \in B_1 \wedge (x,b) \in f) \wedge b \notin B_2\big) \implies \forall b\big((x,b) \in f \to b \notin B_2\big).$$

At the end, we get the desired result by I$\wedge$.

Finally note that $f^{-1}(Y) = X$, since for every $x \in X$ there is $y \in Y$ such that $(x,y) \in f$; because X is the domain of the function f. Hence we have

$$f^{-1}(B^c) = f^{-1}(Y - B) = f^{-1}(Y) - f^{-1}(B) = X - f^{-1}(B) = \big(f^{-1}(B)\big)^c.$$

(vi) Note that $a \in A_1$ implies $a \in A_2$, and $b \in B_1$ implies that $b \in B_2$. Now we have

$$y \in f(A_1) \implies \exists a \in A_1\ (a,y) \in f \implies \exists a \in A_2\ (a,y) \in f \implies y \in f(A_1).$$

We also have

$$x \in f^{-1}(B_1) \implies \exists b \in B_1\ (x,b) \in f \implies \exists b \in B_2\ (x,b) \in f \implies x \in f^{-1}(B_2).$$

(vii) We have

$$\begin{aligned}
x \in (f|_A)^{-1}(B) &\iff \exists b \in B\ (x,b) \in f|_A \\
&\iff \exists b \in B\ (x,b) \in f \cap (A \times Y) \\
&\iff \exists b \in B\ \big((x,b) \in f \wedge (x,b) \in A \times Y\big) \\
&\iff \exists b \in B\ \big((x,b) \in f \wedge x \in A \wedge b \in Y\big) \\
&\iff \big(\exists b \in B\ ((x,b) \in f \wedge b \in Y)\big) \wedge x \in A \\
&\iff \big(\exists b \in B\ (x,b) \in f\big) \wedge x \in A \qquad (\text{since } b \in B \implies b \in Y) \\
&\iff x \in f^{-1}(B) \wedge x \in A \iff x \in A \cap f^{-1}(B).
\end{aligned}$$

(viii) We have

$$\begin{aligned}
y \in f(f^{-1}(B)) &\iff \exists x \in f^{-1}(B)\ (x,y) \in f \\
&\iff \exists x\big(x \in f^{-1}(B) \wedge (x,y) \in f\big) \\
&\iff \exists x\big((\exists b \in B\ (x,b) \in f) \wedge (x,y) \in f\big) \\
&\iff \exists x\big(\exists b \in B\ ((x,b) \in f \wedge (x,y) \in f)\big) \\
&\iff \exists b \in B\ \exists x\big((x,b) \in f \wedge (x,y) \in f\big) \\
&\iff \exists b\big(b \in B \wedge \exists x\big((x,b) \in f \wedge (x,y) \in f\big)\big) \\
&\implies \exists b(b \in B \wedge y = b) \implies \exists b(y \in B) \iff y \in B.
\end{aligned}$$

Note that we exchanged the two quantifiers $\exists x$ and $\exists b \in B$ similarly to the proof of part (ii).

Next we have

$$\begin{aligned}
x \in A &\implies x \in A \wedge x = x \\
&\implies \exists a(a \in A \wedge x = a) \iff \exists a \in A\ (x = a) \\
&\implies \exists a \in A\ \big((\exists y\,(a,y) \in f) \wedge x = a\big) \qquad (\text{since } a \text{ is in the domain of } f) \\
&\iff \exists a \in A\ \big(\exists y((a,y) \in f \wedge x = a)\big) \\
&\iff \exists y\big(\exists a \in A\ ((a,y) \in f \wedge x = a)\big) \\
&\implies \exists y\big(\exists a \in A\ \big((a,y) \in f \wedge (x,y) \in f\big)\big) \\
&\iff \exists y\big(\big(\exists a \in A\ (a,y) \in f\big) \wedge (x,y) \in f\big) \\
&\iff \exists y\big(y \in f(A) \wedge (x,y) \in f\big) \\
&\iff \exists y \in f(A)\ (x,y) \in f \iff x \in f^{-1}(f(A)),
\end{aligned}$$

as desired. Also note that we exchanged the two quantifiers $\exists y$ and $\exists a \in A$ similarly to the proof of part (ii). ■

Theorem 3.13. *Suppose $f : X \to Y$ and $g : Y \to Z$ are functions such that the domain of g is the same as the codomain of f. Then there is a unique function*

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

called the ***composition*** *of g, f, that satisfies $(g \circ f)(x) = g(f(x))$ for every $x \in X$.*

Proof. First note that the uniqueness of a function from X to Z with value $g(f(x))$ at x is guaranteed by Theorem 3.8. Now let

$$h := \{w \in X \times Z : \exists x \in X \ \exists y \in Y \ \exists z \in Z \\ w = (x, z) \text{ and } (x, y) \in f \text{ and } (y, z) \in g\}.$$

Note that for every $x \in X$ we have $(x, f(x)) \in f$ and $(f(x), g(f(x))) \in g$. Hence by definition we have $(x, g(f(x))) \in h$. So the domain of h is X. Next suppose $(x, z_1), (x, z_2) \in h$. Then by definition of h there are $y_1, y_2 \in Y$ such that

$$(x, y_1) \in f, \ (y_1, z_1) \in g, \qquad \text{and} \qquad (x, y_2) \in f, \ (y_2, z_2) \in g.$$

Thus we get $y_1 = y_2$, since f is a function. Then we must have $z_1 = z_2$, since g is a function. Therefore h is a function. We also have $h(x) = g(f(x))$, since we have shown that $(x, g(f(x))) \in h$. Hence there is a unique function with our desired property. ∎

Theorem 3.14. *Suppose $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$ are functions. Then we have*

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

In other words, the composition of functions is associative. In addition, we have

$$f \circ \mathrm{id}_X = f, \qquad \mathrm{id}_Y \circ f = f.$$

Proof. Note that both $h \circ (g \circ f), (h \circ g) \circ f$ are functions from X to W. Let $x \in X$. We have

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

Hence we get $h \circ (g \circ f) = (h \circ g) \circ f$, as desired. Next, note that both $f \circ \mathrm{id}_X, \mathrm{id}_Y \circ f$ are functions from X to Y. Let $x \in X$. We have $(f \circ \mathrm{id}_X)(x) = f(\mathrm{id}_X(x)) = f(x)$. We also have $(\mathrm{id}_Y \circ f)(x) = \mathrm{id}_Y(f(x)) = f(x)$, since $f(x) \in Y$. Therefore we get the desired equalities. ∎

Remark. Sometimes we want to consider the composition of two functions $f : X \to Y_1$ and $g : Y_2 \to Z$, but Y_1, Y_2 are not equal. However, if $Y_1 \cap Y_2 \neq \emptyset$ then we can restrict our attention to $Y_1 \cap Y_2$ and form the composition of f, g. More explicitly, let

$$A := \{x \in X : f(x) \in Y_1 \cap Y_2\} = f^{-1}(Y_1 \cap Y_2).$$

Then we can form the composition of $f|_A : A \to Y_1 \cap Y_2$ and $g|_{Y_1 \cap Y_2} : Y_1 \cap Y_2 \to Z$. We usually denote $g|_{Y_1 \cap Y_2} \circ f|_A$ simply by $g \circ f$; but we keep in mind that the domain of $g \circ f$ is A.

Theorem 3.15. *Suppose $f : X \to Y$ and $g : Y \to Z$ are functions. Let $A \subset X$ and $C \subset Z$. Then we have*

$$(g \circ f)(A) = g\big(f(A)\big), \qquad \text{and} \qquad (g \circ f)^{-1}(C) = f^{-1}\big(g^{-1}(C)\big).$$

Proof. Suppose $z \in (g \circ f)(A)$. Then there is $a \in A$ such that $z = (g \circ f)(a) = g(f(a))$. However, $f(a) \in f(A)$; so we have $z = g(f(a)) \in g(f(A))$. Conversely, suppose $z \in g(f(A))$. Then there is $y \in f(A)$ such that $z = g(y)$. But we also have $y = f(a)$ for some $a \in A$. Hence $z = g(f(a)) = (g \circ f)(a)$. Thus $z \in (g \circ f)(A)$. Therefore $(g \circ f)(A) = g(f(A))$.

Next, suppose $z \in (g \circ f)^{-1}(C)$. Then $(g \circ f)(z) \in C$. Hence $g(f(z)) \in C$. Thus we have $f(z) \in g^{-1}(C)$. So we get $z \in f^{-1}\big(g^{-1}(C)\big)$. Conversely, suppose $z \in f^{-1}\big(g^{-1}(C)\big)$. Then $f(z) \in g^{-1}(C)$. Therefore $g(f(z)) \in C$. Hence $(g \circ f)(z) \in C$. Thus $z \in (g \circ f)^{-1}(C)$. Therefore $(g \circ f)^{-1}(C) = f^{-1}\big(g^{-1}(C)\big)$, as desired. ■

Definition 3.15. Let $f : X \to Y$ be a function. We say that f is **one-to-one**, or **injective**, if for every $x_1, x_2 \in X$ we have

$$f(x_1) = f(x_2) \implies x_1 = x_2.$$

We say that f is **onto**, or **surjective**, if $f(X) = Y$. In other words, f is onto if

$$\text{for every } y \in Y \text{ there exists } x \in X \text{ such that } f(x) = y.$$

We say that f is **bijective**, or a **one-to-one correspondence**, if it is both injective and surjective.

Theorem 3.16. *Suppose $f : X \to Y$ and $g : Y \to Z$ are functions. Then we have*

(i) *If f, g are one-to-one then $g \circ f$ is also one-to-one.*

(ii) *If f, g are onto then $g \circ f$ is also onto.*

Proof. **(i)** Let $x_1, x_2 \in X$. Suppose $(g \circ f)(x_1) = (g \circ f)(x_2)$. Then we have $g(f(x_1)) = g(f(x_2))$. Thus $f(x_1) = f(x_2)$, since g is one-to-one. Hence we get $x_1 = x_2$, since f is one-to-one. Therefore $g \circ f$ is one-to-one.

(ii) Let $z \in Z$. Then there is $y \in Y$ such that $g(y) = z$, since g is onto. Furthermore, there is $x \in X$ such that $f(x) = y$, since f is onto. Now we have

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

Thus $g \circ f$ is onto. ■

***Example* 3.7.** Suppose $A \subset X$, and $f : X \to Y$ is a function. Let $j : A \to X$ be the inclusion map. Then we have

$$f|_A = f \circ j.$$

Because the domain of both of these functions is A, and for every $x \in A$ we have $f|_A(x) = f(x) = f(j(x))$. Now note that j is one-to-one, since if $j(x) = j(y)$ for some $x, y \in A$, then we have $x = j(x) = j(y) = y$. Therefore if f is one-to-one then $f|_A$ is also one-to-one. In other words, the restriction of an injective function is injective.

Definition 3.16. Let $f : X \to Y$ be a function. We say f is **invertible** if there exists a function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$. In other words, for every $x \in X$ and $y \in Y$ we have

$$f(g(y)) = y, \qquad \text{and} \qquad g(f(x)) = x.$$

The function g is called the **inverse** of f, and is denoted by f^{-1}.

Theorem 3.17. *The inverse of an invertible function is unique.*

Proof. Suppose $f : X \to Y$ is an invertible function, and $g, h : Y \to X$ are inverses of f. Let $y \in Y$. Then we have

$$g(y) = g\big(f(h(y))\big) = h(y),$$

since $f \circ h = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$. Therefore $g = h$. ■

***Example* 3.8.** Let X be a set. Then the identity map of X is invertible, and we have

$$\mathrm{id}_X^{-1} = \mathrm{id}_X;$$

because $\mathrm{id}_X \circ \mathrm{id}_X = \mathrm{id}_X$.

Theorem 3.18. *A function is invertible if and only if it is bijective.*

Proof. Let $f : X \to Y$. First suppose f is invertible. Let $x_1, x_2 \in X$. Then we have

$$f(x_1) = f(x_2) \implies x_1 = f^{-1}(f(x_1)) = f^{-1}(f(x_2)) = x_2.$$

Thus f is one-to-one. Now let $y \in Y$. Then we have $y = f(f^{-1}(y))$. So $y \in f(X)$. Hence f is onto. Therefore f is bijective.

Conversely, suppose f is bijective. Let

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

We claim that g is a function, and is the inverse of f. Let $y \in Y$. Then there is $x \in X$ such that $y = f(x)$, since f is onto. So $(x, y) \in f$. Thus by definition $(y, x) \in g$. Therefore the domain of g is Y. Next, suppose $(y, x_1), (y, x_2) \in g$. Then we have $(x_1, y), (x_2, y) \in f$, i.e. $y = f(x_1)$ and $y = f(x_2)$. Hence $f(x_1) = f(x_2)$. Thus $x_1 = x_2$, since f is one-to-one. Therefore g is a function. Now for every $x \in X$

we have $(x, f(x)) \in f$. So by definition we have $(f(x), x) \in g$, which means that $g(f(x)) = x$. On the other hand, for every $y \in Y$ we have $(y, g(y)) \in g$. But this can only happen if we have $(g(y), y) \in f$; which means that $f(g(y)) = y$. Therefore f is invertible, and g is the inverse of f. ■

Remark. Let $f : X \to Y$ be an invertible function, and let $B \subset Y$. It is easy to show that the inverse image of B under f is equal to the image of B under f^{-1}. Hence there is no ambiguity in the notation $f^{-1}(B)$. However, note that the notation $g^{-1}(A)$ for the inverse image of a set A under a function g has no indication that g is invertible.

Theorem 3.19. *Suppose $f : X \to Y$ and $g : Y \to Z$ are invertible functions. Then we have*

(i) *$f^{-1} : Y \to X$ is invertible, and*

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

(ii) *$g \circ f : X \to Z$ is invertible, and*

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

Remark. As a consequence of this theorem and the previous theorem, we get that if f, g are bijective then f^{-1} and $g \circ f$ are also bijective.

Proof. **(i)** We know that $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$. Therefore f^{-1} is invertible by definition. In addition, by the uniqueness of inverse we must have $(f^{-1})^{-1} = f$, as desired.

(ii) Note that $g \circ f : X \to Z$. Also note that $f^{-1} : Y \to X$, and $g^{-1} : Z \to Y$. Therefore we can form the composition $f^{-1} \circ g^{-1}$, and we have $f^{-1} \circ g^{-1} : Z \to X$. Now by associativity of composition of functions we have

$$\begin{aligned}(f^{-1} \circ g^{-1}) \circ (g \circ f) &= ((f^{-1} \circ g^{-1}) \circ g) \circ f \\ &= (f^{-1} \circ (g^{-1} \circ g)) \circ f \\ &= (f^{-1} \circ \mathrm{id}_Y) \circ f = f^{-1} \circ f = \mathrm{id}_X.\end{aligned}$$

We can similarly show that $(g \circ f) \circ (f^{-1} \circ g^{-1}) = \mathrm{id}_Z$. Hence $g \circ f$ is invertible by definition. In addition, by the uniqueness of inverse we must have $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, as desired. ■

Theorem 3.20. *Suppose $f : X \to Y$ and $g : Z \to Y$ are functions that have the same codomain. If $f|_{X \cap Z} = g|_{X \cap Z}$, then $f \cup g$ is a function from $X \cup Z$ to Y. In particular, if $X \cap Z = \emptyset$ then $f \cup g$ is a function from $X \cup Z$ to Y.*

Remark. By Theorem 3.8 we have $f|_{X\cap Z} = g|_{X\cap Z}$ if and only if $f(x) = g(x)$ for every $x \in X \cap Z$.

Remark. Note that the functions f, g are also sets; so $f \cup g$ in the above theorem is the union of the sets f, g. Let us denote the function $f \cup g$ by h, and let $x \in X \cup Z$. Then we use the following notation to describe h:

$$h(x) = \begin{cases} f(x) & x \in X, \\ g(x) & x \in Z. \end{cases}$$

We usually use the above notation to define h, instead of saying that $h := f \cup g$. And we say that h is a *piecewise-defined function.*

Proof. Let $w \in X \cup Z$. If $w \in X$ then $(w, f(w)) \in f \subset f \cup g$. And if $w \in Z$ then $(w, g(w)) \in g \subset f \cup g$. Thus the domain of $f \cup g$ is $X \cup Z$. Now suppose

$$(w, y_1), (w, y_2) \in f \cup g.$$

If both $(w, y_1), (w, y_2)$ belong to f or g, then we have $y_1 = y_2$, since f and g are functions. So suppose one of $(w, y_1), (w, y_2)$ belongs to f, and the other one belongs to g; say $(w, y_1) \in f$ and $(w, y_2) \in g$. Then w must belong to the domains of both f and g, i.e. $w \in X$ and $w \in Z$. Thus $w \in X \cap Z$. Hence by our assumption we have

$$y_1 = f(w) = f|_{X\cap Z}(w) = g|_{X\cap Z}(w) = g(w) = y_2.$$

Therefore $f \cup g$ is a function, as desired. Finally, note that if $X \cap Z = \emptyset$ then we have

$$\begin{aligned} f|_{X\cap Z} &= f \cap ((X \cap Z) \times Y) = f \cap (\emptyset \times Y) = f \cap \emptyset \\ &= \emptyset = g \cap \emptyset = g \cap (\emptyset \times Y) = g \cap ((X \cap Z) \times Y) = g|_{X\cap Z}. \end{aligned}$$

Hence in this case, our previous assumption is satisfied too; so we can also conclude the desired result. ∎

It often happens that we want to define a function f on a quotient set X/R by using a function $F : X \to Y$. The following theorem provides a tool for doing this.

Theorem 3.21. *Suppose R, S are equivalence relations on the sets X, Y respectively. Also suppose that $F : X \to Y$ is a function, and for every $a, b \in X$ we have*

$$aRb \implies F(a)SF(b).$$

Then there is a unique function $f : X/R \to Y/S$ that satisfies

$$f([a]_R) = [F(a)]_S$$

for every $a \in X$.

Remark. In the following proof we define $f([a]_R) := [F(a)]_S$, and we will show that f is a function. In this process we have to check that f maps each equivalence class $[a]_R$ to exactly one equivalence class in Y/S. However we know that if $b \in [a]_R$ then $[b]_R = [a]_R$. So if in the definition of f we use the representative b for the equivalence class $[a]_R$, we obtain $[F(b)]_S$. Therefore we have to make sure that the two equivalence classes $[F(a)]_S$ and $[F(b)]_S$ are the same (which follows easily from our assumption about F). When we define a function in this way, and the required conditions are satisfied, we say that the function is **well defined**.

Remark. A particular case of this theorem is when S is the equality relation on Y; then $F(a)SF(b)$ simply means that $F(a) = F(b)$, and we have $f([a]_R) = \{F(a)\}$. We can easily modify the following proof and show that $f([a]_R) := F(a)$ defines a well-defined function $f : X/R \to Y$.

Proof. First note that by Theorem 3.8 there is at most one function whose domain is X/R, and maps $[a]_R$ to $[F(a)]_S$. Now let us define $f([a]_R) := [F(a)]_S$. More explicitly, this means that

$$f := \{(x, y) \in X/R \times Y/S : \exists c \in x \text{ such that } F(c) \in y\}.$$

Note that the domain of f is all of X/R, since to every $[a]_R \in X/R$ we can at least assign one value $[F(a)]_S$. Now let us show that this value is the only value that is assigned to $[a]_R$. Suppose $([a]_R, y) \in f$. Then by definition there is $c \in [a]_R$ such that $F(c) \in y$. But $c \in [a]_R$ means that aRc. Hence by our assumption we have $F(a)SF(c)$. Thus $F(c) \in [F(a)]_S$. Therefore the two equivalence classes $y, [F(a)]_S$ have a nonempty intersection; so they must be equal, i.e. $y = [F(a)]_S$. Thus the value of f at $[a]_R$ is uniquely determined, as desired. Hence f is a function. ■

3.4 Order Relations

In this section, we study another important type of relations, namely order relations.

Definition 3.17. Let $R \subset X \times X$ be a relation on a set X. We say that R is a **(non-strict) partial order** if it is reflexive, antisymmetric, and transitive; i.e. if for every $x, y, z \in X$ we have

(i) xRx.
(ii) If xRy and yRx then $x = y$.
(iii) If xRy and yRz then xRz.

***Example* 3.9.** Inclusion is a partial order on the power set of any set, i.e. if for $x, y \in \mathcal{P}(X)$ we define xRy to mean $x \subset y$, then R is a partial order.

***Example* 3.10.** Equality is the only equivalence relation which is also a partial order. Because if a relation R is both symmetric and antisymmetric, then xRy implies that yRx; and therefore we must have $x = y$. On the other hand, for every x we have xRx due to reflexivity. Hence xRy if and only if $x = y$.

Definition 3.18. Let $S \subset X \times X$ be a relation on a set X. We say that S is a **strict partial order** if it is irreflexive and transitive; i.e. if for every $x, y, z \in X$ we have

(i) $x \not S x$.

(ii) If xSy and ySz then xSz.

Theorem 3.22. *Suppose S is a strict partial order on a set X. If for $x, y \in X$ we have xSy, then $y \not S x$.*

Proof. Suppose xSy. Also, suppose to the contrary that ySx. Then by transitivity of S we get xSx, which contradicts the irreflexivity of S. Hence we cannot have ySx; so $y \not S x$. ■

Theorem 3.23. *Let $R, S \subset X \times X$ be two relations on a set X. Suppose we have*

$$S = R - \{(x, x) : x \in X\}, \qquad \text{and} \qquad R = S \cup \{(x, x) : x \in X\}.$$

Then R is a non-strict partial order if and only if S is a strict partial order.

Remark. The relationship between R, S can also be expressed as follows

$$xRy \qquad \Longleftrightarrow \qquad xSy \ \text{ or } \ x = y.$$

Proof. Suppose S is a strict partial order. Now R is a reflexive relation, because for every $x \in X$ we have $(x, x) \in R$, i.e. xRx. Let us show that R is antisymmetric. Suppose xRy and yRx. If $x \neq y$ then we must have xSy and ySx. However, this contradicts the last theorem, since S is a strict partial order. Therefore we must have $x = y$, as desired. Finally, let us show that R is transitive. Suppose xRy and yRz. If $x = y$, or $y = z$, then we trivially have xRz. So suppose $x \neq y$ and $y \neq z$. Then we must have xSy and ySz. Then we get xSz, since S is transitive. Hence we also have xRz, as desired.

Conversely, suppose R is a non-strict partial order. Note that S is irreflexive, since for every $x \in X$ we have $(x, x) \notin S$, i.e. $x \not S x$. To show that S is transitive, suppose that xSy and ySz. Then we also have xRy and yRz. Hence we get xRz, since R is transitive. Now note that we cannot have $x = z$, since in this case we have xRy and yRx, and hence by antisymmetry of R we get $x = y$, which contradicts the irreflexivity of S. Therefore $x \neq z$. Thus xRz implies that xSz, as desired. ■

We usually denote partial orders by $\leq, \preceq, \sqsubseteq, ...$, and their corresponding strict partial orders by $<, \prec, \sqsubset, ...$ Thus in this notation, $x \preceq y$ if and only if $x \prec y$ or $x = y$. Although, an exception to this convention is the non-strict partial order $\subset$. We will also write $y \succeq x$ to mean that $x \preceq y$, and $y \succ x$ to mean that $x \prec y$. Similar inverted notation can be used for other symbols too. Also, for several elements like x, y, z, u, v, we may write $x, y, z, u \prec v$ to denote the fact that $x \prec v$, and $y \prec v$, and $z \prec v$, and $u \prec v$.

If for some x, y we have $x \preceq y$, then we say that x is **less** than y, or y is **greater** than x. We may also say that x is **smaller** than y, or y is **larger** than x. And if we have $x \prec y$, then we may say that x is **strictly less** than y, or y is **strictly greater** than x. We may also say that x is **strictly smaller** than y, or y is **strictly larger** than x. Also, we say x, y are **comparable** if at least one of the relations $x \preceq y$ or $y \preceq x$ holds.

Theorem 3.24. *Suppose $\preceq$ is a partial order on a set X. Let $x, y, z \in X$. If $x \preceq y$ and $y \prec z$ then $x \prec z$. Also, if $x \prec y$ and $y \preceq z$ then $x \prec z$.*

Proof. We prove the first claim; the other one can be proved similarly. Suppose $x \preceq y$ and $y \prec z$. If $x \prec y$ then by transitivity of $\prec$ we have $x \prec z$. And if $x = y$ then we trivially have $x \prec z$, since we know that $y \prec z$. In other words, in set theoretic notation we have $(y, z) \in \prec$ and $(x, z) = (y, z)$, hence we also have $(x, z) \in \prec$. ■

Definition 3.19. Suppose $\preceq$ is a partial order on a set X. Let $z \in X$, and $A \subset X$. Then

(i) z is a **maximum** or **greatest** (or **largest**) element of A if $z \in A$, and for every $a \in A$ we have $a \preceq z$.

(ii) z is a **minimum** or **least** (or **smallest**) element of A if $z \in A$, and for every $a \in A$ we have $a \succeq z$.

(iii) z is a **maximal** element of A if $z \in A$, and there is no $a \in A$ such that $a \succ z$.

(iv) z is a **minimal** element of A if $z \in A$, and there is no $a \in A$ such that $a \prec z$.

(v) z is an **upper bound** for A if for every $a \in A$ we have $a \preceq z$. If A has an upper bound, then we say that A is **bounded above**.

(vi) z is a **lower bound** for A if for every $z \in A$ we have $a \succeq z$. If A has a lower bound, then we say that A is **bounded below**.

(vii) A is **bounded** if it is bounded above and bounded below. And A is called **unbounded** if it is not bounded.

(viii) z is a **least upper bound** or **supremum** of A if z is an upper bound for A, and for any $w \in X$ which is an upper bound for A we have $w \succeq z$.

(ix) z is a **greatest lower bound** or **infimum** of A if z is a lower bound for A, and for any $w \in X$ which is a lower bound for A we have $w \preceq z$.

Remark. Note that the maximum, minimum, maximal, and minimal elements of A must belong to A, if they exist; while upper and lower bounds, supremum, and infimum of A need not belong to A.

Also note that a maximal element is not necessarily a maximum. Because if no element of A is greater than b, then it does not follow that b is greater than every element of A; since there might be some $a \in A$ which is not comparable to b, i.e. $a \not\preceq b$ and $b \not\preceq a$. Similarly, a minimal element is not necessarily a minimum.

Theorem 3.25. *Suppose $\preceq$ is a partial order on a set X. Let $A \subset X$. Then the maximum, minimum, supremum, and infimum of A are unique, if they exist.*

Proof. Suppose $z, x \in A$ are both maximums of A. Then by definition we must have $x \preceq z$, since z is a maximum, and $z \preceq x$, since x is a maximum. Therefore, by antisymmetry of $\preceq$ we get $z = x$. The uniqueness of minimum can be proved similarly.

Now suppose $z, x \in X$ are both infimums of A. Then by definition we have $x \preceq z$, since z is an infimum, and x is a lower bound for A. We also have $z \preceq x$, since x is an infimum, and z is a lower bound for A. Therefore, by antisymmetry of $\preceq$ we get $z = x$. The uniqueness of supremum can be proved similarly. ■

Remark. Note that the maximal or minimal elements of a set are not necessarily unique; because a set can have several maximal or minimal elements which are not comparable.

Definition 3.20. Suppose $\preceq$ is a partial order on a set X. Then we say that $\preceq$ is a **total order** or **linear order** if it satisfies the **trichotomy law**, i.e. if for every $x, y \in X$ exactly one of the following three conditions holds:

$$x \prec y, \qquad x = y, \qquad x \succ y.$$

Remark. Note that the fact that at most one of the above three conditions can hold is a consequence of the fact that $\preceq$ is a partial order. Thus a partial order $\preceq$ is a total order if for every x, y at least one of the above three conditions holds. This fact is also manifested in the following theorem.

Theorem 3.26. *Suppose $\preceq$ is a partial order on a set X. Then $\preceq$ is a total order if and only if every two elements of X are comparable.*

Proof. If $\preceq$ is a total order, then for every x, y one of the conditions $x \prec y, x = y, x \succ y$ holds. Hence in either case we must have $x \preceq y$ or $y \preceq x$. Thus x, y are comparable.

Conversely, suppose for every x, y we have $x \preceq y$ or $y \preceq x$. If $x = y$ then we cannot have $x \prec y$ or $y \prec x$, due to the irreflexivity of $\prec$. Now suppose $x \neq y$. Then we must have $x \prec y$ or $y \prec x$. However, both of these conditions cannot hold

simultaneously; since otherwise we would get $x \prec x$ by transitivity of $\prec$, and this contradicts the irreflexivity of $\prec$. Therefore, in any case, exactly one of the three conditions $x \prec y, x = y, x \succ y$ holds. Thus $\preceq$ is a total order. ■

Definition 3.21. Suppose $\preceq, \sqsubseteq$ are partial orders on the sets X, Y respectively. Let $f : X \to Y$ be a function, and let $x_1, x_2 \in X$. Then we say that

(i) f is an **increasing** function if $x_1 \preceq x_2$ implies that $f(x_1) \sqsubseteq f(x_2)$.
(ii) f is a **decreasing** function if $x_1 \preceq x_2$ implies that $f(x_1) \sqsupseteq f(x_2)$.
(iii) f is a **strictly increasing** function if $x_1 \prec x_2$ implies that $f(x_1) \sqsubset f(x_2)$.
(iv) f is a **strictly decreasing** function if $x_1 \prec x_2$ implies that $f(x_1) \sqsupset f(x_2)$.
(v) f is **monotone** if it is increasing or decreasing, and f is **strictly monotone** if it is strictly increasing or strictly decreasing.

Theorem 3.27. *Suppose $\preceq, \sqsubseteq$ are total orders on the sets X, Y respectively. Let $f : X \to Y$ be a function. Then we have*

(i) *If f is strictly monotone then it is one-to-one.*
(ii) *If f is monotone and one-to-one then it is strictly monotone.*
(iii) *If f is strictly increasing and invertible then f^{-1} is also strictly increasing.*
(iv) *If f is strictly decreasing and invertible then f^{-1} is also strictly decreasing.*

Proof. **(i)** Suppose to the contrary that $f(x_1) = f(x_2)$ and $x_1 \neq x_2$. Then either $x_1 \prec x_2$ or $x_1 \succ x_2$, because $\preceq$ is a total order. But then we get $f(x_1) \sqsubset f(x_2)$, or $f(x_1) \sqsupset f(x_2)$, since f is strictly monotone. However, this contradicts our initial assumption that $f(x_1) = f(x_2)$. Hence we must have $x_1 = x_2$. Thus f is one-to-one.

(ii) Note that if $x_1 \prec x_2$ then we must have $f(x_1) \neq f(x_2)$, since $x_1 \neq x_2$ and f is one-to-one. Hence if f is monotone, it must be strictly monotone.

(iii) Note that $f^{-1} : Y \to X$. Suppose $y_1, y_2 \in Y$, and $y_1 \sqsubset y_2$. We have to show that $f^{-1}(y_1) \prec f^{-1}(y_2)$. Suppose to the contrary that $f^{-1}(y_1) \not\prec f^{-1}(y_2)$. Then we must have $f^{-1}(y_1) = f^{-1}(y_2)$, or $f^{-1}(y_1) \succ f^{-1}(y_2)$; because $\preceq$ is a total order. In the first case we have

$$y_1 = f(f^{-1}(y_1)) = f(f^{-1}(y_2)) = y_2,$$

which contradicts our assumption. And in the second case, by using the fact that f is strictly increasing, we get

$$y_1 = f(f^{-1}(y_1)) \sqsupset f(f^{-1}(y_2)) = y_2,$$

which also contradicts our assumption. Hence we cannot have $f^{-1}(y_1) \not\prec f^{-1}(y_2)$. Thus $f^{-1}(y_1) \prec f^{-1}(y_2)$, and f^{-1} is strictly increasing.

(iv) The proof is similar to the previous part. ■

Definition 3.22. Suppose $\preceq$ is a total order on a set X. Then we say that $\preceq$ is a **well-ordering** if every nonempty subset of X has a least element.

Chapter 4

Natural Numbers and Finite Sets

4.1 Natural Numbers

Natural numbers are probably the first mathematical object that people get familiar with, and learn how to use them for counting. But, what is a natural number? For example, what is "two"? It turns out that from a mathematical point of view, the inherent nature of "two" is not very important. The important thing about "two" is that it represents *twoness*. In other words, we need a concept of "two" that represents what two apples, two oranges, and two chairs have in common, i.e. their quantity. One way to do this is to define "two" to be some set which has exactly two elements. Then we can say that a collection of objects has two elements if it has the same number of elements as "two". Note that in this process, we actually assigned a representative to the concept of "two"; and we do not claim that this representative is the *true* representative of the *real* "two", if there is any such thing.

The following definition, due to von Neumann, implements the above idea in an elegant way.

$$
\begin{aligned}
&0 := \emptyset = \{\,\},\\
&1 := 0 \cup \{0\} = \{0\} = \{\emptyset\},\\
&2 := 1 \cup \{1\} = \{0, 1\},\\
&3 := 2 \cup \{2\} = \{0, 1, 2\},\\
&4 := 3 \cup \{3\} = \{0, 1, 2, 3\},\\
&\vdots
\end{aligned}
$$

Thus we define each natural number n to be the set of all natural numbers smaller than n. We can continue this way and build every natural number, however, we cannot show that the collection of all natural numbers is a set. In other words, by using the axioms we have expressed so far we can construct infinitely many sets,

i.e. all the natural numbers, but we cannot show that a set with infinitely many elements exists. Therefore in order to show that a set containing all the natural numbers exists we need a new axiom.

Axiom of Infinity.

$$\vdash \exists I\big((\emptyset \in I) \wedge (\forall a \in I\ a \cup \{a\} \in I)\big).$$

Remark. Remember that $\emptyset \in I$ is a shorthand notation for

$$\exists y\big((y \in I) \wedge \forall z(z \notin y)\big).$$

Also, $\forall a \in I\ a \cup \{a\} \in I$ is a shorthand notation for

$$\forall a\big[a \in I \to \exists y\big((y \in I) \wedge \forall z[z \in y \leftrightarrow (z \in a \vee z = a)]\big)\big].$$

The axiom of infinity says that there is a set I which contains $\emptyset$, and for every $a \in I$ we have $a \cup \{a\} \in I$. It follows that I contains $\emptyset \cup \{\emptyset\} = 1$, and $1 \cup \{1\} = 2$, and $2 \cup \{2\} = 3$, and so on. Thus, intuitively, we can see that I contains every natural number; so it has infinitely many elements. Although it might also have elements which are not natural numbers. Notice that the axiom of infinity states that an infinite set exists, without getting into the details of what an infinite set is. Before going further, let us introduce a few notions.

Definition 4.1. Let a be a set. The **successor** of a is

$$a^+ := a \cup \{a\}.$$

Remark. The successor of any set exists due to the axioms of pairing and union. Also note that for every set a we have $a \in a^+$, and $a \subset a^+$.

With the above notation we have

$$0 = \emptyset, \qquad 1 = 0^+, \qquad 2 = 1^+, \qquad 3 = 2^+, \qquad \dots$$

Definition 4.2. We say a set I is **inductive** if $\emptyset \in I$, and for every $a \in I$ we have $a^+ \in I$.

So, the axiom of infinity states that an inductive set exists. Intuitively, an inductive set contains every natural number. However, it can have other elements too. Our next step is to show that an inductive set exists which does not have any element besides the natural numbers. But we have not yet defined what a natural number is. Of course we have defined $0, 1, 2, 3, \dots$, but we did not say when a given set is a natural number. For a precise definition of natural numbers, we rely on our intuition about inductive sets.

Definition 4.3. A **natural number** is a set which belongs to every inductive set.

For example $0 = \emptyset$ is a natural number, because every inductive set contains $\emptyset$ by definition.

Theorem 4.1. *If n is a natural number then n^+ is also a natural number.*

Remark. As a consequence, $1 = 0^+$ is a natural number, $2 = 1^+$ is a natural number, and so on.

Proof. Suppose n is a natural number. Let I be an inductive set. Then by definition we have $n \in I$. Hence we must have $n^+ \in I$, since I is inductive. Therefore n^+ belongs to every inductive set; because I is an arbitrary inductive set. Thus n^+ is also a natural number, as desired. ■

Remark. Let us also present a more formal proof for the above theorem. Let $\text{ind}(I)$ be the formula which says that I is an inductive set. We have

1	$\vdash \forall I(\text{ind}(I) \to n \in I)$;	(since n is a natural number)
2	$\vdash \text{ind}(I) \to n \in I$;	(by E$\forall$)
3	$\text{ind}(I) \vdash n \in I$;	(by E$\to$)
4	$n \in I, \text{ind}(I) \vdash n^+ \in I$;	(by definition of inductive sets)
5	$\text{ind}(I) \vdash n^+ \in I$;	(by cut rule applied to lines 3,4)
6	$\vdash \text{ind}(I) \to n^+ \in I$;	(by I$\to$)
7	$\vdash \forall I(\text{ind}(I) \to n^+ \in I)$.	(by I$\forall$)

The last formula says that n^+ is a natural number, as desired.

Note that the sentences "Let I be an inductive set. Then by definition we have $n \in I$." in the previous proof correspond to the application of E$\forall$ in this proof. Also note that the sentence "Therefore n^+ belongs to every inductive set; because I is an arbitrary inductive set." in the previous proof corresponds to the application of I$\forall$ in this proof. In informal proofs, these are common ways to state that we are employing E$\forall$ or I$\forall$, respectively.

Theorem 4.2. *There exists a unique set whose elements are exactly the natural numbers. In other words*

$$\vdash \exists! A \forall x \big(x \in A \leftrightarrow \forall I(\text{ind}(I) \to x \in I)\big),$$

where $\text{ind}(I)$ *is the formula* $(\emptyset \in I) \wedge (\forall a \in I\ a \cup \{a\} \in I)$*, which states that I is an inductive set.*

Remark. Informally, the above formula says that $x \in A$ if and only if x belongs to every inductive set. Hence $x \in A$ if and only if x is a natural number. We can also say that the set A is the intersection of all inductive sets.

Proof. The uniqueness of A follows from Theorem 2.2, since A is defined by a formula, namely $\forall I(\mathrm{ind}(I) \to x \in I)$. Thus we only need to prove the existence of A. The idea is to consider an inductive set J, which we know exists due to the axiom of infinity, and then using the axiom of separation, define A as the subset of J that consists of only natural numbers. Since we know that by definition all natural numbers belong to the inductive set J, the set A constructed in this way will contain all the natural numbers. Let us implement this idea rigorously.

To simplify the notation, we denote $\forall I(\mathrm{ind}(I) \to x \in I)$ by $\mathrm{nat}(x)$. Note that $\mathrm{nat}(x)$ says that x belongs to every inductive set, i.e. x is a natural number. By the axiom of infinity we know that

$$\vdash \exists J\, \mathrm{ind}(J). \tag{$*$}$$

On the other hand, by the axiom of separation we have

$$\vdash \exists A \forall x\big(x \in A \leftrightarrow (x \in J) \wedge \mathrm{nat}(x)\big). \tag{$**$}$$

Now by E$\forall$ we have $\mathrm{nat}(x) \vdash \mathrm{ind}(J) \to x \in J$. Hence by E$\to$ and the cut rule we get

$$\mathrm{nat}(x),\, \mathrm{ind}(J) \vdash x \in J.$$

Thus by I$\wedge$ we have $\mathrm{nat}(x),\, \mathrm{ind}(J) \vdash (x \in J) \wedge \mathrm{nat}(x)$. Therefore by E$\leftrightarrow$ we get

$$\mathrm{nat}(x),\, \mathrm{ind}(J),\, x \in A \leftrightarrow (x \in J) \wedge \mathrm{nat}(x) \vdash x \in A.$$

In addition, by E$\leftrightarrow$ and E$\wedge$ we have

$$x \in A,\, \mathrm{ind}(J),\, x \in A \leftrightarrow (x \in J) \wedge \mathrm{nat}(x) \vdash (x \in J) \wedge \mathrm{nat}(x) \vdash \mathrm{nat}(x).$$

Hence by I$\leftrightarrow$ we obtain

$$\mathrm{ind}(J),\, x \in A \leftrightarrow (x \in J) \wedge \mathrm{nat}(x) \vdash x \in A \leftrightarrow \mathrm{nat}(x).$$

So by applying E$\forall$, and then I$\forall$, we get

$$\mathrm{ind}(J),\, \forall x\big(x \in A \leftrightarrow (x \in J) \wedge \mathrm{nat}(x)\big) \vdash \forall x\big(x \in A \leftrightarrow \mathrm{nat}(x)\big).$$

Now by applying I$\exists$, and then E$\exists$, we get

$$\mathrm{ind}(J),\, \exists A \forall x\big(x \in A \leftrightarrow (x \in J) \wedge \mathrm{nat}(x)\big) \vdash \exists A \forall x\big(x \in A \leftrightarrow \mathrm{nat}(x)\big).$$

Thus by $(**)$ we obtain $\mathrm{ind}(J) \vdash \exists A \forall x\big(x \in A \leftrightarrow \mathrm{nat}(x)\big)$. Hence by E$\exists$ we get

$$\exists J\, \mathrm{ind}(J) \vdash \exists A \forall x\big(x \in A \leftrightarrow \mathrm{nat}(x)\big).$$

Therefore by $(*)$ we get $\vdash \exists A \forall x\big(x \in A \leftrightarrow \mathrm{nat}(x)\big)$, as desired. ■

Notation. The unique set A given by the above theorem is denoted by ω, and is called the **set of natural numbers**. Thus we have

$$\omega = \{0, 1, 2, 3, \dots\}.$$

We also set $\mathbb{N} := \{n \in \omega : n \neq 0\}$. Hence we have

$$\mathbb{N} = \{1, 2, 3, \dots\}.$$

The set $\mathbb{N}$ is also referred to as the set of natural numbers. However, in these notes, we consider 0 to be a natural number.

Theorem 4.3. *ω is an inductive set, and it is a subset of any inductive set.*

Remark. In other words, ω is the smallest inductive set with respect to inclusion.

Proof. Remember that the elements of ω are exactly those sets which belong to every inductive set. First note that $\emptyset \in \omega$, since $\emptyset$ belongs to every inductive set. Now let $a \in \omega$. Then a is a natural number. Hence as we have shown before, a^+ is also a natural number. Thus $a^+ \in \omega$. Therefore ω is inductive.

Finally, let I be an inductive set. Let $a \in \omega$. Then we have $a \in I$, since a belongs to every inductive set. Therefore $\omega \subset I$. So, ω is a subset of every inductive set, because I was an arbitrary inductive set. ■

Principle of Mathematical Induction. *Every inductive subset of ω is equal to ω. In other words, if $A \subset \omega$ satisfies*

(i) *$0 \in A$, and*

(ii) *for every $n \in \omega$, $n \in A$ implies that $n^+ \in A$,*

then $A = \omega$.

Proof. Note that A is an inductive set. Because $\emptyset \in A$. And if $a \in A$ then $a \in \omega$, since $A \subset \omega$; so by the above assumption we get $a^+ \in A$. Therefore by the previous theorem we must have $\omega \subset A$. Hence $A = \omega$, as desired. ■

Let $\phi(n)$ be a formula that has n as a free variable. Suppose we want to show that $\phi(n)$ is true for every $n \in \omega$. Let

$$A := \{n \in \omega : \phi(n)\}.$$

Note that by definition, $n \in A$ is equivalent to $n \in \omega \wedge \phi(n)$. Hence for $n \in \omega$ we have $n \in A$ if and only if $\vdash \phi(n)$. Thus our goal is to show that $A = \omega$. To this end, we can use the above theorem, and show that

(i) $0 \in A$, i.e. $\vdash \phi(0)$, and

(ii) if $n \in A$ then $n^+ \in A$, i.e. $\vdash \phi(n) \to \phi(n^+)$.

Therefore we conclude that $\vdash \phi(n)$ for every $n \in \omega$. This method is called **proof by induction**. The first part, i.e. showing that $\phi(0)$ holds, is called the **base case**, or the **basis**. The second part, i.e. showing that if $\phi(n)$ holds then $\phi(n^+)$ holds too, is called **induction step**. In the induction step we assume that $\phi(n)$ is true, and then we have to show that $\phi(n^+)$ is true. The assumption that $\phi(n)$ holds is called the **induction hypothesis**.

Intuitively, proof by induction works as follows. We know that $\vdash \phi(0)$. If we set $n = 0$ in the induction step, we have $\vdash \phi(0) \to \phi(1)$. Thus by E$\to$ we get $\vdash \phi(1)$. Now if we set $n = 1$ in the induction step, we have $\vdash \phi(1) \to \phi(2)$. Thus by E$\to$ we get $\vdash \phi(2)$. If we repeat this argument we obtain $\vdash \phi(3)$, and then $\vdash \phi(4)$, and so forth.

Remark. Note that $\phi(0)$ can be formally expressed as

$$\exists y\big(\phi(y) \wedge \forall z(z \notin y)\big).$$

Also, $\phi(n) \to \phi(n^+)$ can be formally expressed as

$$\forall a \forall b\big[\mathrm{nat}(a) \wedge (b = a \cup \{a\}) \to \big(\phi(a) \to \phi(b)\big)\big],$$

where $\mathrm{nat}(a)$ expresses that a belongs to every inductive set, i.e. it is a natural number.

Remark. Not every subset of ω is of the form $\{n \in \omega : \phi(n)\}$ for some formula ϕ. Because ω has more subsets than there are formulas, as we will explain when we discuss the notion of *cardinality* of sets. Therefore the principle of mathematical induction is actually stronger than its special case discussed above.

Theorem 4.4. *Let $n \in \omega$. Then we have*
(i) $n = \bigcup n^+$.
(ii) *If $n \neq 0$ then there is a unique $m \in \omega$ such that $n = m^+$.*

Remark. When $n = m^+$, i.e. when n is the successor of m, we say m is the *predecessor* of n.

Proof. **(i)** The proof is by induction on n. First note that for every set A we can easily show that $\bigcup\{A\} = A$. Now for $n = 0$ we have

$$0 = \emptyset = \bigcup\{\emptyset\} = \bigcup 1 = \bigcup 0^+.$$

Next suppose for some n we have $n = \bigcup n^+$. We know that $n^{++} = n^+ \cup \{n^+\}$. Therefore by Exercise 2.2 we get

$$\bigcup n^{++} = \bigcup n^+ \cup \bigcup\{n^+\} = n \cup n^+ = n^+,$$

since $n \subset n^+$.

(ii) The proof is again by induction on n. Let

$$A := \{n \in \omega : n = 0 \text{ or } n = m^+ \text{ for a unique } m \in \omega\}.$$

It is obvious that $0 \in A$. Now let $n \in A$. We have to show that $n^+ = m^+$ for some $m \in \omega$. But if we take $m = n$ then we have $n^+ = m^+$, as desired. (Note that here we did not need to use the induction hypothesis.) In addition, note that if we have $n^+ = m^+$ for some $m \in \omega$, then by part (i) we get

$$m = \bigcup m^+ = \bigcup n^+ = n.$$

Therefore $n^+ = m^+$ for a unique $m \in \omega$; so $n^+ \in A$ by definition. Thus $A = \omega$. Hence we have shown that every nonzero natural number is the successor of a unique natural number, as desired. ■

Definition 4.4. A **Peano system** is a set N with a distinguished element $e \in N$, equipped with a function $s : N \to N$, that satisfies the following conditions:

(i) $s(a) \neq e$ for all $a \in N$.
(ii) s is one-to-one.
(iii) If $M \subset N$ satisfies
  (a) $e \in M$, and
  (b) $a \in M$ implies $s(a) \in M$,
  then $M = N$.

Remark. The function s is called the **successor function**, and the above conditions are called **Peano axioms**. These axioms represent the basic properties of natural numbers, and are sufficient to imply all their elementary properties. The next theorem shows that the set of natural numbers that we constructed by set theoretic tools satisfies the Peano axioms.

Theorem 4.5. *The set ω, equipped with $0 \in \omega$ and the successor function*

$$\begin{array}{c} \omega \to \omega \\ n \mapsto n^+ \end{array},$$

is a Peano system.

Proof. First note that $n \mapsto n^+$ is a function, since $n^+ = n \cup \{n\}$ is uniquely determined by n. (To be more precise, note that as we proved in Section 2.2, the singleton $\{n\}$ is uniquely determined by its element n, and the union of the two set $n, \{n\}$ is uniquely determined by them.) In addition, the function $n \mapsto n^+$ is one-to-one; because by the previous theorem we have

$$m^+ = n^+ \implies m = \bigcup m^+ = \bigcup n^+ = n.$$

Furthermore, the third condition of a Peano system is satisfied due to the principle of mathematical induction. Thus we only need to show that for every $n \in \omega$ we have

$$n^+ \neq 0.$$

Now note that for every n we have $n \in n^+$, while $0 = \emptyset$ is empty. So $n^+ \neq 0$ as desired. ■

4.2 Arithmetic

Sometimes, when we define a function f, we need to use some of the values of f in order to assign the value of f at another point of its domain. For example, the power n^k is usually defined as $n \cdot n^{k-1}$, i.e. the value of this exponential function at k is defined in terms of its value at $k-1$. The next theorem provides the tool for defining functions in this way.

Recursion Theorem. *Let A be a set, and let $a \in A$. Suppose $F : \omega \times A \to A$ is a function. Then there exists a unique function $f : \omega \to A$ such that*

(i) $f(0) = a$, *and*
(ii) $f(n^+) = F(n, f(n))$ *for every $n \in \omega$.*

Proof. Consider a function $h : B \to A$, where $B \subset \omega$. In this proof, we say h is *admissible* if the following two conditions hold:

(i) If $0 \in B$ then $h(0) = a$.
(ii) If $n^+ \in B$ then $n \in B$, and we have $h(n^+) = F(n, h(n))$.

Let

$$X := \{h : h \text{ is an admissible function from some } B \subset \omega \text{ into } A\}.$$

Note that $h \subset B \times A \subset \omega \times A$; so $X \subset \mathcal{P}(\omega \times A)$ and its existence is guaranteed by the subset axiom. Now let $f := \bigcup X$, i.e. f is the union of all admissible functions. We claim that f is a function from ω to A that satisfies the requirements of the theorem, i.e. f is itself an admissible function. First note that the elements of X are functions; so the elements of $\bigcup X = f$ are ordered pairs. For example, $\{(0, a)\}$ is an admissible function from $\{0\}$ into A; hence $\{(0, a)\} \in X$. Thus we have $(0, a) \in f$.

Next consider the set

$$S := \{n \in \omega : \text{there is a unique } x \in A \text{ such that } (n, x) \in f\}.$$

Let us show by induction that $S = \omega$. Note that $0 \in S$; because as we have seen $(0, a) \in f$. And if we have $(0, x) \in h \subset f$ for some admissible function h, then by definition we must have $x = a$. Now suppose $n \in S$. Then there is a unique $x_n \in A$ such that $(n, x_n) \in f$. Hence there is an admissible function h such that

$(n, x_n) \in h \subset f$. If n^+ is in the domain of h we have $(n^+, F(n, x_n)) \in h$, since h is admissible. And if n^+ does not belong to the domain of h, let

$$\hat{h} := h \cup \{(n^+, F(n, x_n))\}.$$

Then it is easy to see that $\hat{h}$ is also an admissible function. It is obvious that $\hat{h}$ is a function, since n^+ does not belong to the domain of h. Also, if 0 is in the domain of $\hat{h}$ we have $\hat{h}(0) = h(0) = a$, since $n^+ \neq 0$. Now suppose k^+ is in the domain of $\hat{h}$. If $k^+ \neq n^+$ then k^+ is in the domain of h, and the required condition holds, since h is admissible. And if $k^+ = n^+$ then $k = n$. So k belongs to the domain of h, which is included in the domain of $\hat{h}$. Thus we have

$$\hat{h}(k^+) = \hat{h}(n^+) = F(n, x_n) = F(n, h(n)) = F(k, \hat{h}(k)).$$

Hence $\hat{h}$ is an admissible function, and we have $(n^+, F(n, x_n)) \in \hat{h}$. Therefore in either case we have $(n^+, F(n, x_n)) \in f$.

On the other hand, if there is an admissible function $\tilde{h}$ such that $(n^+, \tilde{h}(n^+)) \in \tilde{h} \subset f$, then n is also in the domain of $\tilde{h}$, and we must have $\tilde{h}(n^+) = F(n, \tilde{h}(n))$. However, $(n, \tilde{h}(n)) \in \tilde{h} \subset f$; so by induction hypothesis we get $\tilde{h}(n) = x_n$. Therefore $\tilde{h}(n^+) = F(n, x_n)$. Thus $x = F(n, x_n)$ is the unique element of A that satisfies $(n^+, x) \in f$, as desired. Hence $S = \omega$. But this means that f is a function from ω to A. In addition, we have shown that for every n we have

$$(n, x_n) \in f \implies (n^+, F(n, x_n)) \in f.$$

In other words, $f(n) = x_n$ implies that $f(n^+) = F(n, x_n) = F(n, f(n))$. Thus f is an admissible function.

Finally, let us show that f is unique. Suppose $g : \omega \to A$ is also an admissible function. Let us show that $f = g$. Let

$$\tilde{S} := \{n \in \omega : f(n) = g(n)\}.$$

Then we have $0 \in \tilde{S}$, since $f(0) = a = g(0)$. Now suppose $n \in \tilde{S}$. The $f(n) = g(n)$. Hence we have

$$f(n^+) = F(n, f(n)) = F(n, g(n)) = g(n^+),$$

because F is a function. Thus we get $n^+ \in \tilde{S}$. Therefore $\tilde{S} = \omega$, and we have $f = g$ as desired. ■

Remark. A definition of a function f by using the recursion theorem is called a **recursive definition**, or an **inductive definition**. Sometimes, we are only interested in the objects $f(n)$, rather than the function f. In other words, we would like to construct the object $f(n^+)$ by using the object $f(n)$. The recursion theorem provides a sufficient tool for these kinds of constructions too.

Remark. Intuitively, a recursive definition works as follows. We define $f(0) := a$. Then we can define

$$f(1) = f(0^+) := F(0, a) = F(0, f(0)).$$

Then we can define $f(2) := F(1, F(0, a)) = F(1, f(1))$. If we continue in this way, we can define $f(3)$, $f(4)$, and so on. Of course, we have to show that we can continue this process indefinitely, and we obtain a unique function through it. This has been shown in the recursion theorem.

Remark. A special important case of the recursion theorem is when F does not depend on n, i.e. when there exists a function $G : A \to A$ such that $F = G \circ \pi_2$, where $\pi_2 : \omega \times A \to A$ is the projection on the second component. In other words, when for every $n \in \omega$ and $x \in A$ we have $F(n, x) = G(x)$. Then there is a unique function $f : \omega \to A$ such that

(i) $f(0) = a$, and
(ii) $f(n^+) = G(f(n))$ for every $n \in \omega$.

Next we will show that every Peano system is essentially the same as the Peano system of ω. Intuitively, this means that any Peano system can be obtained from ω by renaming its elements.

Theorem 4.6. *Suppose N, equipped with $e \in N$ and the successor function $s : N \to N$, is a Peano system. Then N is **isomorphic** to the Peano system of ω, i.e. there exists a one-to-one and onto function*

$$h : \omega \to N$$

such that $h(0) = e$, and $h(n^+) = s(h(n))$ for every $n \in \omega$.

Proof. By the special case of the recursion theorem discussed above there is a function $h : \omega \to N$ such that $h(0) = e$, and $h(n^+) = s(h(n))$ for every n. We only need to show that h is one-to-one and onto. Let

$$S := \{n \in \omega : h(k) = h(n) \text{ implies } k = n \text{ for every } k \in \omega\}.$$

First let us show that $0 \in S$. Suppose to the contrary that $k \neq 0$, and $h(k) = h(0) = e$. Then we know that $k = m^+$ for some $m \in \omega$. Hence we get

$$e = h(k) = h(m^+) = s(h(m)),$$

which contradicts the fact that in a Peano system e is not in the image of s. Thus $0 \in S$. Now suppose $n \in S$. Suppose for some k we have $h(k) = h(n^+)$. We know that $h(n^+) \neq e$, since $n^+ \neq 0$. Thus $k \neq 0$ too. So we have $k = m^+$ for some m. Hence we get

$$s(h(m)) = h(m^+) = h(k) = h(n^+) = s(h(n)).$$

Therefore $h(m) = h(n)$, because s is one-to-one. Thus by the induction hypothesis we obtain $m = n$. So $k = m^+ = n^+$. Hence $n^+ \in S$. Thus $S = \omega$, and therefore h is one-to-one.

Next let us show that h is onto. Consider the set

$$M := \{a \in N : a = h(n) \text{ for some } n \in \omega\}.$$

Then $e \in M$ since $e = h(0)$. Now suppose $a \in M$. Then we have $a = h(n)$ for some n. Hence

$$s(a) = s(h(n)) = h(n^+).$$

So $s(a) \in M$. Therefore by induction in the Peano system of N we get $M = N$. Thus h is onto, as desired. ■

Theorem 4.7. *There exists a unique function*

$$\begin{array}{c} \omega \times \omega \longrightarrow \omega \\ (m, n) \mapsto m + n \end{array},$$

called ***addition****, such that for every $m, n \in \omega$ we have*

(i) $m + 0 = m$,
(ii) $m + n^+ = (m + n)^+$.

Remark. Note that the above properties are sufficient for computing $m + n$ for any m, n. Because we know that $m + 0 = m$. Hence by the second property we get

$$m + 1 = m + 0^+ = (m + 0)^+ = m^+.$$

We can repeat this argument to compute $m + 2$, $m + 3$, and so forth.

Remark. $m + n$ is called the **sum** of m, n.

Proof. Fix $m \in \omega$. Then by recursion theorem there is a unique function $f_m : \omega \to \omega$ such that $f_m(0) = m$, and $f_m(n^+) = (f_m(n))^+$ for every n. Now we define

$$m + n := f_m(n).$$

Then it is obvious that $m + n$ has the required properties. To show the uniqueness, suppose $g : \omega \times \omega \to \omega$ also has the properties of addition. Then for fixed m we have $g(m, 0) = m$, and $g(m, n^+) = (g(m, n))^+$ for every n. Hence by the uniqueness of f_m we get

$$g(m, n) = f_m(n) = m + n.$$

But m is arbitrary, so g is the same function as addition. ■

***Example* 4.1.** We have

$$2 + 2 = 2 + 1^+ = (2 + 1)^+ = (2^+)^+ = 3^+ = 4.$$

Theorem 4.8. *Suppose $m, n, k \in \omega$. Then we have*

(i) *Associativity* :

$$(m + n) + k = m + (n + k).$$

(ii) *Identity element* :

$$0 + n = n = n + 0.$$

(iii) *Commutativity* :

$$m + n = n + m.$$

Remark. During the proof of this theorem, we will also show that

$$n^+ = 1 + n = n + 1.$$

Proof. **(i)** The proof is by induction on k. For $k = 0$ we have

$$m + (n + 0) = m + n = (m + n) + 0.$$

Now suppose the equality holds for some k. Then for k^+ we have

$$\begin{aligned} m + (n + k^+) &= m + (n + k)^+ = (m + (n + k))^+ \\ &= ((m + n) + k)^+ = (m + n) + k^+, \end{aligned}$$

as desired.

(ii) The equality $n + 0 = n$ holds due to the defining properties of addition. For the other equality we use induction on n. For $n = 0$ we have $0 + 0 = 0$. Now suppose $0 + n = n$ for some n. Then for n^+ we have

$$0 + n^+ = (0 + n)^+ = (n)^+ = n^+.$$

(iii) First let us prove by induction that $1 + n = n^+$ for every n. For $n = 0$ we have $1 + 0 = 1 = 0^+$. Now if we have $1 + n = n^+$, then we get

$$1 + n^+ = (1 + n)^+ = (n^+)^+,$$

as desired. Next, we prove the commutativity of addition by induction on n. For $n = 0$ we have

$$m + 0 = m = 0 + m,$$

as we have shown in the previous part. Now suppose that $m + n = n + m$ for some n. Then we have

$$\begin{aligned} m + n^+ &= (m + n)^+ = (n + m)^+ \\ &= 1 + (n + m) = (1 + n) + m = n^+ + m, \end{aligned}$$

as wanted. ■

Theorem 4.9. *There exists a unique function*

$$\begin{array}{c} \omega \times \omega \longrightarrow \omega \\ (m,n) \mapsto m \cdot n \end{array},$$

called ***multiplication****, such that for every* $m, n \in \omega$ *we have*

(i) $m \cdot 0 = 0$,
(ii) $m \cdot n^+ = (m \cdot n) + m$.

Remark. Note that the above properties are sufficient for computing $m \cdot n$ for any m, n. Because we know that $m \cdot 0 = 0$. Hence by the second property we get

$$m \cdot 1 = m \cdot 0^+ = (m \cdot 0) + m = 0 + m = m.$$

We can repeat this argument to compute $m \cdot 2$, $m \cdot 3$, and so forth.

Notation. We usually denote $m \cdot n$ simply by mn. We will also assume that multiplication binds stronger than addition; thus, for example, $mn+k$ means $(mn)+k$, not $m(n+k)$.

Remark. mn is called the **product** of m, n.

Proof. Fix $m \in \omega$. Then by recursion theorem (applied to addition) there is a unique function $f_m : \omega \to \omega$ such that $f_m(0) = 0$, and $f_m(n^+) = f_m(n) + m$ for every n. Now we define

$$m \cdot n := f_m(n).$$

Then it is obvious that $m \cdot n$ has the required properties. To show the uniqueness, suppose $g : \omega \times \omega \to \omega$ also has the properties of multiplication. Then for fixed m we have $g(m, 0) = 0$, and $g(m, n^+) = (g(m, n)) + m$ for every n. Hence by the uniqueness of f_m we get

$$g(m, n) = f_m(n) = m \cdot n.$$

But m is arbitrary, so g is the same function as multiplication. ■

***Example* 4.2.** We have

$$2 \cdot 2 = 2 \cdot 1^+ = 2 \cdot 1 + 2 = 2 + 2 = 4.$$

Theorem 4.10. *Suppose* $m, n, k \in \omega$. *Then we have*

(i) *Distributivity* :

$$m(n + k) = mn + mk.$$

(ii) *Associativity* :

$$(mn)k = m(nk).$$

(iii) *Commutativity* :

$$mn = nm.$$

(iv) *Identity element* :

$$1n = n = n1.$$

Remark. During the proof of this theorem, we will also show that

$$0n = 0 = n0.$$

Proof. **(i)** The proof is by induction on k. For $k = 0$ we have

$$m(n + 0) = mn = mn + 0 = mn + m0.$$

Now suppose the equality holds for some k. Then for k^+ we have

$$\begin{aligned} m(n + k^+) &= m(n + k)^+ = m(n + k) + m \\ &= (mn + mk) + m = mn + (mk + m) = mn + mk^+, \end{aligned}$$

as desired.

(ii) The proof is by induction on k. For $k = 0$ we have

$$m(n0) = m0 = 0 = (mn)0.$$

Now suppose the equality holds for some k. Then for k^+ we have

$$m(nk^+) = m(nk + n) = m(nk) + mn = (mn)k + mn = (mn)k^+,$$

as desired.

(iii) First let us prove by induction that $0m = 0$ for every m. For $m = 0$ we have $0 \cdot 0 = 0$. Now if we have $0m = 0$, then we get

$$0m^+ = 0m + 0 = 0 + 0 = 0,$$

as desired. Next, we prove by induction that $n^+m = nm + m$ for every m. For $m = 0$ we have

$$n^+0 = 0 = 0 + 0 = n0 + 0.$$

Now suppose $n^+m = nm + m$ for some m. Then for m^+ we have

$$\begin{aligned} n^+m^+ &= n^+m + n^+ = (nm + m) + n^+ \\ &= nm + (m + n^+) = nm + (m + n)^+ = nm + (n + m)^+ \\ &= nm + (n + m^+) = (nm + n) + m^+ = nm^+ + m^+. \end{aligned}$$

Finally, we prove the commutativity of multiplication by induction on n. For $n = 0$ we have

$$m0 = 0 = 0m.$$

Now suppose that $mn = nm$ for some n. Then we have

$$n^+m = nm + m = mn + m = mn^+,$$

as wanted.

(iv) First note that we have

$$n1 = n0^+ = n0 + n = 0 + n = n.$$

The other equality follows from commutativity of multiplication. ■

Remark. Similarly to the case of addition and multiplication, we can show that there exists a unique function

$$\begin{aligned} \omega \times \omega &\longrightarrow \omega \\ (m, n) &\mapsto m^n \end{aligned},$$

called **exponentiation**, such that for every $m, n \in \omega$ we have

(i) $m^0 = 1$,

(ii) $m^{n^+} = (m^n) \cdot m$.

We can prove the usual properties of exponentiation of natural numbers by induction; however, we will postpone this to Sections 5.2 and 5.6, in which we define the notion of power for more general objects. Let us only mention that we assume that exponentiation binds stronger than multiplication and addition; so, for example, $l + m^n k$ means $l + ((m^n)k)$. We also use the convention that m^{n^k} means $m^{(n^k)}$.

4.3 Order of Natural Numbers

Theorem 4.11. *Let $n \in \omega$. Then for every set a we have*

$$a \in n \implies a \subset n.$$

In other words, for every sets a, b we have

$$a \in n \text{ and } b \in a \implies b \in n.$$

Remark. Sets with the above property are called **transitive sets**. So the theorem says that natural numbers are transitive sets.

Proof. The proof is by induction on n. For $n = 0$ the claim is vacuously true, since 0 is the empty set. Now suppose n is transitive. Let $a \in n^+ = n \cup \{n\}$. If $a \in n$ then $a \subset n \subset n^+$. And if $a \in \{n\}$ then $a = n \subset n^+$. Thus n^+ is also transitive. ■

Definition 4.5. Let $m, n \in \omega$. Then we say $n < m$ if $n \in m$.

Remark. As usual, $n \leq m$ means that $n < m$ or $n = m$. Also, $m > n$ means $n < m$, and $m \geq n$ means $n \leq m$.

Theorem 4.12. *The relation $<$ on ω is a strict partial order, i.e. for every $n, m, k \in \omega$ we have*

(i) *Irreflexivity* : $n \not< n$.

(ii) *Transitivity* : *If $n < m$ and $m < k$, then $n < k$.*

Remark. As we showed in Theorem 3.23, it follows that $\leq$ is a non-strict partial order on ω.

Remark. The fact that $n \not< n$, which means $n \notin n$, follows from the axiom of regularity. But here we show that for natural numbers this fact also follows without using the axiom of regularity.

Proof. **(i)** The proof is by induction. For $n = 0$ we have $0 \notin 0$, since $0 = \emptyset$. Next suppose for some n we have $n \notin n$. Suppose to the contrary that $n^+ \in n^+ = n \cup \{n\}$. If $n^+ \in n$ then we get $n \in n$, because $n \in n^+$ and n is transitive. However, this contradicts the induction hypothesis. So we must have $n^+ \in \{n\}$. Hence $n^+ = n$. But this also leads to the contradiction $n \in n$, since we know that $n \in n^+$. Therefore $n^+ \notin n^+$, as desired.

(ii) If $m \in k$ then $m \subset k$, since k is a transitive set. Therefore $n \in m$ implies that $n \in k$. ■

Theorem 4.13. *For every $n \in \omega$ we have $0 \leq n$. In other words, 0 is the least element of ω. On the other hand, for every $n \in \omega$ we have $n < n^+$; so, ω does not have a largest element.*

Proof. It is obvious that $n \in n^+ = n \cup \{n\}$, i.e. $n < n^+$. Let us show by induction that $0 \leq n$ for every n. For $n = 0$ we obviously have $0 \leq 0$. Now suppose $0 \leq n$ for some n. If $0 = n$ then $0 < n^+$, since $n < n^+$. And if $0 < n$ then $0 < n^+$ by transitivity of $<$. Hence in either case we have $0 \leq n^+$, as wanted. ■

Theorem 4.14. *For every $n, m \in \omega$ we have*

$$n \leq m \qquad \Longleftrightarrow \qquad n \subset m.$$

Proof. If $n = m$ then obviously $n \subset m$. Also if $n \in m$ then $n \subset m$, since m is transitive. So $n \leq m$ implies $n \subset m$. For the reverse implication we use induction on m. Let

$$S := \{m \in \omega : n \subset m \text{ implies } n \leq m \text{ for every } n \in \omega\}.$$

Then $0 \in S$, because $n \subset 0 = \emptyset$ implies $n = \emptyset = 0$ (since we also have $\emptyset \subset n$). Now suppose $m \in S$. Let $n \subset m^+ = m \cup \{m\}$. If $m \notin n$ then every element of n must belong to m (since it cannot belong to $\{m\}$); hence we have $n \subset m$. Therefore $n \leq m < m^+$. Thus $n < m^+$ by transitivity of $<$. So let us assume that $m \in n$. Then $\{m\} \subset n$. We also have $m \subset n$, since n is a transitive set. Hence $m^+ = m \cup \{m\} \subset n$. Therefore we must have $n = m^+$, since we already knew that $n \subset m^+$. Thus $m^+ \in S$. ■

Theorem 4.15. *The partial order $<$ on ω is a linear order, i.e. it satisfies the trichotomy law. More explicitly, for every $n, m \in \omega$ exactly one of the three conditions*

$$n < m, \qquad n = m, \qquad n > m,$$

holds.

Proof. Note that if we have $n < m$ and $m < n$ then by transitivity we get $n < n$, which contradicts the irreflexivity of $<$. Similarly, if any other two of the above conditions hold we get a contradiction. Thus at most one of the above three conditions can hold. Let us show that at least one of them holds too. Let

$$S := \{n \in \omega : \text{for every } m \in \omega \text{ we have } n < m \text{ or } n = m \text{ or } n > m\}.$$

Then $0 \in S$, since as we have seen $0 \leq m$ for every m. Now suppose $n \in S$. Let $m \in \omega$. If $m \leq n$ then $m < n^+$, since $n < n^+$, and $<$ is transitive. So suppose $m > n$. This means that $n \in m$, which implies $\{n\} \subset m$. We also have $n \subset m$, since m is a transitive set. Therefore $n^+ = n \cup \{n\} \subset m$. Thus $n^+ \leq m$ by the above theorem. So $n^+ \in S$, as desired. ■

Theorem 4.16. *For any $n \in \omega$ there is no $k \in \omega$ such that*

$$n < k < n^+.$$

Hence for every $k \in \omega$ we have

$$\begin{aligned} k > n \qquad &\iff \qquad k \geq n^+, \\ k \leq n \qquad &\iff \qquad k < n^+. \end{aligned}$$

Proof. Suppose $n < k$, i.e. $n \in k$. Then $\{n\} \subset k$. We also know that $n \subset k$. Therefore $n^+ = n \cup \{n\} \subset k$. Hence $n^+ \leq k$. Thus there is no k such that $n < k < n^+$. The conclusions can also be proved easily. So far we have shown that $n < k$ implies $n^+ \leq k$. Conversely, $n^+ \leq k$ implies $n < k$, since $n < n^+$. On the other hand, $k \leq n$ implies $k < n^+$. Also, if $k < n^+$ then we cannot have $n < k$; hence by trichotomy law we must have $k \leq n$. ■

Theorem 4.17. *Let $n, m, k \in \omega$. Then we have*
(i) *If $n < m$ then $n + k < m + k$.*
(ii) *If $k > 0$ and $n < m$, then $nk < mk$.*

Proof. **(i)** The proof is by induction on k. For $k = 0$ we have

$$n + 0 = n < m = m + 0.$$

Suppose that $n + k < m + k$ for some k. And suppose to the contrary that $n + k^+ \not< m + k^+$. Thus by trichotomy law we have

$$n + k^+ \geq m + k^+ = (m + k)^+.$$

Then by the previous theorem we get

$$(n + k)^+ = n + k^+ > m + k > n + k.$$

Hence $m+k$ is a natural number between $n+k$ and $(n+k)^+$, which is a contradiction. Therefore we must have $n + k^+ < m + k^+$, as desired.

(ii) Let

$$S := \{k \in \omega : k = 0, \text{ or } nk < mk \text{ for every } n, m \in \omega \text{ where } n < m\}.$$

Then $0 \in S$ by definition. Suppose $k \in S$. If $k = 0$ we have

$$n0^+ = n1 = n < m = m1 = m0^+.$$

So $k^+ \in S$ in this case. And if $k \neq 0$, then we have $nk < mk$ for every $n < m$. Now by part (i) we have

$$nk^+ = nk + n < mk + n = n + mk < m + mk = mk + m = mk^+.$$

Thus $nk^+ < mk^+$ by transitivity. Hence $k^+ \in S$ in this case too. Therefore $S = \omega$, and the assertion of the theorem is true. ■

Cancellation Laws. *Let $n, m, k \in \omega$. Then we have*
(i) *If $n + k = m + k$ then $n = m$.*
(ii) *If $k \neq 0$ and $nk = mk$, then $n = m$.*

Proof. **(i)** Suppose to the contrary that $n \neq m$. Then by trichotomy law we have $n < m$ or $m < n$. Suppose $n < m$; the other case is similar. Then we get $n + k < m + k$, which is a contradiction. Hence we must have $n = m$.

(ii) Suppose to the contrary that $n \neq m$. Then we have $n < m$ or $m < n$. Suppose $n < m$; the other case is similar. Now note that $k \neq 0$ implies that $k > 0$, since we always have $0 \leq k$. Then we get $nk < mk$, which is a contradiction. Hence we must have $n = m$. ■

The order of natural numbers seems to be defined in terms of the specific structure of the elements of the set ω. However, the next theorem shows that the order of natural numbers can be expressed in terms of their addition, which is itself defined inductively by using the successor function. Thus it shows that we can define the order of natural numbers by only using the Peano axioms; although we are not going to follow that route in these notes.

Theorem 4.18. *For every $n, m \in \omega$ we have $n \leq m$ if and only if there exists $k \in \omega$ such that $m = n + k$.*

Proof. The proof is by induction on n. Let

$$S := \{n \in \omega : \text{for every } m \in \omega,\ n \leq m \text{ if and only if } m = n + k \text{ for some } k \in \omega\}.$$

Then $0 \in S$, because for every m we know that $0 \leq m$, and $m = 0 + m$. Now suppose $n \in S$. Let $n^+ \leq m$. Then $n < m$, since $n < n^+$. Hence there is k such that $m = n + k$. Also note that $k \neq 0$. Because otherwise we would have

$$n = n + 0 = n + k = m,$$

which contradicts $n < m$. Thus there is j such that $k = j^+$. Then we have

$$m = n + k = n + j^+ = (n + j)^+ = (j + n)^+ = j + n^+ = n^+ + j.$$

Conversely, suppose there are m, k such that $m = n^+ + k$. Then we have

$$m = n^+ + k = k + n^+ = (k + n)^+ = (n + k)^+ = n + k^+.$$

Hence by induction hypothesis we get $n \leq m$. However we cannot have $n = m$, since otherwise by cancellation law we would get

$$n + 0 = n = m = n + k^+ \implies 0 = k^+,$$

which is a contradiction. Thus we must have $n < m$. But this implies that $n^+ \leq m$ by Theorem 4.16. Therefore $n^+ \in S$, as desired. ■

Principle of Strong Induction. *Let $A \subset \omega$. Suppose that for every $n \in \omega$ we have*

$$\text{"if for every natural number } m < n \text{ we have } m \in A, \text{ then } n \in A."$$

Then we have $A = \omega$.

Remark. The principle of strong induction is not logically stronger than the principle of mathematical induction. In fact the two principles are equivalent, because we will show in this and the next two theorems that we have

$$\begin{aligned} \text{mathematical induction} &\implies \text{strong induction} \\ &\implies \text{well ordering principle} \\ &\implies \text{mathematical induction.} \end{aligned}$$

The difference of the two induction principles is that in strong induction the induction hypothesis is that every number less than n is in A, whereas in mathematical induction the induction hypothesis is that the predecessor of n is in A. The stronger induction hypothesis in strong induction can be helpful in some situations.

Remark. Note that when $n = 0$, the requirement of strong induction becomes

$$\text{"if for every natural number } m < 0 \text{ we have } m \in A, \text{ then } 0 \in A." \qquad (*)$$

Since there is no natural number less than 0, the phrase

$$\text{"for every natural number } m < 0 \text{ we have } m \in A"$$

is vacuously true. Thus, the conditional sentence $(*)$ is true if and only if $0 \in A$. Therefore when we want to check that the requirement of the strong induction holds for some set A, we also need to check the base case, i.e. whether 0 belongs to A. Although, sometimes the argument for $n > 0$ automatically works for $n = 0$ too. In those cases we can subsume the base case in the induction step, and we do not need to check it separately.

On the other hand, in some occasions we may have to check more than one base cases, because in their corresponding argument for general n we need to (tacitly) assume that $n > k$ for some $k \geq 1$. It should be noted that we do not need to prove a separate induction principle for these types of problems, and the principle of strong induction is strong enough for tackling them. We will see a problem of this kind in Example 4.8.

Proof. Let A be a set which satisfies the requirement of strong induction. Let

$$B := \{n \in \omega : \text{for every } m \leq n \text{ we have } m \in A\}.$$

We prove by the principle of mathematical induction that $B = \omega$. Then it is obvious that we will also have $A = \omega$. First note that by the above remark we must have $0 \in A$. Then we also have $0 \in B$, since $m \leq 0$ implies $m = 0$. Now suppose $n \in B$. Then by definition, for every $m \leq n$ we have $m \in A$. Now note that $m \leq n$ if and only if $m < n^+$. Hence for every natural number $m < n^+$ we have $m \in A$. Thus by our assumption about A we must have $n^+ \in A$. Therefore for every $m \leq n^+$ we have $m \in A$. So $n^+ \in B$. Thus $B = \omega$, as desired. ■

Well-Ordering Principle. *The linear order $<$ on ω is a well-ordering, i.e. every nonempty subset of ω has a least element.*

Proof. Let A be a subset of ω, and suppose that A does not have a least element. Consider $A^c = \omega - A$. We prove by strong induction that $A^c = \omega$. First note that $0 \in A^c$. Because if $0 \in A$ then 0 must be the least element of A. Now suppose for some n, every $m < n$ belongs to A^c. Then n cannot belong to A. Because otherwise n would be the least element of A, since all the natural numbers less than n do not belong to A by induction hypothesis. Hence we must have $n \in A^c$. Therefore by strong induction we get $A^c = \omega$. However, this means that A is empty. Hence if A is nonempty it must have a least element. ■

Remark. Note that in the above proof we could have subsumed the base case in the induction step, since the argument for general n also works for $n = 0$.

Theorem 4.19. *The well-ordering principle is equivalent to the principle of mathematical induction.*

Proof. We have already seen that the principle of mathematical induction implies the well-ordering principle. Let us prove the reverse implication. Let $A \subset \omega$ be such that $0 \in A$, and if $n \in A$ then $n^+ \in A$. We have to show that $A = \omega$. Suppose to the contrary that $A \neq \omega$. Then $A^c = \omega - A$ is nonempty. Hence by well-ordering principle A^c has a least element m. Note that $m \neq 0$, since $0 \in A$. Hence there is k such that $m = k^+$. But $k \notin A^c$, because $k < k^+$, and k^+ is the least element of A^c. Hence we must have $k \in A$. But by our assumption this implies that $k^+ \in A$, which is a contradiction. Therefore $A = \omega$, as desired. ■

Occasionally, we need the induction principles to prove that a property holds for every natural number n such that $n \geq k$, or $k \leq n \leq m$, for some $k, m \in \omega$. The following theorem allows us to do this.

Theorem 4.20. *Let $A \subset \omega$. Suppose $k, m \in \omega$, and $k \leq m$.*

(i) *Suppose $k \in A$, and for every natural number n where $k \leq n < m$, $n \in A$ implies that $n^+ \in A$. Then we have*

$$\{n \in \omega : k \leq n \leq m\} \subset A.$$

(ii) *Suppose $k \in A$, and for every natural number $n \geq k$, $n \in A$ implies that $n^+ \in A$. Then we have*

$$\{n \in \omega : n \geq k\} \subset A.$$

(iii) *Suppose for every natural number n where $k \leq n \leq m$ we have*

"if for every natural number $k \leq j < n$ we have $j \in A$, then $n \in A$."

Then we have

$$\{n \in \omega : k \leq n \leq m\} \subset A.$$

(iv) *Suppose for every natural number $n \geq k$ we have*

"if for every natural number $k \leq j < n$ we have $j \in A$, then $n \in A$."

Then we have

$$\{n \in \omega : n \geq k\} \subset A.$$

Remark. This theorem is particularly true when $k = 1$. Hence the induction principles also hold for $\mathbb{N} = \{n \in \omega : n \geq 1\}$.

Remark. Note that similarly to the case of $k = 0$, when we use this version of strong induction, we also need to check the base case, i.e. we have to check that $k \in A$.

Proof. **(i)** Let

$$B := \{n \in \omega : n < k \text{ or } n \in A \text{ or } n > m\}.$$

Then $0 \in B$. Because if $0 < k$ then $0 \in B$ by definition, and if $0 = k$ then $0 \in A \subset B$ (note that we cannot have $0 > k$). Now suppose $n \in B$. If $n^+ < k$ or $n^+ > m$, then $n^+ \in B$ by definition. If $n^+ = k$ then $n^+ \in A \subset B$. So suppose $k < n^+ \leq m$. Then we must have $k \leq n < m$; which also implies that $n \in A$, since $n \in B$. Thus by our assumption about A we get $n^+ \in A \subset B$. Therefore $B = \omega$. Hence for every n where $k \leq n \leq m$ we must have $n \in A$, since $n \not< k$ and $n \not> m$.

(ii) The proof is similar to part (i).

(iii) Similarly to the above let

$$B := \{n \in \omega : n < k \text{ or } n \in A \text{ or } n > m\}.$$

We want to show by strong induction that $B = \omega$. Let $n \in \omega$. Suppose for every $j < n$ we have $j \in B$. We have to show that $n \in B$. If $n < k$ or $n > m$, then $n \in B$ by definition. So suppose $k \leq n \leq m$. Then for every $k \leq j < n$ we have $j \in A$. Because we know that $j \in B$, and we have $k \leq j \leq m$. Hence by our assumption about A we get $n \in A \subset B$. Therefore by strong induction we obtain that $B = \omega$. Hence for every n where $k \leq n \leq m$ we must have $n \in A$, since $n \not< k$ and $n \not> m$.

(iv) The proof is similar to part (iii). ■

4.4 Finite Sets

The next theorem confirms our initial intuition that every natural number is the set of natural numbers less than itself. It also implies that ω is a transitive set too.

Theorem 4.21. *Let n be a natural number. Then we have*

$$n = \{k : k \text{ is a natural number, and } k < n\}.$$

Proof. By definition we know that a natural number k satisfies $k < n$ if and only if $k \in n$. Thus the only nontrivial part of the theorem is that the elements of a natural number are themselves natural numbers. We prove this by induction. Let

$$S := \{n \in \omega : \text{if } k \in n \text{ then } k \in \omega\}.$$

Then $0 \in S$, because the condition is vacuously true for $n = 0$. Suppose $n \in S$. Let $k \in n^+ = n \cup \{n\}$. If $k \in n$ then k is a natural number by induction hypothesis. And if $k \in \{n\}$ then $k = n$ is trivially a natural number. So $n^+ \in S$. ■

Remark. More explicitly, the above theorem says that

$$n = \{0, 1, 2, \ldots, n-1\},$$

where $n - 1$ is the unique number that satisfies $(n-1) + 1 = (n-1)^+ = n$, when $n > 0$. In particular, note that informally, n is a set that has n elements.

Definition 4.6. Let A be a set. We say A is **finite** if there exist a natural number n, and a bijective function $f : n \to A$. An **infinite** set is a set which is not finite.

Remark. In the light of the above theorem, we can say that A is finite if there exist a natural number n, and a bijective function $f : \{0, 1, 2, \ldots, n-1\} \to A$.

***Example* 4.3.** Every natural number n is a finite set, since the identity function from n to n is a bijective function.

Theorem 4.22. *Suppose A is a finite set, and $F : A \to A$. Then F is one-to-one if and only if it is onto.*

Remark. In fact, the converse of the above theorem is also true, i.e. if for every function $F : A \to A$, being one-to-one is equivalent to being onto, then A is a finite set. However, its proof requires a weak form of the axiom of choice; so we will postpone it until Chapter 7.

Proof. By our assumption there exists a bijective function $g : n \to A$ for some $n \in \omega$. Consider the function $f : n \to n$ where $f := g^{-1} \circ F \circ g$. Then $F = g \circ f \circ g^{-1}$. Hence F is one-to-one if and only f is one-to-one, since g, g^{-1} are one-to-one. Similarly, F is onto if and only if f is onto, since g, g^{-1} are onto. Thus it suffices to show that f is one-to-one if and only if f is onto. We prove this by induction on n. Let

$$S := \{n \in \omega : \text{every function } f : n \to n \text{ is one-to-one if and only if it is onto}\}.$$

First note that $0 \in S$, because the only function from 0 to 0 is $\emptyset$, which is vacuously one-to-one and onto. Now suppose $n \in S$. Let

$$f : n \cup \{n\} = n^+ \longrightarrow n^+ = n \cup \{n\}.$$

Suppose f is one-to-one. We have to show that f is onto. If $f(n) = n$, then for every $l \in n$ we have $f(l) \neq n$, since f is one-to-one. Hence for every $l \in n$ we must have $f(l) \in n$. Therefore $f|_n : n \to n$ is a function. Note that $f|_n$ is one-to-one, since if $f|_n(l_1) = f|_n(l_2)$ for some $l_1, l_2 \in n$, then we have

$$f(l_1) = f|_n(l_1) = f|_n(l_2) = f(l_2);$$

hence $l_1 = l_2$. Therefore by the induction hypothesis $f|_n$ is onto. Thus for every $k \in n$ there is $l \in n$ such that $k = f|_n(l) = f(l)$. So k is in the image of f. We also know that n is in the image of f, since $f(n) = n$. Hence f is onto.

Next suppose $f(n) \neq n$. Then we must have $f(n) = k$ for some $k \in n$. If n does not belong to the image of f, then the image of f is a subset of n. Hence similarly to the above, we can show that $f|_n : n \to n$ is a one-to-one function. Thus $f|_n$ is also onto. Hence there is $l \in n$ such that $k = f|_n(l) = f(l)$. So we get $f(l) = k = f(n)$, which contradicts the fact that f is one-to-one. Therefore n belongs to the image of f. Then we have $f(j) = n$ for some $j \in n$. We will switch the values of f at j, n to construct another function, i.e. we let $\tilde{f} : n^+ \to n^+$ to be defined by

$$\tilde{f}(l) := \begin{cases} k & l = j, \\ n & l = n, \\ f(l) & l \neq j, n. \end{cases} \qquad (*)$$

Let us show that $\tilde{f}$ is also a one-to-one function. First note that

$$\tilde{f} = f|_{n^+ - \{j,n\}} \cup \{(j, k), (n, n)\}.$$

So $\tilde{f}$ is a (piecewise-defined) function due to Theorem 3.20. Suppose $\tilde{f}(l_1) = \tilde{f}(l_2)$ for some $l_1, l_2 \in n^+$. If $l_1, l_2 \neq j, n$ we have

$$f(l_1) = \tilde{f}(l_1) = \tilde{f}(l_2) = f(l_2);$$

hence $l_1 = l_2$. And if $l_2 = j$ we get $\tilde{f}(l_1) = \tilde{f}(j) = k = f(n)$. However, for $l_1 \neq n$ we have $f(l_1) \neq k$. So we cannot have $l_1 \neq j, n$, since otherwise $\tilde{f}(l_1) = f(l_1) \neq k$. Thus l_1 is equal to j or n. But if $l_1 = n$ we get $\tilde{f}(l_1) = n \neq k$. Hence $l_1 = j = l_2$, as desired. Similarly, when $l_2 = n$ we can show that $l_1 = l_2$. Therefore $\tilde{f} : n^+ \to n^+$ is a one-to-one function that satisfies $\tilde{f}(n) = n$. Hence by the argument of the previous paragraph $\tilde{f}$ is onto. So for $i \in n^+$ there is $\tilde{l} \in n^+$ such that $i = \tilde{f}(\tilde{l})$. However, if $\tilde{l} \neq j, n$ we have $\tilde{f}(\tilde{l}) = f(\tilde{l})$, and otherwise we have $\tilde{f}(j) = f(n)$ and $\tilde{f}(n) = f(j)$. Therefore in either case we have $\tilde{j} = f(l)$ for some $l \in n^+$. So f is onto, as desired.

Conversely, suppose f is onto. We have to show that f is one-to-one. First suppose $f(n) \neq n$. Then we must have $f(n) = k$ for some $k \in n$. We also have $f(j) = n$ for some $j \in n$, since f is onto. Similarly to the above, we will switch the values of f at j, n to construct another function $\tilde{f}$ by equation $(*)$. Then $\tilde{f}$ is also onto. Because, similarly to the above, we can show that for every $l \in n^+$ there is $\tilde{l} \in n^+$ such that $\tilde{f}(\tilde{l}) = f(l)$. Thus $\tilde{f} : n^+ \to n^+$ is an onto function that satisfies $\tilde{f}(n) = n$. Hence by the argument of the next paragraph $\tilde{f}$ is one-to-one. However, this implies that f is also one-to-one, by an argument similarly to the one in the above paragraph. (Note that the relation between $f, \tilde{f}$ is symmetric, i.e. f can also be obtained from $\tilde{f}$ by switching its values at j, n. Hence in the arguments of the above paragraph we can switch the roles of $f, \tilde{f}$ to get appropriate arguments for this paragraph.)

Next suppose $f(n) = n$. Then $n \in f^{-1}(\{n\})$, where $f^{-1}(\{n\})$ is the preimage of $\{n\}$. Suppose to the contrary that $n \cap f^{-1}(\{n\}) \neq \emptyset$. Then in particular we have $n \neq \emptyset$. Let $k \in n$. Consider the function

$$g := f|_{n - f^{-1}(\{n\})} : n - f^{-1}(\{n\}) \longrightarrow n.$$

Then g is onto, since every element of n is in the image of f, and $f(f^{-1}(\{n\})) = \{n\}$ does not intersect $n = \{0, 1, \ldots, n-1\}$. Let us extend g to all of n by mapping all the elements of $n \cap f^{-1}(\{n\})$ to k, i.e. let

$$\tilde{g}(l) := \begin{cases} g(l) = f(l) & l \in n - f^{-1}(\{n\}), \\ k & l \in n \cap f^{-1}(\{n\}). \end{cases}$$

Then $\tilde{g}$ is an onto function from n to n, because g is onto. However, there is $l \in n - f^{-1}(\{n\})$ such that $g(l) = k$, since g is onto. Thus $\tilde{g}$ cannot be one-to-one, which contradicts the induction hypothesis. Therefore we must have $n \cap f^{-1}(\{n\}) = \emptyset$. In other words, the image of the set n under f does not intersect $\{n\}$. Hence the image of the set n under f is contained in n. Now consider the function $f|_n : n \to n$. Then $f|_n$ is onto, since every $j \in n$ is in the image of f, and $f(n) = n \notin n$. Thus $f|_n$ is one-to-one by induction hypothesis. Let us show that f is also one-to-one.

Suppose $f(l_1) = f(l_2)$ for some $l_1, l_2 \in n^+$. If $l_1, l_2 \in n$ then we have

$$f|_n(l_1) = f(l_1) = f(l_2) = f|_n(l_2);$$

hence $l_1 = l_2$. And if $l_1 = n$ then $f(l_2) = f(n) = n \notin n$. Thus $l_2 \notin n$. So we must have $l_2 = n = l_1$. Therefore f is one-to-one. Hence $n^+ \in S$, as desired. ■

Theorem 4.23. *ω is an infinite set.*

Proof. Consider the function $n \mapsto n^+$ from ω to ω. We know that this function is one-to-one; but it is not onto, since 0 is not in its image. Therefore by the previous theorem ω cannot be a finite set. Hence it is an infinite set. ■

Theorem 4.24. *Suppose A is a set, and n, m are natural numbers. If there exist bijective functions $f : n \to A$ and $g : m \to A$, then we must have $n = m$.*

Proof. Note that the function $h := g^{-1} \circ f : n \to m$ is bijective. Suppose to the contrary that $n \neq m$. Then $n < m$ or $n > m$. Suppose $n > m$; otherwise we can switch n, m and consider h^{-1} instead of h. Now we have $h : n \to m \subset n$. We know that h is one-to-one, and its image is m. However $m \neq n$, and this contradicts Theorem 4.22, since n is a finite set. Therefore we must have $n = m$. ■

Definition 4.7. Let A be a finite set. The unique natural number n for which there exists a bijective function $f : n \to A$ is called the **number of elements** of A. The number n is also called the **size** of A, or the **cardinality** of A, and is denoted by

$$|A|.$$

***Example* 4.4.** For every natural number n we have $|n| = n$. Because the identity function from n to n is a bijective function. In particular we have $|\emptyset| = 0$.

***Example* 4.5.** Suppose B is a finite set, and $|B| = 0$. Then $B = \emptyset$. Because there is a bijective function $h : 0 \to B$. Now if B is nonempty then there is some $b \in B$. And since h is onto we must have $(a, b) \in h$ for some $a \in 0 = \emptyset$, which is a contradiction. Hence B must be empty.

***Example* 4.6.** For a set A we have $|A| = 1$ if and only if A is a singleton. Because if $A = \{a\}$ is a singleton then $\{(\emptyset, a)\}$ is a bijective function from $1 = \{\emptyset\}$ to A. Conversely suppose $f : 1 \to A$ is bijective. Then there is $a \in A$ such that $(\emptyset, a) \in f$. In addition, for every $x \in A$ we must have $(c, x) \in f$ for some $c \in 1$, since f is onto. But this implies $c = \emptyset$; so $x = a$, because f is a function. Therefore A is a singleton.

Pigeonhole Principle. *Suppose A, B are finite sets, and we have $|A| > |B|$. Then a function $f : A \to B$ cannot be one-to-one.*

Remark. Thus there are at least two elements $a_1, a_2 \in A$ such that $f(a_1) = f(a_2)$. Informally, we can say that if there are more pigeons (elements of A) than pigeonholes (elements of B), then there must be a pigeonhole that contains at least two pigeons.

Proof. Suppose $|A| = m$ and $|B| = n$. Then $m > n$, and there are bijective functions $g : m \to A$ and $h : n \to B$. Now consider the function $F := h^{-1} \circ f \circ g : m \to n \subset m$. If f is one-to-one then F is also one-to-one, since h^{-1}, g are one-to-one. Also, the image of F is contained in n, because the image of h^{-1} is n. Thus the image of F is not all of m, which contradicts Theorem 4.22. Therefore f cannot be one-to-one. ■

Theorem 4.25. *Let A, B be two sets, and let $f : A \to B$.*

(i) *If A, B are finite, and $|A| < |B|$, then f cannot be onto.*
(ii) *If A is finite, and f is onto, then B is also finite, and $|A| \geq |B|$.*
(iii) *If B is finite, and f is one-to-one, then A is also finite, and $|A| \leq |B|$.*
(iv) *If one of the two sets A, B is finite, and f is bijective, then the other set is also finite, and we have $|A| = |B|$.*

Proof. **(i)** Suppose $|A| = m$ and $|B| = n$. Then $m < n$, and there are bijective functions $g : m \to A$ and $h : n \to B$. Consider the function $F := h^{-1} \circ f \circ g : m \to n$. If f is onto then F is also onto, since h^{-1}, g are onto. We know that $m \in n$. Also, note that the domain of F is $m \subset n$. Now we extend F to all of n by mapping all the elements of $\{l : m \leq l < n\}$ to m, i.e. let

$$\tilde{F}(l) := \begin{cases} F(l) & l < m, \\ m & l \geq m. \end{cases}$$

Then $\tilde{F}$ is an onto function from n to n, because F is onto. However, there is $l < m$ such that $F(l) = m$, since F is onto. Thus $\tilde{F}$ cannot be one-to-one, which contradicts Theorem 4.22. Therefore f cannot be onto.

(ii) If B is finite then by the previous part we obtain $|A| \geq |B|$, since otherwise f cannot be onto. Thus we only need to show that B is finite. Suppose $|A| = m$, and $g : m \to A$ is bijective. Then $h := f \circ g : m \to B$ is onto. We prove by induction on m that B is finite. Let

$$S := \{m \in \omega : \text{if there exists an onto function } h : m \to B \text{ for some set } B, \text{ then } B \text{ is finite}\}.$$

First let us show that $0 \in S$. Suppose to the contrary that B is nonempty. So there is some $b \in B$. Then since h is onto we must have $(a, b) \in h$ for some $a \in 0 = \emptyset$, which is a contradiction. Thus $B = \emptyset$; hence it is finite. Now suppose $m \in S$.

Let $h : m^+ \to B$ be an onto function. Keep in mind that $m^+ = m \cup \{m\}$. Let $c := h(m)$. Then for every $b \in B - \{c\}$ there is $k \in m$ such that $h(k) = b$. Consider $h|_m : m \to B$. If c belongs to the image of $h|_m$, then $h|_m$ is onto. Thus B is finite by induction hypothesis. So suppose c does not belong to the image of $h|_m$. Then $h|_m : m \to B - \{c\}$ is an onto function. Hence by induction hypothesis $B - \{c\}$ is finite. Thus there is a bijective function $F : n \to B - \{c\}$ for some $n \in \omega$. Let $\hat{F} := F \cup \{(n, c)\}$. Then it is easy to see that $\hat{F} : n^+ \to B$ is also a bijective function. Hence B is finite too. Therefore $m^+ \in S$, as desired.

(iii) If A is finite then by the pigeonhole principle we obtain $|A| \leq |B|$, since otherwise f cannot be one-to-one. Thus we only need to show that A is finite. Suppose $|B| = n$, and $h : n \to B$ is bijective. Then $g := f \circ h^{-1} : A \to n$ is one-to-one. We prove by induction on n that A is finite. Let

$$S := \{n \in \omega : \text{if there exists a one-to-one function } g : A \to n \text{ for some set } A, \text{ then } A \text{ is finite}\}.$$

First let us show that $0 \in S$. Suppose to the contrary that A is nonempty. So there is some $a \in A$. Then since g is a function we must have $(a, b) \in g$ for some $b \in 0 = \emptyset$, which is a contradiction. Thus $A = \emptyset$; hence it is finite. Now suppose $n \in S$. Let $g : A \to n^+$ be a one-to-one function. Keep in mind that $n^+ = n \cup \{n\}$. If n does not belong to the image of g, then $g : A \to n$ is a one-to-one function. Thus A is finite by induction hypothesis. So suppose n belongs to the image of g. Then there is a unique $c \in A$ such that $g(c) = n$, since g is one-to-one. Now consider $\tilde{g} := g|_{A-\{c\}} : A - \{c\} \to n$. Then $\tilde{g}$ is a one-to-one function. Hence by induction hypothesis $A - \{c\}$ is finite. Thus there is a bijective function $F : m \to A - \{c\}$ for some $m \in \omega$. Let $\hat{F} := F \cup \{(m, c)\}$. Then it is easy to see that $\hat{F} : m^+ \to A$ is also a bijective function. Hence A is finite too. Therefore $n^+ \in S$, as desired.

(iv) Note that f is both one-to-one and onto. Hence by the previous two parts we conclude that if A is finite then B is finite, and vice versa. In addition we get $|A| \geq |B|$ and $|A| \leq |B|$. Therefore $|A| = |B|$. ■

Theorem 4.26. *Suppose A is a finite set, and $B \subset A$. Then B is also finite, and we have $|B| \leq |A|$. In addition, $|B| = |A|$ implies that $B = A$.*

Proof. Let $\jmath : B \to A$ be the inclusion map. Then $\jmath$ is one-to-one. Hence by the previous theorem B is also finite, and we have $|B| \leq |A|$. Now suppose $|B| = |A|$. Suppose to the contrary that $B \neq A$. Let $a \in A - B$. Let $n = |B| = |A|$. So there are bijective functions $g : n \to A$ and $h : n \to B$. We can consider the composite function $f := g^{-1} \circ h : n \to n$, since $B \subset A$. Then f is a one-to-one function. Let $k := g^{-1}(a) \in n$. Then k cannot belong to the image of f, since otherwise we would have $k = f(j) = g^{-1}(h(j))$ for some j. But g^{-1} is one-to-one; so we get $h(j) = a$, contradicting the fact that $a \notin B$. Thus f is not onto. However this contradicts Theorem 4.22. Hence we must have $B = A$, as wanted. ■

Addition Principle. *Suppose A, B are finite sets, and we have $A \cap B = \emptyset$. Then $A \cup B$ is also a finite set, and*

$$|A \cup B| = |A| + |B|.$$

Proof. Suppose $|A| = m$ and $|B| = n$. The proof is by induction on n. If $n = 0$ then B is empty. Hence $A \cup B = A \cup \emptyset = A$ is a finite set, and we have

$$|A \cup B| = |A| = m = m + 0.$$

Now suppose the claim holds for some n. Let B be a finite set with $|B| = n^+$, such that $A \cap B = \emptyset$. Suppose $h : n \cup \{n\} \to B$ is a bijective function. Let $c = h(n) \in B$. Then it is easy to check that $h|_n : n \to B - \{c\}$ is a bijective function too. Therefore $B - \{c\}$ is a finite set, and $|B - \{c\}| = n$. In addition we have

$$A \cap (B - \{c\}) \subset A \cap B = \emptyset.$$

Hence by the induction hypothesis $C := A \cup (B - \{c\})$ is a finite set, and we have $|C| = m + n$. Thus there is a bijective function $F : m + n \to C$. Note that $c \notin C$, since $c \notin A$ because of $A \cap B = \emptyset$. Let $\hat{F} := F \cup \{(m + n, c)\}$. Then it is easy to see that

$$\hat{F} : (m + n)^+ \to C \cup \{c\} = A \cup B$$

is also a bijective function. Hence $A \cup B$ is a finite set, and we have $|A \cup B| = (m + n)^+ = m + n^+$, as desired. ■

The addition principle confirms our intuition about addition, namely, that in order to find the sum of n, m we form the union of two disjoint sets, one with n elements, and the other one with m elements, then we count the number of elements of the union. Similarly, the multiplication principle, stated below, confirms our intuition about multiplication. Namely, in order to find nm, we form a grid with n rows and m columns, then we count the number of elements in the grid. Note that intuitively, forming a grid is the same as forming the Cartesian product.

Multiplication Principle. *Suppose A, B are finite sets. Then $A \times B$ is also a finite set, and we have*

$$|A \times B| = |A| \cdot |B|.$$

Proof. Suppose $|A| = m$ and $|B| = n$. The proof is by induction on n. If $n = 0$ then B is empty. Hence $A \times B = A \times \emptyset = \emptyset$ is a finite set, and we have

$$|A \times B| = |\emptyset| = 0 = m \cdot 0.$$

Now suppose the claim holds for some n. Let B be a finite set with $|B| = n^+$. Suppose $h : n \cup \{n\} \to B$ is a bijective function. Let $c = h(n) \in B$, and let

$C := B - \{c\}$. Then it is easy to check that $h|_n : n \to C$ is a bijective function too. Therefore $|C| = n$. Hence by induction hypothesis $A \times C$ is a finite set, and $|A \times C| = mn$. Also note that $B = C \cup \{c\}$, and $C \cap \{c\} = \emptyset$. Hence we have

$$A \times B = (A \times C) \cup (A \times \{c\}),$$

and $(A \times C) \cap (A \times \{c\}) = A \times (C \cap \{c\}) = A \times \emptyset = \emptyset$. Now consider the function $a \mapsto (a, c)$ from A to $A \times \{c\}$. It is easy to see that this function is bijective. Therefore $A \times \{c\}$ is also a finite set, and we have $|A \times \{c\}| = |A| = m$. Thus by addition principle we get

$$|A \times B| = |A \times C| + |A \times \{c\}| = mn + m = mn^+,$$

as desired. ■

Notation. Let $m, n, k \in \omega$.
(i) If $n + k = m$, then we write $k = m - n$.
(ii) If $nk = m$, and $n \neq 0$, then we write $k = \frac{m}{n}$.

Theorem 4.27. *Suppose A, B are finite sets, and $A \subset B$. Then $B - A$ is also a finite set, and we have*

$$|B - A| = |B| - |A|.$$

Proof. First note that $B - A$ is a subset of B, so it is a finite set. Now we have $B = A \cup (B - A)$, and $A \cap (B - A) = \emptyset$. Thus by addition principle we get

$$|B| = |A| + |B - A|,$$

which gives the desired. ■

Theorem 4.28. *Let $k, m \in \omega$, and suppose $k \leq m$. Then the set*

$$\{k, \ldots, m\} := \{n \in \omega : k \leq n \leq m\}$$

has $m - k + 1$ elements.

Proof. Note that $\{k, \ldots, m\} = \{n : n \leq m\} - \{n : n < k\} = m^+ - k$, where the difference of m^+, k is their set theoretic difference. Now by the above theorem we get

$$|\{k, \ldots, m\}| = |m^+| - |k| = m^+ - k.$$

We know that there is j such that $k + j = m$. So by definition we have $j = m - k$. Hence $k + j^+ = (k + j)^+ = m^+$. So $j^+ = m^+ - k$, i.e. $m^+ - k = (m - k) + 1$. Thus we get the desired. ■

Inclusion-Exclusion Principle. *Suppose A, B are finite sets. Then $A \cup B$ and $A \cap B$ are also finite sets, and we have*

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Proof. First note that $A \cap B$ is a subset of B; so it is a finite set. Let $C := B - A \cap B$. Then C is also a finite set, and we have

$$|C| = |B| - |A \cap B|.$$

We also have $A \cap C = \emptyset$. Because for $a \in A \cap C$ we have $a \in A$ and $a \in C$. But $a \in C$ means that $a \in B$ and $a \notin A \cap B$. So we get $a \in A \cap B$ and $a \notin A \cap B$, which is a contradiction. In addition we have $A \cup C = A \cup B$. Note that $A \cup C \subset A \cup B$, since $C \subset B$. Now let $a \in A \cup B$. If $a \in A$ then $a \in A \cup C$. And if $a \notin A$ then we must have $a \in B$. So we have $a \in B$, while $a \notin A \cap B \subset A$. Hence $a \in C \subset A \cup C$. Thus in either case we get $a \in A \cup C$. Therefore $A \cup C = A \cup B$. However, by addition principle we know that $A \cup C$ is a finite set. Thus $A \cup B$ is a finite set, and we have

$$|A \cup B| = |A \cup C| = |A| + |C| = |A| + |B| - |A \cap B|,$$

as wanted. ■

Theorem 4.29. *Suppose A, B are finite sets. Then B^A, i.e. the set of all functions from A to B, is also a finite set, and we have*

$$\left|B^A\right| = |B|^{|A|}.$$

Proof. Suppose $|A| = m$ and $|B| = n$. The proof is by induction on m. If $m = 0$ then A is empty. Hence $A \times B = \emptyset \times B = \emptyset$. But any function $f : A \to B$ is a subset of $A \times B = \emptyset$. Thus the only function from A to B is $\emptyset$. Therefore

$$|B^A| = 1 = n^0.$$

Now suppose the claim holds for some m. Let A be a finite set with $|A| = m^+$. Then A is nonempty, since $|A| > 0$. Let $c \in A$, and $C := A - \{c\}$. Then $|C| = m$. Thus by induction hypothesis B^C is a finite set, and $|B^C| = n^m$. Consider the function $F : B^A \to B^C \times B$ whose value at any function $f \in B^A$ is given by

$$F(f) := (f|_C, f(c)).$$

We claim that F is bijective. Suppose $F(f) = F(g)$. Then $f|_C = g|_C$, and $f(c) = g(c)$. Hence for every $a \in A$ we have $f(a) = g(a)$. Because we already know

that the equality holds when $a = c$; and when $a \neq c$ then $a \in C$, so we have $f(a) = f|_C(a) = g|_C(a) = g(a)$. Therefore $f = g$. Thus F is one-to-one.

Next let us show that F is onto. Let $(h, b) \in B^C \times B$. Let $f : A \to B$ be defined as follows:

$$f(a) := \begin{cases} h(a) & a \in C, \\ b & a = c. \end{cases}$$

Then $f|_C = h$. Hence $F(f) = (h, b)$. So F is onto; and therefore it is bijective. Thus B^A is a finite set, since $B^C \times B$ is a finite set by multiplication principle. In addition we have

$$|B^A| = |B^C \times B| = |B^C| \cdot |B| = n^m n = n^{m^+},$$

as desired. ■

Theorem 4.30. *Suppose A is a finite set. Then $\mathcal{P}(A)$, i.e. the power set of A, is also a finite set, and we have*

$$|\mathcal{P}(A)| = 2^{|A|}.$$

Proof. Consider 2^A, the set of all functions from A to $2 = \{0, 1\}$. Let $F : \mathcal{P}(A) \to 2^A$ be the function whose value at any subset $B \subset A$ is the function $F_B : A \to 2$, which is defined by

$$F_B(a) := \begin{cases} 1 & a \in B, \\ 0 & a \in A - B. \end{cases}$$

In other words, $F_B = (B \times \{1\}) \cup ((A - B) \times \{0\})$. We claim that F is bijective. Suppose $F_B = F_C$. Then for every $a \in A$ we have

$$a \in B \iff F_B(a) = 1 \iff F_C(a) = 1 \iff a \in C.$$

Thus $B = C$. So F is one-to-one. Next let us show that F is onto. Let $f \in 2^A$. Let

$$B := \{a \in A : f(a) = 1\}.$$

Then for $a \in A$ we have

$$F_B(a) = 1 \iff a \in B \iff f(a) = 1,$$
$$F_B(a) = 0 \iff a \notin B \iff f(a) \neq 1.$$

Now note that $f(a) \neq 1$ is equivalent to $f(a) = 0$, because the codomain of f is $\{0, 1\}$. Thus $F_B = f$. Hence F is bijective. Therefore $\mathcal{P}(A)$ is a finite set, since 2^A is a finite set by the previous theorem. We also have

$$|\mathcal{P}(A)| = |2^A| = |2|^{|A|} = 2^{|A|},$$

as desired. ■

Remark. Note that in the above proof, we did not use the finiteness of A in the construction of the one-to-one correspondence between $\mathcal{P}(A)$ and 2^A.

At the end of this section, we prove a very useful fact about finite subsets of linearly ordered sets.

Theorem 4.31. *Suppose $\preceq$ is a linear order on a set X. Let $A \subset X$ be a nonempty finite subset. Then A has a maximum and a minimum.*

Proof. Suppose $|A| = n$. Then $n \geq 1$, since A is nonempty. The proof is by induction on n. If $n = 1$ then A is a singleton. Thus $A = \{a\}$ for some $a \in X$. Then it is obvious that a is both the maximum and the minimum of A. Now suppose the claim holds for some n. Let A be a subset with $|A| = n^+$. Let $c \in A$, and $C := A - \{c\}$. Then $|C| = n$. Hence by induction hypothesis C has a minimum $a \in C$, and a maximum $b \in C$. Note that in a linearly ordered set every two elements are comparable. Now if $c \leq a$ then c is the minimum of A. Because for every $x \in A$, if $x \neq c$ we have $x \in C$, so $c \leq a \leq x$. And if $x = c$ we trivially have $c \leq x$. Similarly, if $a \leq c$ then a is the minimum of A. Also, we can similarly show that if $b \leq c$ then c is the maximum of A, and if $c \leq b$ then b is the maximum of A. Therefore A has maximum and minimum. Hence the claim holds for n^+ too. ■

4.5 Finite Sequences

The recursion theorem allows us to construct a function in infinitely many steps. Sometimes we only need finitely many steps to construct a function. In addition, sometimes we need to define a function starting from some natural number k, instead of 0. The following theorem allows us to do these. (At the end of this section, we will see a more general version of recursion theorem that allows us to define functions whose values at n depend not only on their values at the predecessor of n, but also on their values at several numbers less than n.)

Theorem 4.32. *Let $k, m \in \omega$, and suppose $k < m$. Let J be one of the two sets*

$$\{n \in \omega : n \geq k\}, \qquad \text{or} \qquad \{n \in \omega : k \leq n \leq m\}.$$

Let A be a set, and let $a \in A$. Suppose $F : J \times A \to A$ is a function. Then there exists a unique function $f : J \to A$ such that

(i) *$f(k) = a$, and*
(ii) *$f(n+1) = F(n, f(n))$ for every $n \in J$ where $n + 1 \in J$.*

Proof. We extend F to a function from $\omega \times A$ to A by mapping all the elements of $J^c \times A$ to a, i.e. we set

$$\tilde{F}(n, x) := \begin{cases} F(n, x) & n \in J, \\ a & n \notin J. \end{cases}$$

Then by recursion theorem there is a function $\tilde{f} : \omega \to A$ such that $\tilde{f}(0) = a$, and $\tilde{f}(n+1) = \tilde{F}(n, \tilde{f}(n))$ for every n. Now let $f := \tilde{f}|_J$. Suppose $n, n+1 \in J$. Then we have

$$f(n+1) = \tilde{f}(n+1) = \tilde{F}(n, \tilde{f}(n)) = F(n, f(n)).$$

Also, if $k = 0$ then $f(k) = f(0) = \tilde{f}(0) = a$. So suppose $k \neq 0$. Then $k = j + 1$ for some j. In addition, note that $j \notin J$, since $j < k$. Hence we have

$$f(k) = \tilde{f}(k) = \tilde{f}(j+1) = \tilde{F}(j, \tilde{f}(j)) = a,$$

as desired. Finally, let us show that f is unique. Suppose to the contrary that there exists another function $g : J \to A$ that satisfies the requirements of the theorem. Then the set $\{n \in J : f(n) \neq g(n)\}$ must be nonempty. Let l be its least element. Then $l > k$, since $f(k) = a = g(k)$. Hence $l = i + 1$ for some $i \geq k$. Then $f(i) = g(i)$, because $i < l$, and l is the least natural number at which f, g differ. But $i, i+1 \in J$; so we get

$$f(l) = f(i+1) = F(i, f(i)) = F(i, g(i)) = g(i+1) = g(l),$$

which is a contradiction. Therefore $g = f$, and hence f is unique. ■

Definition 4.8. Let $k, m \in \omega$, and suppose $k \leq m$. A **sequence** in a set X is a function

$$\begin{array}{c} \{n \in \omega : n \geq k\} \to X \\ n \mapsto a_n \end{array}.$$

A **finite sequence** in X is a function

$$\begin{array}{c} \{n \in \omega : k \leq n \leq m\} \to X \\ n \mapsto a_n \end{array}.$$

The **length** of the above finite sequence is the size of the set $\{n \in \omega : k \leq n \leq m\}$, which we have shown to be $m - k + 1$.

Remark. When we want to emphasize that a sequence is not a finite sequence, we call it an *infinite sequence*. Also, in the notation a_n, the number n is called the **index**.

Now we want to define addition and multiplication of several numbers. Remember that addition and multiplication of two numbers are functions from $\omega \times \omega$ to ω; thus their input must be an ordered pair. Although, we have shown later that the order of a pair of numbers does not affect their sum or their product. Similarly, when we define the sum or product of several numbers, we have to initially assume that they have an order. Later we can prove that the order of the numbers does not affect their sum or their product. Thus we are going to define the sum and the

product of a finite sequence of numbers. These definitions have a similar structure, and in fact, they are special cases of a general notion of applying an operation to a finite sequence of objects. Since we will use this notion for other operations too, we are going to define it in the utmost generality.

Definition 4.9. A **binary operation** on a set S is a function

$$\star : S \times S \to S.$$

For two elements $a, b \in S$, we usually write $a \star b$ instead of $\star(a, b)$.

***Example* 4.7.** Addition and multiplication are binary operations on ω.

Definition 4.10. Suppose $\star$ is a binary operation on a set S, and $a_k, \ldots, a_m$ is a finite sequence in S. We inductively define the *(standard) product* of $a_k, \ldots, a_m$ to be

(i) $\prod_{j=k}^{k} a_j := a_k$,
(ii) $\prod_{j=k}^{n+1} a_j := \left(\prod_{j=k}^{n} a_j\right) \star a_{n+1}$ for $k \leq n < m$.

Remark. In the terminology of recursion theorem, we have constructed the function $f(n) = \prod_{j=k}^{n} a_j$, by using the function $F(n, s) = s \star a_{n+1}$. Note that the sequence $j \mapsto a_j$ is given to us; so we know a priori what a_{n+1} is.

Remark. We also denote $\prod_{j=k}^{n} a_j$ by $\prod_{k \leq j \leq n} a_j$. When k is 0 or 1, we may also denote it simply by $\prod_{j \leq n} a_j$. The variable j in the notation $\prod_{j=k}^{n} a_j$ is called a *dummy variable*, because we can change it without causing any harm. For example, we can denote $\prod_{j=k}^{n} a_j$ by $\prod_{i=k}^{n} a_i$, or by $\prod_{l=k}^{n} a_l$.

Remark. When the binary operation is denoted by $a \cdot b$, or simply by ab, then we keep using the notation $\prod$ for the product of several elements. But when the binary operation is denoted by $a + b$, we use the notation $\sum$ instead of $\prod$, and we use the term "sum" instead of "product". So if $n_k, \ldots, n_m$ is a finite sequence of natural numbers, their product will be denoted by $\prod_{j=k}^{m} n_j$, and their sum will be denoted by $\sum_{j=k}^{m} n_j$. We may also denote the sum and product of $n_k, \ldots, n_m$ by

$$n_k + \cdots + n_m, \qquad n_k \cdots n_m,$$

respectively.

***Exercise* 4.1.** Let $n \geq 1$. Show by induction that

(i) $\sum_{j=1}^{n} 1 = n$.
(ii) $1 + \cdots + n = \frac{1}{2} n(n+1)$.
(iii) $1^2 + \cdots + n^2 = \frac{1}{6} n(n+1)(2n+1)$.
(iv) $1^3 + \cdots + n^3 = \frac{1}{4} n^2 (n+1)^2$.

Definition 4.11. Let $j \mapsto A_j$ be a (finite or infinite) sequence of sets in a set X. Then we inductively define

(i) $\bigcup_{j=k}^{k} A_j = A_k$,

(ii) $\bigcup_{j=k}^{n+1} A_j = \left(\bigcup_{j=k}^{n} A_j\right) \cup A_{n+1}$.

And

(i) $\bigcap_{j=k}^{k} A_j = A_k$,

(ii) $\bigcap_{j=k}^{n+1} A_j = \left(\bigcap_{j=k}^{n} A_j\right) \cap A_{n+1}$.

Remark. In the terminology of recursion theorem, we have constructed the function $f(n) = \bigcup_{j=k}^{n} A_j$, by using the function $F(n, A) = A \cup A_{n+1}$, and the function $f(n) = \bigcap_{j=k}^{n} A_j$, by using the function $F(n, A) = A \cap A_{n+1}$. Note that the sequence $j \mapsto A_j$ is given to us; so we know a priori what A_{n+1} is. Also note that the codomain of F, f is $\mathcal{P}(\bigcup X)$. Because if $A \in X$ then $A \subset \bigcup X$; hence $A \in \mathcal{P}(\bigcup X)$. In addition, $\mathcal{P}(\bigcup X)$ contains the intersection or union of its elements, due to Theorem 2.19. Hence we can view $\cap, \cup$ as binary operations on $\mathcal{P}(\bigcup X)$. Then the above definition becomes the definition of the product with respect to these binary operations.

Remark. We also denote $\bigcap_{j=k}^{n} A_j$ by $\bigcap_{k\le j\le n} A_j$, and $\bigcup_{j=k}^{n} A_j$ by $\bigcup_{k\le j\le n} A_j$. In addition, when the sets in the sequence are pairwise disjoint, i.e. for $i \neq j$ we have $A_i \cap A_j = \emptyset$, we denote $\bigcup_{j=k}^{n} A_j$ by $\bigsqcup_{j=k}^{n} A_j$, or $\bigsqcup_{k\le j\le n} A_j$. Furthermore, when k is 0 or 1, we may also use the notations $\bigcap_{j\le n} A_j$, $\bigcup_{j\le n} A_j$, or $\bigsqcup_{j\le n} A_j$.

Theorem 4.33. *Let $j \mapsto A_j$ be a finite sequence of sets in a set X, and suppose $X = \{A_k, \dots, A_m\}$. Then we have*

$$\bigcap X = \bigcap_{j=k}^{m} A_j, \qquad \text{and} \qquad \bigcup X = \bigcup_{j=k}^{m} A_j.$$

Proof. The proof is by induction on m. When $m = k$ we have $\bigcup_{j=k}^{k} A_j = A_k = \bigcap_{j=k}^{k} A_j$. We also have $X = \{A_k\}$. Hence $\bigcap X = A_k = \bigcup X$. Thus the desired equalities hold. Now suppose the claim is true for some m. Consider the sequence $A_k, \dots, A_m, A_{m+1}$. Let $Y := \{A_k, \dots, A_m\}$. Then by induction hypothesis we have

$$\begin{aligned}\bigcup_{j=k}^{m+1} A_j &= \left(\bigcup_{j=k}^{m} A_j\right) \cup A_{m+1} = \left(\bigcup Y\right) \cup A_{m+1} \\ &= \left(\bigcup Y\right) \cup \left(\bigcup \{A_{m+1}\}\right) = \bigcup (Y \cup \{A_{m+1}\}) = \bigcup X.\end{aligned}$$

Note that here we have used the result of Exercise 2.2. The equality for intersections can be proved similarly. ∎

Remark. The above theorem shows that the new notions of union and intersection of several sets is in agreement with the previous notions. However, using sequences instead of sets is preferable in some situations; because unlike sets, the elements of a sequence have an order, and can have repetitions.

Remark. Suppose $j \mapsto A_j$ is an infinite sequence of sets in a set X, and $X = \{A_k, \dots, A_m, \dots\}$, i.e. the sequence is an onto map. Then, motivated by the above theorem, we sometimes denote $\bigcap X$ by $\bigcap_{j=k}^{\infty} A_j$, and $\bigcup X$ by $\bigcup_{j=k}^{\infty} A_j$. We may also use the notations $\bigcap_{j\geq k} A_j$, or $\bigcup_{j\geq k} A_j$.

***Exercise* 4.2.** Let $j \mapsto A_j$ be a finite sequence of sets, and let C be a set. Show that we have

(i) Distributivity :

$$\Big(\bigcap_{j=k}^{m} A_j\Big) \cup C = \bigcap_{j=k}^{m} (A_j \cup C), \qquad \text{and} \qquad \Big(\bigcup_{j=k}^{m} A_j\Big) \cap C = \bigcup_{j=k}^{m} (A_j \cap C).$$

(ii)

$$\Big(\bigcap_{j=k}^{m} A_j\Big) \cap C = \bigcap_{j=k}^{m} (A_j \cap C), \qquad \text{and} \qquad \Big(\bigcup_{j=k}^{m} A_j\Big) \cup C = \bigcup_{j=k}^{m} (A_j \cup C).$$

(iii) De Morgan's laws :

$$C - \Big(\bigcap_{j=k}^{m} A_j\Big) = \bigcup_{j=k}^{m} (C - A_j), \qquad \text{and} \qquad C - \Big(\bigcup_{j=k}^{m} A_j\Big) = \bigcap_{j=k}^{m} (C - A_j).$$

Theorem 4.34. *Suppose $A_1, \dots, A_m$ is a finite sequence of pairwise disjoint sets, i.e. for $i \neq j$ we have $A_i \cap A_j = \emptyset$. Also suppose that each A_j is a finite set. Then $\bigcup_{j\leq m} A_j$ is also a finite set, and we have*

$$\Big|\bigcup_{j\leq m} A_j\Big| = |A_1| + \cdots + |A_m|.$$

Proof. The proof is by induction on m. If $m = 1$ then $\bigcup_{j\leq 1} A_j = A_1$ is a finite set, and we have $|\bigcup_{j\leq 1} A_j| = |A_1|$. Now suppose the claim is true for some m. Consider the sequence $A_1, \dots, A_m, A_{m+1}$. Then by induction hypothesis $\bigcup_{j\leq m} A_j$ is a finite set. Also note that $A_{m+1} \cap (\bigcup_{j\leq m} A_j) = \emptyset$. Because if x belongs to $\bigcup_{j\leq m} A_j = \bigcup\{A_1, \dots, A_m\}$, then it must belong to A_j for some $j \leq m$. However, we assumed that $A_{m+1} \cap A_j = \emptyset$; so $x \notin A_{m+1}$. Hence by addition principle $\bigcup_{j\leq m+1} A_j = (\bigcup_{j\leq m} A_j) \cup A_{m+1}$ is a finite set, and we have

$$\Big|\bigcup_{j\leq m+1} A_j\Big| = \Big|\bigcup_{j\leq m} A_j\Big| + |A_{m+1}| = \Big(\sum_{j\leq m} |A_j|\Big) + |A_{m+1}| = \sum_{j\leq m+1} |A_j|,$$

as desired. ∎

Theorem 4.35. *Suppose $A_1, \ldots, A_m$ is a finite sequence of sets. Also suppose that each A_j is a finite set. Then $\bigcup_{j\leq m} A_j$ is also a finite set, and we have*

$$\Big|\bigcup_{j\leq m} A_j\Big| \leq |A_1| + \cdots + |A_m|.$$

Remark. In other words, the union of finitely many finite sets is finite.

Proof. The proof is by induction on m. If $m = 1$ then $\bigcup_{j\leq 1} A_j = A_1$ is a finite set, and we have $|\bigcup_{j\leq 1} A_j| = |A_1|$. Now suppose the claim is true for some m. Consider the sequence $A_1, \ldots, A_m, A_{m+1}$. Then by induction hypothesis $\bigcup_{j\leq m} A_j$ is a finite set. Hence by inclusion-exclusion principle $\bigcup_{j\leq m+1} A_j = (\bigcup_{j\leq m} A_j) \cup A_{m+1}$ is a finite set, and we have

$$\Big|\bigcup_{j\leq m+1} A_j\Big| \leq \Big|\bigcup_{j\leq m} A_j\Big| + |A_{m+1}| \leq \Big(\sum_{j\leq m} |A_j|\Big) + |A_{m+1}| = \sum_{j\leq m+1} |A_j|,$$

as desired. ∎

Definition 4.12. Let $n \in \omega$. The ***n* factorial** is inductively defined as follows

(i) $0! = 1$,
(ii) $(n + 1)! = (n + 1) \cdot n!$.

Remark. In the terminology of recursion theorem, we have constructed the function $f(n) = n!$, by using the function $F(n, m) = (n + 1)m$.

Remark. It is easy to see that when $n \geq 1$, $n!$ is the product of all nonzero natural numbers less than or equal to n, i.e.

$$n! = n \cdot (n - 1) \cdots 2 \cdot 1.$$

Theorem 4.36. *Let A, B be finite sets. Suppose $|A| = k$, $|B| = n$, and $k \leq n$. Then the set of one-to-one functions from A to B has*

$$\frac{n!}{(n-k)!}$$

elements.

Remark. Note that $n-k$ is the unique natural number that satisfies $k+(n-k) = n$. The existence of $n - k$ is guaranteed by Theorem 4.18, since $k \leq n$; and the uniqueness of $n - k$ follows from cancellation law.

Remark. In particular, the set of sequences of length k with distinct terms in a set of size n has $\frac{n!}{(n-k)!}$ elements.

Proof. Let $S(A;B)$ be the set of one-to-one functions from A to B. Note that the set of functions from A to B is a finite set; so $S(A;B)$ is also a finite set. We show that $|S(A;B)| = \frac{n!}{(n-k)!}$. The proof is by induction on k. Let

$$J := \{k \in \omega : \text{the set of one-to-one functions from a set of size } k \text{ to a set of size } n \text{ has } \tfrac{n!}{(n-k)!} \text{ elements, for every } n \geq k\}.$$

If $k = 0$ then A is empty; and the only function from $\emptyset$ to B is $\emptyset$, which is vacuously one-to-one. Hence there is only one one-to-one function from A to B; and we have $1 = \frac{n!}{(n-0)!}$. Thus $0 \in J$. Now suppose $k \in J$. Let A be a finite set with $|A| = k+1$. Then A is nonempty, since $|A| > 0$. Let $c \in A$, and $C := A - \{c\}$. Then $|C| = k$. Let B be a finite set with $|B| = n \geq k+1$. Then for $f \in S(A;B)$ we have $f|_C \in S(C; B - \{f(c)\})$, because $f|_C$ is one-to-one, so it cannot map any element of C to $f(c)$.

We know that there is a bijective function from n to B. So by using this bijective function we can list all the elements of B as a sequence $b_1, b_2, \ldots, b_n$. Let

$$F : S(A;B) \to \bigcup_{j \leq n} S(C; B - \{b_j\}) \times \{b_j\}$$

be the function whose value at any function $f \in S(A;B)$ is given by

$$F(f) := (f|_C, f(c)).$$

We claim that F is bijective. Suppose $F(f) = F(g)$. Then $f|_C = g|_C$, and $f(c) = g(c)$. Hence for every $a \in A$ we have $f(a) = g(a)$. Because we already know that the equality holds when $a = c$; and when $a \neq c$ then $a \in C$, so we have $f(a) = f|_C(a) = g|_C(a) = g(a)$. Therefore $f = g$. Thus F is one-to-one.

Next let us show that F is onto. Let $(h, b_j) \in S(C; B - \{b_j\}) \times \{b_j\}$ for some j. Let $f : A \to B$ be defined as follows:

$$f(a) := \begin{cases} h(a) & a \in C, \\ b_j & a = c. \end{cases}$$

Then $f|_C = h$. Let us show that f is one-to-one. Suppose $f(a_1) = f(a_2)$ for some $a_1, a_2 \in A$. If $a_1, a_2 \neq c$ we have

$$h(a_1) = f(a_1) = f(a_2) = h(a_2);$$

hence $a_1 = a_2$. And if $a_2 = c$ we get $f(a_1) = f(c) = b_j$. However, for $a_1 \neq c$ we have $f(a_1) = h(a_1) \neq b_j$. So we must have $a_1 = c = a_2$. Therefore we have $f \in S(A;B)$, and $F(f) = (h, b_j)$. Thus F is onto; and therefore it is bijective.

Now note that $n = m+1$ for some m, and $k \leq m$. Thus by induction hypothesis $S(C; B - \{b_j\})$ has $\frac{m!}{(m-k)!}$ elements. Also note that for $i \neq j$ we have

$$\big(S(C; B - \{b_i\}) \times \{b_i\}\big) \cap \big(S(C; B - \{b_j\}) \times \{b_j\}\big) = \emptyset,$$

since $b_i \neq b_j$. Hence we get

$$\begin{aligned} |S(A;B)| &= \Big| \bigcup_{j \leq n} S(C; B - \{b_j\}) \times \{b_j\} \Big| \\ &= \sum_{j \leq n} \big| S(C; B - \{b_j\}) \times \{b_j\} \big| \\ &= \sum_{j \leq n} \tfrac{m!}{(m-k)!} \cdot 1 = \tfrac{m!}{(m-k)!} \sum_{j \leq n} 1 = \tfrac{m!n}{(m-k)!}. \end{aligned}$$

Note that in the last line of the above equation we have used a result from Exercise 4.1, and the generalized distributivity from Section 5.6. Finally, note that $m!n = m!(m+1) = (m+1)! = n!$. In addition, we have $k+(m-k) = m$. So $k+1+(m-k) = m+1 = n$. Thus $m - k = n - (k+1)$. Therefore $|S(A;B)| = \frac{n!}{(n-(k+1))!}$; and hence $k+1 \in J$, as desired. ∎

Remark. Let f be a one-to-one function from A to B. We can consider f as an assignment of a spot in a list labeled by the elements of A to k of the elements of B. In other words, such functions take k elements from the n elements of the set B, and arrange them in a list. The above theorem says that there are $\frac{n!}{(n-k)!}$ ways to do this.

Definition 4.13. Let A be a set. A **permutation** on A is a one-to-one and onto function from A to A. Let $n \in \mathbb{N}$. We denote the set of all permutations on $\{1, \ldots, n\}$ by S_n.

Theorem 4.37. *Suppose A is a finite set, and $|A| = n$. Then the set of permutations on A is a finite set, and has $n!$ elements. In particular we have*

$$|S_n| = n!.$$

Remark. Suppose we have arranged the elements of A in a list. Then we can consider a permutation on A as a rearrangement of its elements. Thus this theorem says that there are $n!$ ways to arrange n objects in a list.

Proof. Note that a function from A to A is one-to-one if and only if it is onto, since A is a finite set. Hence the number of permutations on A is the same as the number of one-to-one functions from A to A, which by the previous theorem is equal to $\frac{n!}{(n-n)!} = \frac{n!}{0!} = \frac{n!}{1} = n!$. ∎

Theorem 4.38. *Suppose A is a finite set, and $|A| = n$. Let $k \leq n$. Then the number of subsets of A which have k elements is*

$$\frac{n!}{k!(n-k)!}.$$

Remark. Informally, this theorem says that there are $\frac{n!}{k!(n-k)!}$ ways to choose k objects among n objects. Note that unlike the previous two theorems, here we do not assign any order to the k objects that we choose.

Proof. We know that the number of subsets of a finite set is finite. So the number of subsets of A which have k elements is also finite. Suppose this number is l. Let $A_1, \ldots, A_l$ be the subsets of A that have k elements. Note that A has at least one subset with k elements, since if $h : n \to A$ is a bijective function, then the image of $h|_k$ has k elements. Let us denote the set of one-to-one functions from a set C to a set B by $S(C; B)$. We claim that

$$S(A_1; A) = \bigsqcup_{j \leq l} S(A_1; A_j).$$

Because if $f : A_1 \to A$ is one-to-one, then $f : A_1 \to f(A_1)$ is bijective. Hence $f(A_1)$ has k elements; so it must be equal to A_j for some j. On the other hand, every one-to-one function from A_1 to some A_j can also be considered as a one-to-one function from A_1 to A. Finally note that if $f \in S(A_1; A_i) \cap S(A_1; A_j)$ for $i \neq j$, then $f(A_1)$ is a subset of both A_i, A_j which has the same number of elements as A_i, A_j. Thus we must have $A_i = f(A_1) = A_j$, which is a contradiction. Hence $S(A_1; A_1), \ldots, S(A_1; A_l)$ are disjoint sets whose union is $S(A_1; A)$. Therefore we have

$$\frac{n!}{(n-k)!} = |S(A_1; A)| = \sum_{j \leq l} |S(A_1; A_j)| = \sum_{j \leq l} k! = k! \sum_{j \leq l} 1 = k!\, l.$$

Note that in the above equation we have used a result from Exercise 4.1, and the generalized distributivity from Section 5.6. Thus we obtain $k!(n-k)!\, l = n!$, which gives the desired formula for l. ∎

Definition 4.14. Let $n, k \in \omega$, and assume $0 \leq k \leq n$. The number

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

is called a **binomial coefficient**.

Remark. Note that by the above theorem, the binomial coefficients are natural numbers, because they are the number of elements of some finite sets.

Theorem 4.39. *Let $X, A_1, \ldots, A_m$ be sets, and for each $j \leq m$ let $f_j : A_j \to X$ be a function. Also suppose that $A_1, \ldots, A_m$ are pairwise disjoint, i.e. for $i \neq j$ we have $A_i \cap A_j = \emptyset$. Then $\bigcup_{j\leq m} f_j$ is a function from $\bigcup_{j\leq m} A_j$ to X.*

Proof. The proof is by induction on m. If $m = 1$ then $\bigcup_{j\leq 1} f_j = f_1$ is a function from $\bigcup_{j\leq 1} A_j = A_1$ to X. Now suppose the claim is true for some m. Consider the sequence $A_1, \ldots, A_m, A_{m+1}$. Then by induction hypothesis $\bigcup_{j\leq m} f_j$ is a function from $\bigcup_{j\leq m} A_j$ to X. Also note that

$$\begin{aligned} A_{m+1} \cap \Big(\bigcup_{j\leq m} A_j \Big) &= A_{m+1} \cap \big(\bigcup\{A_1, \ldots, A_m\}\big) \\ &= \bigcup\{A_{m+1} \cap A_1, \ldots, A_{m+1} \cap A_m\} = \bigcup\{\emptyset\} = \emptyset. \end{aligned}$$

Hence by Theorem 3.20,

$$\bigcup_{j\leq m+1} f_j = \Big(\bigcup_{j\leq m} f_j \Big) \cup f_{m+1}$$

is a function from $\bigcup_{j\leq m+1} A_j = (\bigcup_{j\leq m} A_j) \cup A_{m+1}$ to X. ■

Remark. Note that the functions $f_1, \ldots, f_m$ are also sets; so $\bigcup_{j\leq m} f_j$ in the above theorem is the union of the sets $f_1, \ldots, f_m$. Let us denote the function $\bigcup_{j\leq m} f_j$ by f, and let $x \in \bigcup_{j\leq m} A_j$. Then we use the following notation to describe f:

$$f(x) = \begin{cases} f_1(x) & x \in A_1, \\ \vdots & \\ f_m(x) & x \in A_m. \end{cases}$$

We usually use the above notation to define f, instead of saying that $f := \bigcup_{j\leq m} f_j$. And we say that f is a *piecewise-defined function.*

Definition 4.15. Let $A_k, \ldots, A_m$ be a finite sequence of sets. Then

$$\begin{aligned} &A_k \times \cdots \times A_m \\ &\quad := \{f : \{k, \ldots, m\} \to \bigcup_{k\leq j\leq m} A_j : \text{ such that } f(j) \in A_j \text{ for every } k \leq j \leq m\}. \end{aligned}$$

We also denote $A_k \times \cdots \times A_m$ by $\prod_{j=k}^{m} A_j$, or $\prod_{k\leq j\leq m} A_j$. When k is 0 or 1, we may also use the notation $\prod_{j\leq m} A_j$.

Remark. Thus, the Cartesian product $A_k \times \cdots \times A_m$ is the set of all functions from $\{k, \ldots, m\}$ to $\bigcup_{k\leq j\leq m} A_j$, such that $f(j)$ is an element of A_j. We usually denote f by

$$(a_k, \ldots, a_m),$$

where $a_j = f(j)$. Note that $a_j \in A_j$ for each j. Also note that we have

$$(a_k, \ldots, a_m) = (b_k, \ldots, b_m)$$

if and only if $a_j = b_j$ for every j. The reason is that two functions are equal if and only if they have the same domain, and they have the same value at every point of their common domain.

Definition 4.16. Let $A_1, \ldots, A_m$ be a finite sequence of sets. Then the elements of $A_1 \times \cdots \times A_m$ are called ***m*-tuples**. In the m-tuple $(a_1, \ldots, a_m) \in \prod_{i \leq m} A_i$, a_j is called the jth **component**.

Remark. Note that when $m = 2$, we have two definitions for $A_1 \times A_2$. One of them is the set of all ordered pairs (a_1, a_2), where $a_j \in A_j$; and the other one is the set of all functions $f : \{1, 2\} \to A_1 \cup A_2$ such that $f(j) \in A_j$. Although the two sets are not equal, they essentially look like the same. More formally, there exists a one-to-one correspondence between them, which can be defined naturally as follows

$$f \mapsto (f(1), f(2)).$$

As we mentioned several times so far, the inherent nature of objects is not mathematically as important as their characteristic properties. And the characteristic property of ordered pairs, or 2-tuples, is that they store the information of order of two elements.

Remark. We could have also defined $A_1 \times \cdots \times A_m$ inductively as

$$A_1 \times \cdots \times A_m = (A_1 \times \cdots \times A_{m-1}) \times A_m,$$

where the rightmost product with A_m is the standard Cartesian product of two sets. However, similarly to the case of $m = 2$, this definition is essentially the same as the definition in terms of functions. In addition, the definition in terms of functions has a simpler structure, and can be generalized to the case of infinitely many sets.

Notation. When $A_1 = \cdots = A_m = X$, we denote $A_1 \times \cdots \times A_m$ by X^m.

Remark. Note that the elements of X^m are finite sequences in X of length m.

The next theorem is a sort of generalized associativity property for Cartesian product.

Theorem 4.40. *Let $X_k, \ldots, X_m$ be a finite sequence of sets, and suppose $k \leq l < m$. Then the map*

$$\begin{aligned} (X_k \times \cdots \times X_l) \times (X_{l+1} \times \cdots \times X_m) &\to X_k \times \cdots \times X_m \\ \big((x_k, \ldots, x_l)\ ,\ (x_{l+1}, \ldots, x_m)\big) &\mapsto (x_k, \ldots, x_m) \end{aligned}$$

is a bijective function.

Proof. It is obvious that every $(x_k, \dots, x_m)$ in $X_k \times \cdots \times X_m$ is in the image of the above map. Also, if $(x_k, \dots, x_m) = (y_k, \dots, y_m)$ then $x_j = y_j$ for all j. Hence $(x_k, \dots, x_l) = (y_k, \dots, y_l)$, and $(x_{l+1}, \dots, x_m) = (y_{l+1}, \dots, y_m)$. Thus

$$\big((x_k, \dots, x_l)\ ,\ (x_{l+1}, \dots, x_m)\big) = \big((y_k, \dots, y_l)\ ,\ (y_{l+1}, \dots, y_m)\big).$$

So the map is also one-to-one. ■

Theorem 4.41. *Suppose $A_1, \dots, A_m$ is a finite sequence of sets. Also suppose that each A_j is a finite set. Then $A_1 \times \cdots \times A_m$ is also a finite set, and we have*

$$|A_1 \times \cdots \times A_m| = |A_1| \cdots |A_m|.$$

As a result, for every finite set X we have

$$|X^m| = |X|^m.$$

Proof. The proof is by induction on m. If $m = 1$ then $\prod_{j \le 1} A_j$ is the set of functions from $\{1\}$ to A_1, which we know is a finite set, and has $|A_1|^1 = |A_1|$ elements. Now suppose the claim is true for some m. Consider the sequence $A_1, \dots, A_m, A_{m+1}$. Then by induction hypothesis $\prod_{j \le m} A_j$ is a finite set. Now the previous theorem implies that there is a bijective function from $\prod_{j \le m+1} A_j$ to $(\prod_{j \le m} A_j) \times A_{m+1}$, where we can assume that $\times$ denotes the Cartesian product as defined before in Section 3.1. Hence by multiplication principle, and the induction hypothesis, $\prod_{j \le m+1} A_j$ is a finite set, and we have

$$\Big|\prod_{j \le m+1} A_j\Big| = \Big|\prod_{j \le m} A_j\Big| \cdot |A_{m+1}| = \Big(\prod_{j \le m} |A_j|\Big) \cdot |A_{m+1}| = \prod_{j \le m+1} |A_j|,$$

as desired. The result for $|X^m|$ follows from the fact that $\prod_{j \le m} |X| = |X|^m$, as shown in Theorem 5.40. ■

Definition 4.17. Let $A_1, \dots, A_m$ be a finite sequence of sets. The function

$$\begin{aligned} A_1 \times \cdots \times A_m &\to A_j \\ (a_1, \dots, a_m) &\mapsto a_j \end{aligned}$$

is called the **projection** on the jth component.

Remark. More formally, the projection on the jth component of $A_1 \times \cdots \times A_m$ is the function

$$\{(f, f(j)) : f \in A_1 \times \cdots \times A_m\}.$$

Note that for every $f \in A_1 \times \cdots \times A_m$, the value $f(j)$ is unique, since f is a function; therefore the above relation is indeed a function.

Definition 4.18. Let $X, A_1, \ldots, A_m$ be sets, and let $f : X \to A_1 \times \cdots \times A_m$ be a function. Let π_j be the projection on the jth component of the product $A_1 \times \cdots \times A_m$. Then the jth **component** of f is the function

$$f_j := \pi_j \circ f : X \to A_j.$$

We usually write $f = (f_1, \ldots, f_m)$ to say that $f_1, \ldots, f_m$ are the components of f.

Theorem 4.42. *Let $X, A_1, \ldots, A_m$ be sets, and for each $j \leq m$ let $f_j : X \to A_j$ be a function. Then there is a unique function $f : X \to A_1 \times \cdots \times A_m$ whose jth component is f_j, for each $j \leq m$. In other words, for every $x \in X$ we have*

$$f(x) = (f_1(x), \ldots, f_m(x)).$$

Proof. This follows immediately from Theorem 3.9, since $(f_1(x), \ldots, f_m(x))$ is uniquely determined by x. ■

Remark. A particular case of the above theorem is when $X = A_j$ for some j, $f_j : A_j \to A_j$ is the identity map, and f_i is a constant function with value a_i for $i \neq j$. Then we get a function from A_j to $A_1 \times \cdots \times A_m$ defined by

$$x \mapsto (a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_m)$$

Notation. Let $X, A_1, \ldots, A_m$ be sets, and let $f : A_1 \times \cdots \times A_m \to X$ be a function. Then for $a = (a_1, \ldots, a_m) \in A_1 \times \cdots \times A_m$ we usually denote $f(a)$ by $f(a_1, \ldots, a_m)$.

Theorem 4.43. *Let $X_1, \ldots, X_m, A_1, \ldots, A_m$ be sets, and for each $j \leq m$ let $f_j : X_j \to A_j$ be a function. Then there is a unique function*

$$f : X_1 \times \cdots \times X_m \to A_1 \times \cdots \times A_m$$

such that for every $(x_1, \ldots, x_m) \in \prod_{j \leq m} X_j$ we have

$$f(x_1, \ldots, x_m) = (f_1(x_1), \ldots, f_m(x_m)).$$

Proof. This follows immediately from Theorem 3.9, since $(f_1(x_1), \ldots, f_m(x_m))$ is uniquely determined by $x = (x_1, \ldots, x_m)$. ■

Finally, let us prove a more general version of recursion theorem that allows us to define functions whose values at $n + l$ depend on their values at $n + l - 1, \ldots, n$.

Theorem 4.44. *Let $l, k \in \omega$ with $l > 1$, and*

$$J = \{n \in \omega : n \geq k\}.$$

Let A be a set, and suppose $a_1, \ldots, a_l \in A$ (these elements need not be distinct). Suppose $F : J \times A^l \to A$ is a function. Then there exists a unique function $f : J \to A$ such that

(i) $f(k+i) = a_{i+1}$ *for* $0 \le i < l$, *and*
(ii) $f(n+l) = F\big(n, f(n), \dots, f(n+l-1)\big)$ *for every* $n \ge k$.

Proof. Consider the function $G : J \times A^l \to A^l$ given by

$$G(n, x_1, x_2, \dots, x_l) := \big(x_2, \dots, x_l, F(n, x_1, \dots, x_l)\big).$$

Then we know that there is a unique function $g : J \to A^l$ such that

$$\begin{cases} g(k) = (a_1, \dots, a_l), \\ g(n+1) = G(n, g(n)). \end{cases}$$

If we look at the components of g we get

$$\begin{cases} g_i(k) = a_i & 1 \le i \le l, \\ g_i(n+1) = g_{i+1}(n) & 1 \le i < l, \\ g_l(n+1) = F\big(n, g_1(n), \dots, g_l(n)\big). \end{cases}$$

We claim that for $i < l$ we have $g_1(n+i) = g_{i+1}(n)$. This is trivial for $i = 0$. And if it holds for i then for $i+1$ we have

$$g_1(n+i+1) = g_1(n+1+i) = g_{i+1}(n+1) = g_{i+2}(n).$$

So if we set $f = g_1$ then for $i < l$ we have

$$f(k+i) = g_1(k+i) = g_{i+1}(k) = a_{i+1}.$$

We also have

$$\begin{aligned} f(n+l) &= g_1(n+l) = g_1(n+1+l-1) = g_{l-1+1}(n+1) \\ &= g_l(n+1) = F\big(n, g_1(n), g_2(n), \dots, g_l(n)\big) \\ &= F\big(n, g_1(n), g_1(n+1), \dots, g_1(n+l-1)\big) \\ &= F\big(n, f(n), f(n+1), \dots, f(n+l-1)\big), \end{aligned}$$

as desired.

To prove the uniqueness, suppose $\tilde{f}$ also satisfies the properties stated in the theorem. For $1 \le i \le l$ we set $\tilde{g}_i(n) := \tilde{f}(n+i-1)$. Let $\tilde{g} = (\tilde{g}_1, \dots, \tilde{g}_l)$. Then we have

$$\tilde{g}(k) = \big(\tilde{g}_1(k), \dots, \tilde{g}_l(k)\big) = \big(\tilde{f}(k), \dots, \tilde{f}(k+l-1)\big) = (a_1, \dots, a_l).$$

In addition, $\tilde{g}_i(n+1) = \tilde{f}(n+1+i-1) = \tilde{g}_{i+1}(n)$ for $1 \le i < l$, and

$$\begin{aligned} \tilde{g}_l(n+1) &= \tilde{f}(n+1+l-1) = \tilde{f}(n+l) \\ &= F\big(n, \tilde{f}(n), \dots, \tilde{f}(n+l-1)\big) = F\big(n, \tilde{g}_1(n), \dots, \tilde{g}_l(n)\big). \end{aligned}$$

Hence $\tilde{g}(n+1) = G(n, \tilde{g}(n))$. But due to the uniqueness of g we must have $\tilde{g} = g$, and therefore $\tilde{f} = \tilde{g}_1 = g_1 = f$, as wanted. ∎

***Example* 4.8.** Let us see an application of the previous theorem. Suppose $J = A = \omega$ and $F : J \times A^2 \to A$ maps (n, a, b) to $a + b$. Then we can construct the sequence given by $a_0 = 0$, $a_1 = 1$, and

$$a_{n+2} = a_{n+1} + a_n$$

for $n \geq 0$. This sequence is known as the *Fibonacci sequence*. The first few terms of the sequence are $0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots$.

Let us show that for every n we have $a_n < 2^n$. We prove this by strong induction on n. Let $A = \{n : a_n < 2^n\}$. We must show that if for every natural number $m < n$ we have $m \in A$, then $n \in A$. Suppose for every natural number $m < n$ we have $m \in A$, i.e. $a_m < 2^m$. If $n \geq 2$ then $m = n - 2$ and $m + 1 = n - 1$ are natural numbers less than n, so we have $a_m < 2^m$ and $a_{m+1} < 2^{m+1}$. Hence

$$\begin{aligned} a_n = a_{m+2} = a_{m+1} + a_m &< 2^{m+1} + a_m < 2^{m+1} + 2^m \\ &< 2^{m+1} + 2^{m+1} = 2 \cdot 2^{m+1} = 2^{m+2} = 2^n. \end{aligned}$$

(The fact that $2^m < 2^{m+1}$ can be proved easily by an induction.) However note that this argument only works when $n \geq 2$; thus we need to check the cases $n < 2$, i.e. $n = 0, 1$, separately. In other words, we need to check two base cases: for $n = 0$ we have $a_0 = 0 < 1 = 2^0$, and for $n = 1$ we have $a_1 = 1 < 2 = 2^1$.

Chapter 5

Integers and Rational Numbers

5.1 Integers

The next step in constructing numbers, is to build integers. As we discussed in the beginning of the previous chapter, the inherent nature of numbers is not that important for us. So, similarly to the case of natural numbers, we just need to find a set whose elements represent the integers, and have the properties that we expect from integers. The basic idea in our construction is to consider an integer as the difference of two natural numbers. For example, we can consider -1 as $1-2$, or $6-7$. Since we do not have a notion of subtraction of natural numbers yet, we can store the information of the difference $1-2$ in the ordered pair $(1,2)$. So we can represent -1 by the pair $(1,2)$. However, there are many other ordered pairs that can represent -1, like $(6,7)$, or $(0,1)$. To overcome this ambiguity, we can identify all ordered pairs of natural numbers that represent the same integer, by using an equivalence relation.

Definition 5.1. We define the relation $\sim$ on $\omega \times \omega$ as follows

$$(m,n) \sim (k,l) \qquad \text{if} \qquad m+l = k+n.$$

Remark. Note that informally, $m+l=k+n$ is equivalent to $m-n=k-l$. Thus the two pairs (m,n) and (k,l) are related by $\sim$, if and only if they represent the same difference of natural numbers, i.e. they represent the same integer.

Theorem 5.1. *The relation $\sim$ is an equivalence relation on $\omega \times \omega$.*

Proof. First note that $(m,n) \sim (m,n)$, since $m+n=m+n$. Also if $(m,n) \sim (k,l)$ then $m+l=k+n$; hence we have $k+n=m+l$, which means $(k,l) \sim (m,n)$. So far we have shown that $\sim$ is reflexive and symmetric. Let us show that it is transitive too. Suppose $(m,n) \sim (k,l)$ and $(k,l) \sim (i,j)$. Then we have

$$m+l=k+n, \qquad k+j=i+l.$$

If we add these equalities we get

$$m + l + k + j = k + n + i + l.$$

Note that the sum of equal numbers are equal, since addition is a function. Now by using the associativity and commutativity of addition of ω, we can rearrange the terms in the above equation to obtain

$$m + j + k + l = i + n + k + l.$$

Thus by the cancellation law we get $m + j = i + n$, which means $(m, n) \sim (i, j)$, as desired. ■

Definition 5.2. The set of **integers** is

$$\mathbb{Z} := \omega \times \omega / \sim,$$

i.e. $\mathbb{Z}$ is the set of all equivalence classes of $\omega \times \omega$ under the equivalence relation $\sim$.

We denote the equivalence class of (m, n) by $[(m, n)]$. So, informally, $[(m, n)]$ is the integer $m - n$. Note that all the other pairs in the equivalence class $[(m, n)]$ represent the same difference, i.e. the same integer.

Next, we have to define the addition and multiplication of integers, and show that they have the expected properties. To this end, we need the following version of Theorem 3.21.

Theorem 5.2. *Suppose R, S are equivalence relations on the sets X, Y respectively. Also suppose that $F : X \times X \to Y$ is a function, and for every $a, b, c, d \in X$ we have*

$$aRc \text{ and } bRd \qquad \Longrightarrow \qquad F(a, b)SF(c, d).$$

Then there is a unique function $f : X/R \times X/R \to Y/S$ that satisfies

$$f([a]_R, [b]_R) = [F(a, b)]_S$$

for every $a, b \in X$.

Remark. In the following proof we define $f([a]_R, [b]_R) := [F(a, b)]_S$, and we will show that f is a function. Similarly to the case of Theorem 3.21, when we define a function in this way, and the required conditions are satisfied, we say that the function is **well defined**.

Proof. First note that by Theorem 3.8 there is at most one function whose domain is $X/R \times X/R$, and maps $([a]_R, [b]_R)$ to $[F(a, b)]_S$. Now let us define $f([a]_R, [b]_R) := [F(a, b)]_S$. More explicitly, this means that

$$f := \{((x, y), z) \in (X/R \times X/R) \times Y/S : \exists c \in x \ \exists d \in y \text{ such that } F(c, d) \in z\}.$$

Note that the domain of f is all of $X/R \times X/R$, since to every $([a]_R, [b]_R) \in X/R \times X/R$ we can at least assign one value $[F(a,b)]_S$. Now let us show that this value is the only value that is assigned to $([a]_R, [b]_R)$. Suppose $(([a]_R, [b]_R), z) \in f$. Then by definition there is $c \in [a]_R$ and $d \in [b]_R$ such that $F(c,d) \in z$. But $c \in [a]_R$ means that aRc, and $d \in [b]_R$ means that bRd. Hence by our assumption we have $F(a,b)SF(c,d)$. Thus $F(c,d) \in [F(a,b)]_S$. Therefore the two equivalence classes $z, [F(a,b)]_S$ have a nonempty intersection; so they must be equal, i.e. $z = [F(a,b)]_S$. Thus the value of f at $([a]_R, [b]_R)$ is uniquely determined, as desired. Hence f is a function. ■

Now let $[(m,n)], [(k,l)] \in \mathbb{Z}$. Informally, these numbers represent $m-n$ and $k-l$, respectively. Thus if we want to add and multiply them, the results should be

$$\begin{aligned}(m-n)+(k-l) &= (m+k)-(n+l),\\ (m-n)(k-l) &= (mk+nl)-(ml+nk).\end{aligned}$$

The following theorem and Theorem 5.2 show that if we define addition and multiplication as above, they are well defined.

Theorem 5.3. *Let* $(m,n),(m',n'),(k,l),(k',l') \in \omega\times\omega$. *Suppose* $(m,n) \sim (m',n')$, *and* $(k,l) \sim (k',l')$. *Then we have*

(i) $(m+k, n+l) \sim (m'+k', n'+l')$.
(ii) $(mk+nl, ml+nk) \sim (m'k'+n'l', m'l'+n'k')$.
(iii) $(n,m) \sim (n',m')$.

Proof. We know that

$$m+n' = m'+n, \qquad k+l' = k'+l. \tag{$*$}$$

(i) By adding the above two equations we get

$$m+n'+k+l' = m'+n+k'+l.$$

Now we can rearrange the terms in the above equation to obtain

$$m+k+n'+l' = m'+k'+n+l,$$

as desired.

(ii) If we multiply the first equation of $(*)$ by k and its second equation by m', and add the resulting equations, we get

$$mk+n'k+m'k+m'l' = m'k+nk+m'k'+m'l.$$

Next, if we multiply the first equation of $(*)$ by l and its second equation by n', and add the resulting equations, we get

$$ml + n'l + n'k + n'l' = m'l + nl + n'k' + n'l.$$

Now if we reverse this equation and add it to the previous equation, we obtain

$$\begin{aligned} &mk + n'k + m'k + m'l' + m'l + nl + n'k' + n'l \\ &\quad = ml + n'l + n'k + n'l' + m'k + nk + m'k' + m'l. \end{aligned}$$

Finally, by cancelling equal terms, and rearranging the remaining terms we get

$$mk + nl + m'l' + n'k' = m'k' + n'l' + ml + nk,$$

which is the desired equation.

(iii) We have $n + m' = m' + n = m + n' = n' + m$. ■

Definition 5.3. Let $[(m, n)], [(k, l)] \in \mathbb{Z}$. Then we define the **addition** and **multiplication** of integers as

$$\begin{aligned} [(m, n)] + [(k, l)] &:= [(m + k, n + l)], \\ [(m, n)][(k, l)] &:= [(mk + nl, ml + nk)], \end{aligned}$$

respectively. The **zero** and **identity** of $\mathbb{Z}$ are

$$\begin{aligned} 0 &:= [(0, 0)], \\ 1 &:= [(1, 0)], \end{aligned}$$

respectively. The **opposite** of $[(m, n)]$ is

$$-[(m, n)] := [(n, m)].$$

Remark. It is easy to see that in $\mathbb{Z}$ we have $0 \neq 1$, because in ω we have

$$1 + 0 = 1 \neq 0 = 0 + 0.$$

Also, note that by Theorem 3.21 and the above theorem, the opposite of integers is well defined.

Theorem 5.4. *Suppose $a, b, c \in \mathbb{Z}$. Then we have*

(i) *Associativity* :

$$a + (b + c) = (a + b) + c, \qquad a(bc) = (ab)c.$$

(ii) *Commutativity* :

$$a+b=b+a, \qquad ab=ba.$$

(iii) *Identity elements* :

$$a+0=a, \qquad a1=a.$$

(iv) *Additive inverse* :

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

(v) *Distributivity* :

$$a(b+c)=ab+ac.$$

Proof. Suppose $a=[(m,n)]$, $b=[(k,l)]$, and $c=[(i,j)]$.

(i) We have

$$\begin{aligned} a+(b+c) &= \big[\big(m+(k+i),n+(l+j)\big)\big] \\ &= \big[\big((m+k)+i,(n+l)+j\big)\big] = (a+b)+c. \end{aligned}$$

We also have

$$\begin{aligned} a(bc) &= \big[\big(m,n\big)\big]\big[\big(ki+lj,kj+li\big)\big] \\ &= \big[\big(m(ki+lj)+n(kj+li),m(kj+li)+n(ki+lj)\big)\big] \\ &= \big[\big((mk+nl)i+(ml+nk)j,(mk+nl)j+(ml+nk)i\big)\big] \\ &= \big[\big(mk+nl,ml+nk\big)\big]\big[\big(i,j\big)\big] = (ab)c. \end{aligned}$$

(ii) We have

$$a+b=[(m+k,n+l)]=[(k+m,l+n)]=b+a.$$

We also have

$$ab=[(mk+nl,ml+nk)]=[(km+ln,kn+lm)]=ba.$$

(iii) We have $a+0=[(m+0,n+0)]=[(m,n)]=a$. Also

$$a1=[(m1+n0,m0+n1)]=[(m,n)]=a.$$

(iv) We have

$$a+(-a)=[(m+n,n+m)]=[(m+n,m+n)]=[(0,0)]=0.$$

Note that $(m+n,m+n)\sim(0,0)$, since $m+n+0=m+n+0$.

(v) We have

$$\begin{aligned} a(b+c) &= \big[\big(m,n\big)\big]\big[\big(k+i,l+j\big)\big] \\ &= \big[\big(m(k+i)+n(l+j),m(l+j)+n(k+i)\big)\big] \\ &= \big[\big((mk+nl)+(mi+nj),(ml+nk)+(mj+ni)\big)\big] = ab+ac. \end{aligned}$$

■

Theorem 5.5. *Let $a, b \in \mathbb{Z}$. Then $ab = 0$ implies that either $a = 0$, or $b = 0$.*

Proof. Suppose $a = [(m, n)]$ and $b = [(k, l)]$. Then $ab = [(mk + nl, ml + nk)]$. We will prove the contrapositive of the theorem, i.e. we will show that if $a \neq 0$ and $b \neq 0$, then $ab \neq 0$. Since $(m, n) \nsim (0, 0)$ and $(k, l) \nsim (0, 0)$, we must have $m \neq n$ and $k \neq l$. Thus either $m < n$, or $m > n$. Similarly, either $k < l$, or $k > l$. Suppose $m < n$ and $k < l$. Then there are $i, j > 0$ such that $n = m + i$, and $l = k + j$. Hence we have

$$\begin{aligned} mk + nl &= mk + (m + i)(k + j) \\ &= m(k + j) + (m + i)k + ij = ml + nk + ij > ml + nk. \end{aligned}$$

Next suppose $m < n$ and $k > l$. Then there are $i, j > 0$ such that $n = m + i$, and $k = l + j$. Hence we have

$$\begin{aligned} ml + nk &= ml + (m + i)(l + j) \\ &= m(l + j) + (m + i)l + ij = mk + nl + ij > mk + nl. \end{aligned}$$

The case of $m > n$ and $k < l$ is similar. Finally, suppose $m > n$ and $k > l$. Then there are $i, j > 0$ such that $m = n + i$, and $k = l + j$. Hence we have

$$\begin{aligned} mk + nl &= (n + i)(l + j) + nl \\ &= (n + i)l + n(l + j) + ij = ml + nk + ij > ml + nk. \end{aligned}$$

Therefore, in every case we can show that $mk + nl \neq ml + nk$. Thus we have $(mk + nl, ml + nk) \nsim (0, 0)$; so $ab \neq 0$, as desired. ■

Definition 5.4. Let $a, b \in \mathbb{Z}$. Then we define the **subtraction** of integers as follows

$$a - b := a + (-b).$$

Next, we want to define the order relation on integers. Let $[(m, n)], [(k, l)] \in \mathbb{Z}$. Informally, these numbers represent $m - n$ and $k - l$, respectively. Thus if we want $[(m, n)]$ to be less than $[(k, l)]$, we must have $m - n < k - l$, or equivalently $m + l < k + n$.

Definition 5.5. Let $a, b \in \mathbb{Z}$. Then we say $a < b$ if there are $(m, n) \in a$ and $(k, l) \in b$ such that

$$m + l < k + n.$$

Note that every integer is a set of pairs of natural numbers, i.e. it is the equivalence class of pairs of natural numbers which represent it. Thus the above definition makes sense.

Also, note that $<$ is a relation on $\mathbb{Z}$; and unlike the case of functions, we do not need to check any condition, or use any theorem to show this. In other words, we do not need to check that $<$ is well defined, in order to be sure that it is indeed a relation. Because a relation is merely a set of ordered pairs, and does not have any particular property. So we do not need to make sure that some property holds, in order to show that $<$ is a relation. But, a priori it is not obvious that $<$ is a relation with desirable properties. For example, there might also be some $(m', n') \in a$ and $(k', l') \in b$ such that $m' + l' > k' + n'$; so we might also have $a > b$. However, the next theorem shows that this cannot happen.

Theorem 5.6. *Let $(m, n), (m', n'), (k, l), (k', l') \in \omega \times \omega$. Suppose $(m, n) \sim (m', n')$, and $(k, l) \sim (k', l')$. Then we have*

$$m + l < k + n \qquad \iff \qquad m' + l' < k' + n'.$$

Proof. We know that

$$m + n' = m' + n, \qquad k + l' = k' + l.$$

Suppose $m + l < k + n$. Then we have

$$m + l + n' + l' < k + n + n' + l'.$$

Hence by rearranging the terms and applying the above equations we get

$$m' + n + l + l' < k' + l + n + n'.$$

Thus $m' + l' + n + l < k' + n' + n + l$. This implies that $m' + l' < k' + n'$. Because otherwise we would have $m' + l' \geq k' + n'$. However, then we would get $m' + l' + n + l \geq k' + n' + n + l$, which is a contradiction. Hence we must have $m' + l' < k' + n'$ as desired. The reverse implication follows similarly. ■

Theorem 5.7. *The relation $<$ on $\mathbb{Z}$ is a linear order. In addition, $\mathbb{Z}$ does not have a smallest element, nor a largest element.*

Proof. Let $a = [(m, n)]$, $b = [(k, l)]$, and $c = [(i, j)]$. First note that $a \not< a$, since $m + n \not< m + n$. So $<$ is irreflexive. Now suppose $a < b$ and $b < c$. Then we have

$$m + l < k + n, \qquad k + j < i + l.$$

Hence we have

$$m + l + k + j < k + n + k + j < k + n + i + l.$$

If we cancel $k + l$ from both sides we get $m + j < n + i$, which means $a < c$. So $<$ is also transitive.

Next suppose $a \neq b$. Then $m + l \neq k + n$. Then since the order of natural numbers is a total order, we either have $m + l < k + n$, or $m + l > k + n$. Hence either $a < b$, or $a > b$. So $<$ is a total order on $\mathbb{Z}$. Finally, it is easy to see that we always have

$$[(m, n + 1)] < [(m, n)] < [(m + 1, n)].$$

Therefore no integer like $a = [(m, n)]$ can be the smallest element, nor the largest element of $\mathbb{Z}$. ■

Definition 5.6. Let $a \in \mathbb{Z}$. We say a is **positive** if $a > 0$, and a is **negative** if $a < 0$. We also say a is *nonnegative* or *nonpositive*, if $a \geq 0$ or $a \leq 0$ respectively. The **sign** of a is its property of being positive, negative, or zero.

Remark. Note that by the trichotomy law, a is either positive, zero, or negative; and exactly one of these cases occurs.

Theorem 5.8. *Let $a, b, c \in \mathbb{Z}$. Then we have*
(i) *If $a < b$ then $a + c < b + c$.*
(ii) *If $c > 0$ and $a < b$, then $ac < bc$.*

Remark. As a consequence, note that if $a, b > 0$ then $ab > 0$. Because we have $ab > a0 = 0$.

Proof. Suppose $a = [(m, n)]$, $b = [(k, l)]$, and $c = [(i, j)]$. Also suppose $a < b$, so we have $m + l < k + n$. Hence there is $p > 0$ such that $k + n = m + l + p$.

(i) We have $a + c = [(m + i, n + j)]$, and $b + c = [(k + i, l + j)]$. Now we have

$$m + i + l + j = m + l + i + j < k + n + i + j = k + i + n + j.$$

Thus $a + c < b + c$, as desired.

(ii) We have $ac = [(mi + nj, mj + ni)]$, and $bc = [(ki + lj, kj + li)]$. We also know that $c > 0$; so $i > j$. Then there is $q > 0$ such that $i = j + q$. Now we have

$$\begin{aligned} ki + lj + mj + ni &= (k + n)i + (m + l)j \\ &= (m + l + p)(j + q) + (m + l)j \\ &= (m + l)(j + q) + (m + l + p)j + pq \\ &= (m + l)i + (k + n)j + pq > mi + nj + kj + li. \end{aligned}$$

Hence $ac < bc$, as desired. ■

Theorem 5.9. *Let $a, b \in \mathbb{Z}$. Then $a > b$ if and only if $-a < -b$.*

Remark. As a consequence we have $a < 0$ if and only if $-a > 0$; since $-0 = 0$ by the definition of 0.

Proof. Suppose $a = [(m, n)]$, and $b = [(k, l)]$. Then we have $-a = [(n, m)]$, and $-b = [(l, k)]$. Now $a > b$ means $m + l > k + n$. On the other hand, $-a < -b$ means $n + k < l + m$; which is the same as $m + l > k + n$. Thus $a > b$ is equivalent to $-a < -b$. ■

Cancellation Laws. *Let $a, b, c \in \mathbb{Z}$. Then we have*

(i) *If $a + c = b + c$ then $a = b$.*

(ii) *If $c \neq 0$ and $ac = bc$, then $a = b$.*

Proof. Suppose $a = [(m, n)]$, $b = [(k, l)]$, and $c = [(i, j)]$.

(i) We have $a + c = [(m + i, n + j)]$, and $b + c = [(k + i, l + j)]$. So $a + c = b + c$ means

$$m + i + l + j = k + i + n + j.$$

Now if we cancel $i + j$ from both sides we get $m + l = k + n$, which means $a = b$.

(ii) Since $c \neq 0$ we either have $c > 0$, or $c < 0$. Suppose to the contrary that $a \neq b$. Then either $a < b$, or $a > b$. Suppose $a < b$. Now if $c > 0$ then we get $ac < bc$, which contradicts our assumption. Also, if $c < 0$ we have $-c > 0$. So we get $a(-c) < b(-c)$. But by Proposition 5.2 we have $a(-c) = -ac$, and $b(-c) = -bc$. Thus we have shown that $-ac < -bc$, which is equivalent to $ac > bc$; and this contradicts our assumption too. Therefore we cannot have $a < b$. Similarly, we can show that we cannot have $a > b$. Hence we must have $a = b$, as desired. ■

Theorem 5.10. *For any $a \in \mathbb{Z}$ there is no $b \in \mathbb{Z}$ such that*

$$a < b < a + 1.$$

Hence for every $b \in \mathbb{Z}$ we have

$$\begin{aligned} b > a \quad &\iff \quad b \geq a + 1, \\ b \leq a \quad &\iff \quad b < a + 1. \end{aligned}$$

Proof. Suppose $a = [(m, n)]$, and $b = [(k, l)]$. Then $a+1 = [(m+1, n)]$. Suppose $a < b$. Then we have $m+l < k+n$. Then by Theorem 4.16 we have $m+l+1 \leq k+n$, which means that $a + 1 \leq b$. Hence we have $b \not< a + 1$, as desired. The other conclusions of the theorem follow easily. We have shown that if $b > a$ then $b \geq a+1$. Conversely, if $b \geq a + 1$ then $b > a$, since $a + 1 > a$. Also, if $b \leq a$ then $b < a + 1$, since $a < a + 1$. And if $b < a + 1$ then $b \leq a$, because we must have $a \not< b$. ■

Although $\mathbb{Z}$ does not contain ω as a subset, it has a subset which looks like ω. Informally, we can say that $\mathbb{Z}$ contains a "copy" of ω. This copy of ω is the set

$$\{[(n, 0)] \in \mathbb{Z} : n \in \omega\}.$$

The next theorem shows that the above subset of $\mathbb{Z}$ behaves similarly to ω.

Theorem 5.11. *Let $E : \omega \to \mathbb{Z}$ be defined as $E(n) = [(n, 0)]$, for every $n \in \omega$. Then E is a one-to-one function whose image is the set of nonnegative integers, and for every $n, m \in \omega$ we have*

(i) $E(n + m) = E(n) + E(m)$.
(ii) $E(nm) = E(n)E(m)$.
(iii) $n < m$ *if and only if* $E(n) < E(m)$.
(iv) $[(n, m)] = E(n) - E(m)$.

Remark. Note that in the relation $E(n + m) = E(n) + E(m)$, n, m are added using the addition of ω, and $E(n), E(m)$ are added using the addition of $\mathbb{Z}$. So in some sense, we can say that the function E transforms the addition of ω into the addition of $\mathbb{Z}$. Similar remarks apply to the multiplication and order relation.

Remark. Also note that the last part of the theorem confirms our initial intuition that $[(n, m)]$ represents the integer which is the difference of the natural numbers n, m. Keep in mind that we identify $E(n)$ with n.

Proof. First note that E is one-to-one. Because if $E(n) = E(m)$ then we have $[(n, 0)] = [(m, 0)]$, which means that $n + 0 = m + 0$. So we get $n = m$. Next note that we always have $E(n) \geq 0$, since $n + 0 \geq 0 + 0$. Also, if $[(k, l)] > 0$ then we have $k > l$. Hence there is $n > 0$ such that $k = l + n$. Thus we get $(k, l) \sim (n, 0)$, i.e. $[(k, l)] = [(n, 0)] = E(n)$. Therefore the image of E is the set of nonnegative integers.

(i) We have

$$\begin{aligned} E(n) + E(m) &= [(n, 0)] + [(m, 0)] \\ &= [(n + m, 0 + 0)] = [(n + m, 0)] = E(n + m). \end{aligned}$$

(ii) We have

$$E(n)E(m) = [(n, 0)][(m, 0)] = [(nm + 0 \times 0, n0 + 0m)] = [(nm, 0)] = E(nm).$$

(iii) If $n < m$ then $E(n) < E(m)$, since $n + 0 < m + 0$. Conversely, if $E(n) < E(m)$ then by definition we must have $n + 0 < m + 0$; so $n < m$.

(iv) We have

$$\begin{aligned} [(n, m)] &= [(n, 0)] + [(0, m)] \\ &= [(n, 0)] + \big(-[(m, 0)]\big) = E(n) + \big(-E(m)\big) = E(n) - E(m). \end{aligned}$$

■

The well-ordering of ω has the following consequence for $\mathbb{Z}$.

Theorem 5.12. *Let $A \subset \mathbb{Z}$, and suppose A is nonempty.*

(i) *If A is bounded below then it has a least element.*
(ii) *If A is bounded above then it has a largest element.*

Proof. **(i)** Suppose c is a lower bound for A. Then for every $a \in A$ we have $a \geq c$. Hence $a - c \geq 0$. Consider the set

$$B := \{a - c : a \in A\}.$$

Then B is a subset of the set of nonnegative integers, i.e. B is a subset of the image of E. Now $E^{-1}(B)$ is a nonempty subset of ω, since A and therefore B are nonempty. Let k be the least element of $E^{-1}(B)$. We claim that $E(k) + c$ is the least element of A. Note that $E(k) \in B$, so $E(k) + c \in A$. Let $a \in A$. Then $a - c \in B$. Thus there is $n \in E^{-1}(B)$ such that $E(n) = a - c$; because B is a subset of the image of E. Now we have $k \leq n$. Hence $E(k) \leq E(n) = a - c$. Therefore $E(k) + c \leq a$, as desired.

(ii) Suppose c is an upper bound for A. Then for every $a \in A$ we have $a \leq c$. Hence $-a \geq -c$. Let

$$B := \{-a : a \in A\}.$$

Then $-c$ is a lower bound for B. Also, B is nonempty, since A is. Thus B has a least element, which we call b. Then for every $a \in A$ we have $b \leq -a$, since $-a \in B$. But by definition of B there is $d \in A$ such that $b = -d$. Hence we have $d \geq a$, because $-d \leq -a$. Therefore d is the largest element of A. ■

Definition 5.7. An integer $n \in \mathbb{Z}$ is called **even** if there exists $k \in \mathbb{Z}$ such that

$$n = 2k.$$

An integer which is not even is called **odd**.

Theorem 5.13. *Let $n \in \mathbb{Z}$ be an odd integer. Then there is $k \in \mathbb{Z}$ such that*

$$n = 2k + 1.$$

Proof. Let

$$A := \{l \in \mathbb{Z} : n - 2l \geq 0\}.$$

Then A is nonempty and bounded above. To see this, first suppose that $n \geq 0$. Then $0 \in A$, since $n - 2 \times 0 = n \geq 0$. Also, note that if $l \in A$ then $2l \leq n$. Hence if $l > 0$ then $l < 2l \leq n$; and if $l \leq 0$ then $l \leq 0 \leq n$. Thus in this case n is an upper bound for A. Next suppose $n < 0$. Then we have $n \in A$, since $n - 2n = -n > 0$. In addition, if $l \in A$ then we must have $l \leq 0$, since otherwise we would get $l < 2l \leq n < 0$, which contradicts the assumption of $l > 0$. Thus in this case 0 is an upper bound for A.

Now let k be the largest element of A. Then $n - 2k \geq 0$. Note that we cannot have $n - 2k = 0$, since n is not even. So we must have $n - 2k \geq 1$. However, if we have $n - 2k \geq 2$, then we get

$$n - 2(k + 1) = n - 2k - 2 \geq 2 - 2 = 0.$$

This means that $k+1 \in A$. But $k+1 > k$, and k is the largest element of A. Thus we cannot have $n - 2k \geq 2$, i.e. we must have $n - 2k < 2$. Hence $n - 2k \leq 1$. Therefore we obtain $n - 2k = 1$, which gives us $n = 2k + 1$, as desired. ■

Theorem 5.14. *Let $a, b, c, d \in \mathbb{Z}$. Suppose a, b are even, and c, d are odd. Then we have*

(i) *$a + b$ and $c + d$ are even.*
(ii) *$a + c$ is odd.*
(iii) *ab and ac are even.*
(iv) *cd is odd.*

Proof. Suppose $a = 2n$, $b = 2m$, $c = 2k + 1$, and $d = 2l + 1$.
(i) $a + b = 2(n + m)$, and $c + d = 2(k + l + 1)$.
(ii) $a + c = 2(n + k) + 1$.
(iii) $ab = 2(2nm)$, and $ac = 2(n(2k + 1))$.
(iv) $cd = 4kl + 2k + 2l + 1 = 2(2kl + k + l) + 1$. ■

5.2 Rings

Definition 5.8. A **ring** is a nonempty set R equipped with two binary operations

$$\begin{array}{c} R \times R \longrightarrow R \\ (a, b) \mapsto a + b \end{array}, \qquad \begin{array}{c} R \times R \longrightarrow R \\ (a, b) \mapsto ab \end{array},$$

called respectively **addition** and **multiplication**, such that

(i) The operations are **associative**, i.e. for every $a, b, c \in R$

$$a + (b + c) = (a + b) + c, \qquad a(bc) = (ab)c.$$

(ii) Addition is **commutative**, i.e. for every $a, b \in R$

$$a + b = b + a.$$

(iii) There exist elements $0, 1 \in R$, called respectively additive identity and multiplicative identity, such that for every $a \in R$

$$a + 0 = a, \qquad a1 = a = 1a.$$

(iv) For every $a \in R$ there exists $b \in R$, called its additive inverse, such that

$$a + b = 0.$$

(v) Multiplication is **distributive** over addition, i.e. for every $a, b, c \in R$

$$a(b + c) = ab + ac, \qquad (b + c)a = ba + ca.$$

Remark. Let b be an additive inverse of a. Then note that due to the commutativity of addition we also have

$$0 + a = a, \qquad b + a = 0.$$

Remark. Note that in every ring, we use the same notation for addition, multiplication, and $0, 1$. However, if we want to emphasize that they are associated to a ring R, we denote them by $+_R, \cdot_R, 0_R, 1_R$ respectively.

***Example* 5.1.** $\mathbb{Z}$, equipped with its standard addition and multiplication, is a ring, as we showed in the last section. $\mathbb{N}$ is not a ring as it does not have an additive identity, and its elements do not have additive inverse. Similarly, ω is not a ring, since its elements do not have additive inverse.

Definition 5.9. A multiplicative inverse of an element $a \in R$ is an element $b \in R$ such that

$$ab = 1 = ba.$$

In this case a is called **invertible**.

Proposition 5.1. *Let R be a ring. Then for every $a, b, c \in R$ we have*

(i) **(Cancellation Laws)**
  (a) *If $a + c = b + c$ then $a = b$.*
  (b) *If $ac = bc$, and c is invertible, then $a = b$.*
  (c) *If $ca = cb$, and c is invertible, then $a = b$.*

(ii) *Additive and multiplicative identities of R are unique.*

(iii) *Additive inverse of any element of R is unique; and multiplicative inverse of any invertible element of R is unique*

(iv) $0a = 0 = a0$.

Proof. **(i)** Suppose d is an additive inverse of c. Then we can add d to both sides of $a + c = b + c$ to obtain $(a + c) + d = (b + c) + d$. Now by associativity of addition we have $a + (c + d) = b + (c + d)$. Since $c + d = 0$, we get $a + 0 = b + 0$; and hence $a = b$. The multiplicative cases can be proved similarly using a multiplicative inverse of c.

(ii) Suppose $0, \tilde{0}$ are both additive identities of R. Then we have $\tilde{0} = \tilde{0} + 0$, since 0 is an additive identity. We also have $0 + \tilde{0} = 0$, since $\tilde{0}$ is an additive identity. However we know that $\tilde{0} + 0 = 0 + \tilde{0}$, because addition is commutative. Therefore we must have $\tilde{0} = 0$, as desired. Similarly, for two multiplicative identities $1, \tilde{1}$ we have $\tilde{1} = \tilde{1}1 = 1$.

(iii) Suppose $b, \tilde{b}$ are both additive inverses of a. Then we have

$$\tilde{b} + a = 0 = b + a.$$

Thus by cancellation law we get $\tilde{b} = b$, as desired. The multiplicative case is similar.

(iv) We have

$$0 + 0a = 0a = (0 + 0)a = 0a + 0a.$$

Hence by cancellation law we get $0 = 0a$. The other equality can be proved similarly. ∎

Definition 5.10. Additive and multiplicative identities of a ring are respectively called **zero** and **identity** of the ring.

The unique additive inverse of an element a is denoted by $-a$, and is called its **opposite**. Also for two elements a, b we set $a - b := a + (-b)$.

If an element a has multiplicative inverse, we denote it by a^{-1}, and we call it the **inverse** of a.

Proposition 5.2. *Let R be a ring. Then for every $a, b \in R$ we have*

(i) $-0 = 0$, *and* $1^{-1} = 1$.

(ii) $-(-a) = a$.

(iii) $-(a + b) = (-a) + (-b) = -a - b$.

(iv) *If a is invertible, then a^{-1} is also invertible, and*

$$(a^{-1})^{-1} = a.$$

(v) *If a and b are invertible, then ab is also invertible, and we have*

$$(ab)^{-1} = b^{-1}a^{-1}.$$

(vi) $(-a)b = -ab = a(-b)$. *As a result we have*

$$(-a)(-b) = ab, \qquad -a = (-1)a.$$

Notation. By $-ab$ we mean $-(ab)$. In other words, we assume that multiplication binds stronger than taking the opposite.

Proof. **(i)** We have $0 + 0 = 0$, and $1 \cdot 1 = 1$. Now the result follows from the uniqueness of inverse.

(ii) This is similar to (iv).

(iii) This is similar to (v). Note that the last equality in (iii) holds by definition.

(iv) We know that $a^{-1}a = 1 = aa^{-1}$. Thus a^{-1} is invertible. We also have $(a^{-1})^{-1} = a$, due to the uniqueness of inverse.

(v) First note that

$$\begin{aligned}(b^{-1}a^{-1})(ab) &= b^{-1}(a^{-1}(ab)) = b^{-1}((a^{-1}a)b)\\ &= b^{-1}(1b) = b^{-1}b = 1.\end{aligned}$$

Similarly $(ab)(b^{-1}a^{-1}) = 1$. Therefore ab is invertible. Now the result follows from the uniqueness of inverse.

(vi) We have

$$ab + (-a)b = (a + (-a))b = 0b = 0.$$

Thus uniqueness of additive inverse implies $(-a)b = -ab$. The equality $a(-b) = -ab$ can be proved similarly. Now we have

$$(-a)(-b) = -(-a)b = -(-ab) = ab,$$
$$(-1)a = -(1a) = -a.$$

■

***Exercise* 5.1.** Show that if a is invertible, then $-a$ is also invertible and we have

$$(-a)^{-1} = -a^{-1}.$$

As a consequence we have $(-1)^{-1} = -1^{-1} = -1$.

Solution. By the previous theorem we have

$$(-a^{-1})(-a) = a^{-1}a = 1 = aa^{-1} = (-a)(-a^{-1}).$$

Thus we get the desired result due to the uniqueness of inverse. ■

Definition 5.11. A **commutative ring** is a ring in which multiplication is commutative, i.e. for every elements a, b we have

$$ab = ba.$$

Also, we say two elements a and b in a ring **commute** if $ab = ba$.

***Example* 5.2.** $\mathbb{Z}$ is a commutative ring.

Definition 5.12. Let R be a ring, and let n be a nonnegative integer. We inductively define

(i) $[0] := 0_R$,

(ii) $[n + 1] := [n] + 1_R$.

When m is a negative integer, $n := -m$ is positive. In this case we define

$$[m] = [-n] := -[n].$$

Remark. Note that in the above definition, each one of the $0, 1$, and $\pm$, has two different meanings. Also note that we have

$$[1] = [0] + 1_R = 0_R + 1_R = 1_R.$$

Remark. In the terminology of recursion theorem, we have constructed the function $f(n) = [n]$ for $n \geq 0$, by using the function $F(n, a) = a + 1_R$. Although, to be precise, we have to actually define f on ω; and then for a nonnegative integer n we have to set $[n] = f(E^{-1}(n))$, where E is the one-to-one correspondence between ω and the set of nonnegative integers, constructed in Theorem 5.11.

Notation. Let R be a ring, and let $a \in R$. We usually abuse the notation and write n instead of $[n]$. We also set

$$na := [n]a.$$

Remark. The next proposition shows that the operations of $\mathbb{Z}$ on n's and the operations of R on n's are compatible. Therefore the above abbreviation does not lead to any confusion.

Remark. na is actually the nth *additive power* of a, i.e. the nth power of a with respect to the binary operation $+$, as defined in Section 5.6. To see this note that $0a = [0]a = 0_R$, since $[0] = 0_R$. Also for $n > 0$ we have

$$(n+1)a = [n+1]a = ([n] + 1)a = [n]a + 1a = na + a.$$

In addition, for $m < 0$ set $n = -m$. Then we have

$$ma = (-n)a = [-n]a = (-[n])a = [n](-a) = n(-a).$$

Remark. A consequence of the above remark is that for $n > 0$ we have

$$na = \sum_{j=1}^{n} a = \overbrace{a + a + \cdots + a}^{n \text{ times}},$$

as shown in Theorem 5.40. In particular we have

$$[n] = n1_R = \overbrace{1_R + 1_R + \cdots + 1_R}^{n \text{ times}}.$$

Proposition 5.3. *In any ring R we have*

(i) *For every $n \in \mathbb{Z}$, $[n]$ commutes with all elements of R.*

(ii) *For every $n, m \in \mathbb{Z}$ we have*

$$[n+m] = [n] + [m], \qquad [nm] = [n][m].$$

(iii) *For every $n \in \mathbb{Z}$ we have*

$$[-n] = -[n].$$

Proof. **(i)** Let a be an arbitrary element of R. Then $[0]a = 0a = 0 = a0 = a[0]$, so $[0]$ commutes with every a. Now suppose for some $n > 0$, $[n]$ commutes with every a. Then we have

$$[n+1]a = ([n]+1)a = [n]a + 1a = a[n] + a1 = a([n]+1) = a[n+1].$$

Hence by induction, $[n]$ commutes with every a for all $n > 0$. Next suppose $m = -n < 0$. Then

$$[m]a = (-[n])a = -[n]a = -a[n] = a(-[n]) = a[m].$$

(ii) Since $[n] = n1$ is the nth additive power of $1 \in R$, these relations are special cases of the properties of powers as proved in Section 5.6. For example, for the second equality we can say that $[nm]$ is the nmth power of 1, so it is equal to the mth power of the nth power of 1. But the nth power of 1 is $[n]$. Hence the nmth power of 1 equals the mth power of $[n]$, which is $[m][n]$. But by (i) we have $[m][n] = [n][m]$. Therefore $[nm] = [n][m]$ as desired.

(iii) For every $n \in \mathbb{Z}$ we have

$$[n] + [-n] = [n + (-n)] = [0] = 0;$$

so we must have $[-n] = -[n]$, due to the uniqueness of opposite. ■

Remark. Note that if $R = \mathbb{Z}$ then the map $[\,] : \mathbb{Z} \to \mathbb{Z}$ is the identity map, i.e. for every $n \in \mathbb{Z}$ we have

$$[n] = n.$$

We already know this holds for $n = 0, 1$. So we only need to show the equality for $n > 0$, since then for $m = -n < 0$ we have $[m] = [-n] = -[n] = -n = m$. We proceed by induction on $n > 0$ (more precisely, by induction on $E^{-1}(n)$, where E is the one-to-one correspondence between ω and the set of nonnegative integers, constructed in Theorem 5.11). The base of induction is already known to hold. For the induction step, by using the definition of $[\,]$ we get $[n+1] = [n] + 1 = n + 1$, as desired.

As a result, for every positive integer n we have

$$\overbrace{1 + 1 + \cdots + 1}^{n \text{ times}} = [n] = n > 0,$$

where $1 = 1_{\mathbb{Z}}$. In other words, every positive integer is the sum of several 1's.

Proposition 5.4. *Let R be a ring. Then for every $m, n \in \mathbb{Z}$ and $a, b \in R$ we have*

(i) $(-n)a = n(-a) = -(na)$.

(ii) $(n+m)a = na + ma$.

(iii) $m(na) = (mn)a$.
(iv) $n(a + b) = na + nb$.
(v) $(ma)(nb) = mn(ab) = (na)(mb)$.
(vi) *If a commutes with b, then na commutes with mb.*

Proof. Parts (i) to (iv) are true for any notion of power as proved in Section 5.6. They can also be proved directly, as we do below. We have

$$\begin{aligned}
(-n)a &= [-n]a = (-[n])a = [n](-a) = n(-a),\\
(-n)a &= [-n]a = (-[n])a = -[n]a = -na,\\
(n + m)a &= [n + m]a = ([n] + [m])a = [n]a + [m]a = na + ma,\\
m(na) &= [m]([n]a) = ([m][n])a = [mn]a = (mn)a,\\
n(a + b) &= [n](a + b) = [n]a + [n]b = na + nb.
\end{aligned}$$

For part (v) we have

$$(ma)(nb) = ([m]a)([n]b) = [m]a[n]b = [m][n]ab = [mn]ab = mn(ab).$$

Note that we used the generalized associativity of the product of R, and the fact that $[n]$ commutes with all elements of R. Now for the second equality we use the first one to obtain

$$(ma)(nb) = mn(ab) = nm(ab) = (na)(mb).$$

Finally, for part (vi) we have

$$(na)(mb) = nm(ab) = mn(ab) = mn(ba) = (mb)(na).$$ ■

Definition 5.13. Let R be a ring, and let n be a nonnegative integer. Let $a \in R$. We inductively define the **powers** of a as follows
(i) $a^0 := 1_R$,
(ii) $a^{n+1} := a^n a$.
When m is a negative integer, $n := -m$ is positive. In this case, if a is invertible we define

$$a^m = a^{-n} := (a^{-1})^n.$$

Remark. Note that we have

$$a^1 = a^0 a = 1a = a.$$

Remark. In the terminology of recursion theorem, we have constructed the function $f(n) = a^n$ for $n \geq 0$, by using the function $F(n, s) = sa$. Although, to be precise, we have to actually define f on ω; and then for a nonnegative integer n we have to set $a^n = f(E^{-1}(n))$, where E is the one-to-one correspondence between ω and the set of nonnegative integers, constructed in Theorem 5.11.

Remark. As a consequence of Theorem 5.40, for $n > 0$ we have

$$a^n = \prod_{j=1}^{n} a = \overbrace{aa\cdots a}^{n \text{ times}}.$$

Theorem 5.15. *Let R be a ring. Let $n, m \in \mathbb{Z}$. Then for every $a, b \in R$ we have*

(i) *If a commutes with b, then a^n commutes with b^m, for all $m, n \geq 0$. If one or both of a, b are invertible, we can allow n and/or m to be negative too.*

(ii) *If a is invertible, then a^n is also invertible for all $n \in \mathbb{Z}$, and*

$$(a^n)^{-1} = a^{-n} = (a^{-1})^n.$$

(iii) *$a^n a^m = a^{n+m}$ for all $m, n \geq 0$. If a is invertible, we can allow m, n to be negative too.*

(iv) *$(a^n)^m = a^{nm}$ for all $m, n \geq 0$. If a is invertible, we can allow m, n to be negative too.*

(v) *If a, b commute, we have $a^n b^n = (ab)^n$ for all $n \geq 0$. If a, b are invertible, we can allow n to be negative too.*

Proof. All the proofs are by induction. We will only write the induction steps below, since the base of inductions can be checked easily.

(i) For $m \geq 0$ we have

$$ab^{m+1} = ab^m b = b^m ab = b^m ba = b^{m+1}a.$$

If b is invertible we have

$$ab^{-1} = 1ab^{-1} = b^{-1}bab^{-1} = b^{-1}abb^{-1} = b^{-1}a.$$

Thus by the first part $b^{-m} = (b^{-1})^m$ commutes with a. By repeating this argument with fixed m, we see that a^n commutes with b^m too.

(ii) When $n \geq 0$ we have

$$(a^{n+1})a^{-n-1} = a^n a(a^{-1})^{n+1} = aa^n(a^{-1})^n a^{-1} = aa^n a^{-n}a^{-1} = aa^{-1} = 1.$$

When $n = -m < 0$ we have $a^{-m} = (a^{-1})^m$. Hence by the previous part we get

$$(a^{-m})^{-1} = ((a^{-1})^m)^{-1} = (a^{-1})^{-m} = ((a^{-1})^{-1})^m = a^m.$$

The second equality holds by definition when $n > 0$. When $n = 0$ we have

$$(a^{-1})^0 = 1 = a^0 = a^{-0}.$$

And when $n = -m < 0$ we have

$$(a^{-1})^{-m} = ((a^{-1})^{-1})^m = a^m = a^{-n}.$$

(iii) When $n, m \geq 0$ we have

$$a^n a^{m+1} = a^n a^m a = a^{n+m} a = a^{n+m+1}.$$

Now suppose a is invertible. Then we have

$$\begin{aligned} a^{-n}a^{m+1} &= a^{-n}a^m a \\ &= a^{-n+m}a = \begin{cases} (a^{-1})^{n-m}a = (a^{-1})^{n-m-1}a^{-1}a & \text{if } -n+m < 0, \\ \qquad = (a^{-1})^{n-m-1} = a^{-n+m+1} & \\ a^{-n+m+1} & \text{if } -n+m \geq 0. \end{cases} \end{aligned}$$

We also have

$$\begin{aligned} a^n a^{-m} &= (a^{-1})^{-n}(a^{-1})^m = (a^{-1})^{-n+m} = a^{n-m}, \\ a^{-n}a^{-m} &= (a^{-1})^n (a^{-1})^m = (a^{-1})^{n+m} = a^{-n-m}. \end{aligned}$$

(iv) For $n, m \geq 0$ we have

$$(a^n)^{m+1} = (a^n)^m a^n = a^{nm} a^n = a^{nm+n} = a^{n(m+1)}.$$

If a is invertible we have

$$\begin{aligned} (a^{-n})^m &= ((a^{-1})^n)^m = (a^{-1})^{nm} = a^{-nm}, \\ (a^{\pm n})^{-m} &= ((a^{\pm n})^m)^{-1} = (a^{\pm nm})^{-1} = a^{\mp nm}. \end{aligned}$$

(v) For $n \geq 0$ we have

$$a^{n+1}b^{n+1} = a^n ab^n b = a^n b^n ab = (ab)^n ab = (ab)^{n+1},$$

and if a, b are invertible we have

$$a^{-n}b^{-n} = (a^{-1})^n (b^{-1})^n = (a^{-1}b^{-1})^n = ((ba)^{-1})^n = (ba)^{-n} = (ab)^{-n}. \qquad \blacksquare$$

Remark. Suppose k, n are natural numbers, and $0 \leq k \leq n$. Remember that n factorial is $n! = n \cdot (n-1) \cdots 2 \cdot 1$, and the binomial coefficient is

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

It is easy to see that $\binom{n}{0} = 1 = \binom{n}{n}$ for all $n \geq 0$. We have seen in Section 4.4 that $\binom{n}{k}$ is the natural number which is the number of ways to choose k objects among n objects.

Theorem 5.16. *For all integers $1 \le k \le n$ we have*

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

Proof. We have

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

■

Theorem 5.17. *For two commuting elements a, b in a ring, and a positive integer n, we have*

(i) **Binomial Theorem:**

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

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

Proof. **(i)** The proof is by induction on n. The case of $n = 1$ is obvious. For the induction step we have

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

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

(We replaced k with $j-1$ in the 2nd sum.)

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

(We replaced j with k in the 2nd sum.)

$$= a^{n+1} + \left(\sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1} \right] a^{n+1-k} b^k \right) + b^{n+1}$$

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

Note that since the binomial coefficients are positive integers, we can multiply the ring elements with them.

(ii) We have

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

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

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

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

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

(We replaced k with $j-1$ in the 2nd sum.)

$$= a^n + \left(\sum_{k=1}^{n-1} a^{n-k} b^k + \sum_{k=1}^{n-1} (-1) a^{n-k} b^k \right) + (-1) b^n$$

(We replaced j with k in the 2nd sum.)

$$= a^n + \left(\sum_{k=1}^{n-1} a^{n-k} b^k + (-1) \sum_{k=1}^{n-1} a^{n-k} b^k \right) - b^n$$

$$= a^n - b^n.$$

■

Definition 5.14. A **field** is a commutative ring in which $1 \neq 0$, and all nonzero elements are invertible.

Notation. For two elements a, b in a field, when $b \neq 0$, we set $a/b = \frac{a}{b} := ab^{-1}$.

Definition 5.15. An **integral domain** is a commutative ring in which $1 \neq 0$, and for every elements a, b we have

$$ab = 0 \implies a = 0 \text{ or } b = 0.$$

***Example* 5.3.** $\mathbb{Z}$ is an integral domain.

Theorem 5.18. *Suppose R is an integral domain, and $a, b, c \in R$. Then the cancellation law holds for the multiplication of R, i.e.*

$$ac = bc,\ c \neq 0 \implies a = b.$$

Proof. We have $(a - b)c = 0$. Hence $a - b = 0$, since $c \neq 0$. ■

Theorem 5.19. *Every field is also an integral domain.*

Proof. If $ab = 0$, and $a \neq 0$, we have

$$b = 1b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0 = 0.$$ ■

Definition 5.16. Suppose R is a ring, and $S \subset R$. We say S is a **subring** of R if S contains the identity of R, and for all $a, b \in S$ we have $-a, a + b, ab \in S$.

Proposition 5.5. *Suppose R is a ring, and S is a subring of R. Then $0 \in S$, and S is itself a ring with the addition and multiplication inherited from R.*

Proof. We have $-1 \in S$, since $1 \in S$. Then $0 = 1 + (-1) \in S$. The associativity, commutativity, and distributivity laws are trivially satisfied in S, since they are satisfied in R. Also by definition S contains the opposite of each of its elements. Hence S is a ring. ■

Definition 5.17. Suppose in a ring R we have $n = \overbrace{1 + 1 + \cdots + 1}^{n \text{ times}} = 0$ for some positive integer n. Then the smallest such n is called the **characteristic** of R. If this never happens we say that R has characteristic zero.

Theorem 5.20. *The characteristic of an integral domain is either zero or a prime positive integer.*

Remark. Prime numbers are positive integers defined in the next section, which have the property that they cannot be written as the product of two smaller positive integers.

Proof. If the conclusion does not hold, the characteristic of the integral domain R is $n = pq$ for some positive integers $p, q < n$. But in R we have $pq = n = 0$; hence $p = 0$ or $q = 0$ in R, which is in contradiction with the fact that n is the smallest positive integer with this property. ■

Remark. Suppose F is a field in which $n \neq 0$. Then we have

$$\overbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}^{n \text{ times}} = n\left(\frac{1}{n}\right) = n(1n^{-1}) = nn^{-1} = 1.$$

In other words, the nth additive power of $\frac{1}{n}$ is 1.

***Exercise* 5.2.** Suppose R is a ring, and A is a nonempty set. On the space of all functions from A into R we define the binary operations of *pointwise* addition and multiplication of functions, i.e. for two functions $f, g : A \to R$ and every $a \in A$ we define

$$(f + g)(a) := f(a) + g(a),$$
$$(fg)(a) := f(a)g(a).$$

Show that this space is a ring with these operations. (Note that $f + g, fg$ are functions, since they assign a uniquely determined value to every $a \in A$.)

5.3 Factorization of Integers

In this section, the absolute value of integers is needed in some of the results. So we go over its definition and basic properties first.

Definition 5.18. Let $a \in \mathbb{Z}$. Then the **absolute value** of a is

$$|a| := \begin{cases} a & \text{if } a \geq 0, \\ -a & \text{if } a < 0. \end{cases}$$

Theorem 5.21. *For all $a, b \in \mathbb{Z}$ we have*

(i) $|a| \geq 0$, *and for $a \neq 0$ we have $|a| > 0$.*
(ii) $|-a| = |a|$.
(iii) $|a| = |b|$ *if and only if* $a = \pm b$.
(iv) $|ab| = |a||b|$.
(v) *For $b > 0$ we have $|a| < b$ if and only if $-b < a < b$; and for $b \geq 0$ we have $|a| \leq b$ if and only if $-b \leq a \leq b$.*
(vi) $-|a| \leq a \leq |a|$.

(vii) **Triangle Inequality:**

$$|a+b| \leq |a| + |b|.$$

Proof. The proofs are the same as of Theorem 6.11. ■

Proposition 5.6. *Let $n \in \mathbb{Z}$, and suppose $n \neq 0$. Then we have $|n| \geq 1$.*

Proof. Since $n \neq 0$ we have $n > 0$ or $n < 0$. Thus we must have $n \geq 1$ or $n \leq -1$, because there is no integer between $0, 1$ nor between $-1, 0$. In the first case we have $|n| = n \geq 1$, and in the second case we have $|n| = -n \geq -(-1) = 1$. ■

Now we move to the main topic of this section. We start by the notion of divisibility.

Definition 5.19. Let $a, b \in \mathbb{Z}$. If there exists $r \in \mathbb{Z}$ such that $b = ra$, then we say a **divides** b, or a is a **divisor** of b, or b is a **multiple** of a; and we write $a \mid b$. If a is not a divisor of b we write $a \nmid b$.

Proposition 5.7. *Let $a, b, c \in \mathbb{Z}$, then we have*

(i) $1 \mid a$, $a \mid a$, *and* $a \mid 0$.
(ii) *If* $a \mid b$ *and* $b \neq 0$ *then* $|a| \leq |b|$.
(iii) $a \mid 1$ *if and only if* $a = \pm 1$.
(iv) $a \mid b$ *and* $b \mid a$ *if and only if* $a = \pm b$.
(v) *If* $a \mid b$ *and* $b \mid c$ *then* $a \mid c$.
(vi) $a \mid b$ *implies* $a \mid bc$ *and* $ac \mid bc$.
(vii) *If* $a \mid b$ *and* $a \mid c$ *then* $a \mid bx + cy$ *for every* $x, y \in \mathbb{Z}$, *in particular* $a \mid b + c$.

Remark. Note that if $0 \mid a$ then for some r we have $a = r0 = 0$. Hence 0 is the only number that has 0 as a divisor.

Proof. **(i)** We have $a1 = 1a = a$, and $0a = 0$.

(ii) There is $r \in \mathbb{Z}$ such that $b = ra$. Note that $a, r \neq 0$ since $b \neq 0$. Thus $|r| \geq 1$. Hence $|b| = |ra| = |r||a| \geq 1|a| = |a|$, since $|a| > 0$.

(iii) As a consequence of the above part, we can see that if $a \mid 1$ then $|a| \leq 1$, and thus $-1 \leq a \leq 1$. Hence a is either 1,-1, or 0. However, 0 does not divide 1; so we must have $a = \pm 1$.

(iv) First suppose $a \mid b$ and $b \mid a$. Then we have $|a| \leq |b|$ and $|b| \leq |a|$. Hence $|a| = |b|$, and therefore $a = \pm b$. Conversely, if $a = b$ or $a = -b$, then we trivially have $a \mid b$ and $b \mid a$.

(v) There are $r, s \in \mathbb{Z}$ such that $b = ra$ and $c = sb$. Hence $c = sra$, so $a \mid c$.

(vi) There is $r \in \mathbb{Z}$ such that $b = ra$. Thus $bc = rac = rca$. So $a \mid bc$, and $ac \mid bc$.

(vii) There are $r, s \in \mathbb{Z}$ such that $b = ra$ and $c = sa$. Therefore $bx + cy = (rx + sy)a$, so $a \mid bx + cy$. For $x = y = 1$ we get $a \mid b + c$. ■

Division Algorithm. *For every $a \in \mathbb{Z}$ and every nonzero $b \in \mathbb{Z}$, there are unique $q, r \in \mathbb{Z}$ such that*

$$a = bq + r, \qquad \text{with } 0 \le r < |b|.$$

Here q is called the ***quotient*** *and r is called the* ***remainder****.*

Proof. We prove the existence of q, r by strong induction on $|a|$. (Note that $|a| \ge 0$, so it can be regarded as a natural number. More precisely, for a given b, we prove by strong induction on $n \in \omega$ that q, r exist for a with $|a| = E(n)$, where $E : \omega \to \mathbb{Z}$ is defined in Theorem 5.11.) Now if $|a| < |b|$ we can simply put $q = 0$ and $r = a$ when $a \ge 0$. And when $a < 0$ we can choose $q = \mp 1$ so that $bq = -|b|$; then we set $r = |b| + a$ (note that $-|b| < a < 0$ in this case, so $0 < |b| + a < |b|$ as desired). Also, if $|a| = |b|$ then we have $a = \pm b$, so we can put $q = \pm 1$ and $r = 0$.

Now suppose $|a| > |b|$ and the conclusion holds for all integers with absolute value less than $|a|$. Then we have $0 < |a| - |b| < |a|$, since $|b| > 0$. But, according to the signs of a, b, the positive number $|a| - |b|$ has one of the four possible values $\pm(a + b)$ or $\pm(a - b)$. So for some choice of $\pm$'s we have

$$|a \pm b| = |\pm(a \pm b)| = |a| - |b| < |a|.$$

In other words, there is $u \in \{-1, 1\}$ such that $|a + ub| < |a|$. Thus by the induction hypothesis we have

$$a + ub = bq + r,$$

for some q, r with $0 \le r < |b|$. Therefore we have $a = b(q - u) + r$ as desired.

Next, to prove the uniqueness, suppose to the contrary that we have

$$bq_1 + r_1 = a = bq_2 + r_2,$$

where $0 \le r_i < |b|$. Then we have $b(q_1 - q_2) = r_2 - r_1$, so $b | r_2 - r_1$. If $r_2 - r_1 \ne 0$ then we must have $|b| \le |r_2 - r_1|$. However, $|r_2 - r_1| < |b|$ since $r_2 - r_1 \le r_2 < |b|$ and $r_2 - r_1 \ge -r_1 > -|b|$. Thus we have arrived at a contradiction. Hence $r_1 = r_2$, and therefore $q_1 = q_2$. ∎

Second Proof. Let us give another proof for the existence of q, r. Consider the set

$$A := \{a - bc : c \in \mathbb{Z} \text{ and } a - bc \ge 0\}.$$

First note that $A \ne \emptyset$. To see this, let $u \in \{-1, 1\}$ be such that $|b| = ub$. We also know that $|b| \ge 1$ since $b \ne 0$. Then for $c = -u|a|$ we have

$$a - bc = a + bu|a| = a + |b||a| \ge a + 1|a| = a + |a| \ge 0,$$

since $a \ge -|a|$.

Now let r be the least element of A. Then, by definition, there is $q \in \mathbb{Z}$ such that $r = a - bq \geq 0$, or equivalently $a = bq + r$ with $r \geq 0$. We claim that $r < |b| = ub$ (for some $u \in \{\pm 1\}$). Suppose to the contrary that $r \geq |b|$. Then we have

$$a = bq + r = bq + |b| - |b| + r = b(q + u) + (r - |b|).$$

Thus $r - |b| = a - b(q + u)$ and $r - |b| \geq 0$. So $r - |b| \in A$. However, $r - |b| < r$ (as $|b| > 0$), and this contradicts the fact that r is the least element of A. Hence we must have $r < |b|$, as desired. ■

Remark. Note that, in the above (second) proof we show that the remainder r is the least element of a set A of integers. We also know that the least element of a set is unique. But, based on this information alone, we cannot conclude that the remainder r is unique, and we need a separate proof for this fact. Because, there might be other remainders, constructed in different ways, which are not the least element of the set A. For example, if, similarly to the division algorithm in more general rings (see Section A.1), we merely require that the absolute value of the remainder is less than $|b|$, then although the least of the set A is still a remainder, there is also another remainder which does not belong to A.

Theorem 5.22. *Let $n, b \in \mathbb{N}$, and suppose $b > 1$. Then there is a unique integer $m \geq 0$, and unique integers $0 \leq r_0, r_1, \ldots, r_m < b$ with $r_m \neq 0$, such that*

$$n = r_m b^m + r_{m-1} b^{m-1} + \cdots + r_1 b + r_0.$$

Remark. This is the *base b representation* of the number n. It is sometimes denoted by $n = (r_m \ldots r_0)_b$.

Proof. First we prove the existence. The proof is by strong induction on n. If $1 \leq n < b$, then we can put $m = 0$ and $r_0 = n > 0$. Suppose the conclusion holds for all $k < n$. We can assume that $n \geq b$. Now we have $n = bq + r$ where $0 \leq r < b$. But q cannot be nonpositive since then we would have $n \leq r < b$, contrary to our assumption. Thus we must have $q > 0$. Then $n = bq + r > q + r \geq q$. Therefore by the induction hypothesis we have

$$q = s_m b^m + \cdots + s_1 b + s_0,$$

for some $0 \leq s_i < b$ with $s_m \neq 0$. Then we have

$$n = bq + r = s_m b^{m+1} + \cdots + s_1 b^2 + s_0 b + r,$$

as desired.

For the uniqueness, again the proof is by strong induction on n. If $1 \leq n < b$, then the representation is obviously unique. Because $b^i \geq b > n$, so we cannot have

$m > 0$. Suppose the uniqueness holds for all positive integers less than n. We can assume that $n \geq b$. Suppose that

$$s_k b^k + \cdots + s_1 b + s_0 = n = r_m b^m + \cdots + r_1 b + r_0,$$

where $0 \leq r_i, s_j < b$ and $r_m, s_k \neq 0$. Then $m, k > 0$, since $n \geq b$. Now we have

$$(s_k b^{k-1} + \cdots + s_1)b + s_0 = n = (r_m b^{m-1} + \cdots + r_1)b + r_0.$$

Therefore r_0, s_0 are the remainder in the division of n by b. Hence $r_0 = s_0$. Since the quotient is also unique, we have

$$s_k b^{k-1} + \cdots + s_1 = r_m b^{m-1} + \cdots + r_1.$$

But, in the above paragraph we showed that when $n \geq b$ the quotient is a positive integer strictly less than n. Therefore by the induction hypothesis we have $m - 1 = k - 1$, and $r_i = s_i$ for $1 \leq i \leq m$. Hence $m = k$, and the base b representation of n is unique. ■

Definition 5.20. The **greatest common divisor (g.c.d)** of two nonzero integers a, b is an integer c that satisfies

(i) $c \mid a$ and $c \mid b$.

(ii) If $r \in \mathbb{Z}$ is a common divisor of a, b, i.e. if $r \mid a$ and $r \mid b$, then $r \leq c$.

Remark. Note that if the g.c.d of two nonzero integers a, b exists it is unique. Because if c, c' are both g.c.d of a, b we have $c' \mid a$ and $c' \mid b$ so $c' \leq c$, and conversely $c \mid a$ and $c \mid b$ so $c \leq c'$. Thus we must have $c = c'$. The following theorem, the well-known Euclidean algorithm, shows that the g.c.d of any two nonzero integers exists.

Euclidean Algorithm. *Let $a, b \in \mathbb{Z}$ be nonzero. Consider the following sequence of divisions*

$$\begin{aligned} a &= bq_1 + r_1, \\ b &= r_1 q_2 + r_2, \\ r_1 &= r_2 q_3 + r_3, \\ &\vdots \end{aligned}$$

In this sequence, after finitely many divisions the remainder becomes zero, i.e. for some $n \geq 0$ we have (we set $r_0 = b$ and $r_{-1} = a$)

$$r_{n-1} = r_n q_{n+1}.$$

Then $|r_n|$ is the greatest common divisor of a, b.

Remark. Note that if $n > 0$ then $r_n = |r_n|$ is the g.c.d of a, b, because the remainders in a division are nonnegative. However, if $n = 0$ (in which case we have $a = bq_1$) then $|b| = |r_0|$ is the g.c.d of a, b.

Remark. The above sequence of divisions can be constructed using recursion theorem. To see this consider the function $G : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}$ defined by (r' is the remainder of the division of a' by b')

$$G(a', b') = \begin{cases} (b', r') & \text{if } b' \neq 0, \\ (0, 0) & \text{if } b' = 0. \end{cases}$$

Now the recursion theorem provides us a function $f : \omega \to \mathbb{Z} \times \mathbb{Z}$ that satisfies $f(0) = (a, b)$, and

$$f(n + 1) = G(f(n)).$$

Then $f(1) = G(a, b) = (b, r_1)$. Next we have

$$f(2) = G(b, r_1) = \begin{cases} (r_1, r_2) & \text{if } r_1 \neq 0, \\ (0, 0) & \text{if } r_1 = 0, \end{cases}$$

and so on. Then in the last two steps we will have $f(n) = (r_{n-1}, r_n)$ and $f(n+1) = (r_n, 0)$. After that we get $f(n + j) = (0, 0)$ for $j > 1$.

Proof. First note that the remainders satisfy $|b| > r_1 > r_2 > \cdots \geq 0$. Then we must have $r_j \leq |b| - j$, since $r_1 < |b|$ implies $r_1 \leq |b| - 1$. And if the inequality holds for j then for $j + 1$ we have

$$r_{j+1} < r_j \implies r_{j+1} \leq r_j - 1 \leq |b| - j - 1 = |b| - (j + 1),$$

as wanted. Hence if $r_1, \ldots, r_{|b|}$ are all nonzero then for $j = |b| + 1$ we would have $r_{|b|+1} \leq |b| - |b| - 1 = -1$, which is in contradiction with remainders being nonnegative. Thus for some $0 \leq n \leq |b| - 1$ we must have $r_{n+1} = 0$ and $r_j \neq 0$ for $j \leq n$ (we can simply take n to be the largest index for which $r_n \neq 0$, noting that $r_{-1} = a$ and $r_0 = b$ are nonzero by assumption). Then we have

$$r_{n-1} = r_n q_{n+1} + r_{n+1} = r_n q_{n+1} + 0 = r_n q_{n+1}.$$

Now let us show that r_n is a common divisor of a, b. First note that $r_n \mid r_{n-1}$ by the above relation. It is also obvious that $r_n \mid r_n$. We claim that $r_n \mid r_{n-i}$ for every $i \leq n + 1$. We have already seen this for $i = 0, 1$. Suppose $i \geq 2$ and the claim holds for every $k < i$. Then we have

$$r_{n-i} = r_{n-i+1} q_{n-i+2} + r_{n-i+2} = r_{n-(i-1)} q_{n-i+2} + r_{n-(i-2)}.$$

By the induction hypothesis we have $r_n \mid r_{n-(i-2)}$ and $r_n \mid r_{n-(i-1)}$. Therefore we get $r_n \mid r_{n-(i-1)}q_{n-i+2} + r_{n-(i-2)}$, which means $r_n \mid r_{n-i}$. Hence the claim is proved. In particular, we have $r_n \mid r_0 = b$ and $r_n \mid r_{-1} = a$. Hence r_n is a common divisor of a, b.

Next suppose $r \mid a$ and $r \mid b$. We claim that $r \mid r_j$ for every $j \leq n$. By assumption, this holds when $j = -1, 0$. Suppose $j \geq 1$ and the claim holds for every $k < j$. Then we have

$$r_{j-2} = r_{j-1}q_j + r_j \implies r_j = r_{j-2} - q_j r_{j-1}.$$

Now by the induction hypothesis we have $r \mid r_{j-2}$ and r_{j-1}. So $r \mid r_{j-2} - q_j r_{j-1}$, which means $r \mid r_j$. Therefore the claim is proved. In particular, we have $r \mid r_n$. Hence for $n > 0$ we get $r \leq |r| \leq |r_n| = r_n$, since $r_n > 0$ when $n > 0$. Thus r_n is the g.c.d of a, b. And for $n = 0$ we have $r_0 = b$ and $a = bq_1$. In this case we can still conclude that $r \leq |r_0| = |b|$. Now note that $|b| = \pm b$ is a divisor of b, and it is also a divisor of a, so it is the g.c.d of a, b. ■

Proposition 5.8. *Let $a, b \in \mathbb{Z}$ be nonzero. Then $c \in \mathbb{Z}$ is the g.c.d of a, b if and only if $c > 0$ and satisfies*

(i) *$c \mid a$ and $c \mid b$.*
(ii) *If $r \mid a$ and $r \mid b$, then $r \mid c$.*

Proof. If c satisfies the above conditions, then for any common divisor r of a, b we have $r \mid c$, so $r \leq |r| \leq |c| = c$, as wanted. On the other hand, suppose c is the g.c.d of a, b. Then we have $c \mid a$ and $c \mid b$; so $a = a'c$ and $b = b'c$ for some a', b'. Thus we get $a = (-a')(-c)$ and $b = (-b')(-c)$, and therefore $-c \mid a$ and $-c \mid b$. Hence we must have $-c \leq c$. This implies $c \geq 0$, since otherwise we would have

$$0 = -c + c \leq c + c = 2c < 0,$$

which is a contradiction. However, $c \neq 0$ because it is a divisor of a, b and they are nonzero; so $c > 0$. Thus we only need to show that if r is a common divisor of a, b then $r \mid c$. But we have already checked this in the proof of Euclidean algorithm (this also follows from the next theorem, since $r \mid ax + by = c$). ■

Theorem 5.23. *Let $a, b \in \mathbb{Z}$ be nonzero, and let c be their greatest common divisor. Then we have*

$$c = ax + by$$

for some $x, y \in \mathbb{Z}$.

Remark. The above relation is also known as the *Bézout's identity.*

Remark. Note that x, y in the above relation are not unique. For example, the g.c.d of $7, 3$ is 1, and we have

$$1 = 7 \times 4 - 3 \times 9 = 7 \times 7 - 3 \times 16.$$

Proof. Consider the sequence of divisions in the Euclidean algorithm

$$r_{j-2} = r_{j-1}q_j + r_j$$

for $j \geq 1$ (recall that we set $r_{-1} = a$ and $r_0 = b$). We know that $c = r_n$, and $r_{n+1} = 0$, i.e. $r_{n-1} = r_n q_{n+1}$. We claim that for each $j \leq n$ we have

$$r_j = ax_j + by_j$$

for some $x_j, y_j \in \mathbb{Z}$. For $j = -1, 0$ we have $r_{-1} = a = a1 + b0$ and $r_0 = b = a0 + b1$. Suppose $j \geq 1$ and the claim holds for every $k < j$. Then we have

$$\begin{aligned} r_{j-2} = r_{j-1}q_j + r_j \implies r_j &= r_{j-2} - q_j r_{j-1} \\ &= ax_{j-2} + by_{j-2} - q_j(ax_{j-1} + by_{j-1}) \\ &= a(x_{j-2} - q_j x_{j-1}) + b(y_{j-2} - q_j y_{j-1}), \end{aligned}$$

as desired. Hence $c = r_n$ can be written in the desired way too. ■

Second Proof. Consider the set

$$A := \{ax + by : x, y \in \mathbb{Z} \text{ and } ax + by > 0\}.$$

First note that $A \neq \emptyset$ because for some $u \in \{-1, 1\}$ we have $|a| = au + b0 \in A$ (note that $|a| > 0$ as a is nonzero). Let d be the least element of A. Note that by definition $d > 0$, and $d = ax_0 + by_0$ for some integers x_0, y_0. Also, if $r \mid a$ and $r \mid b$ then $r \mid ax_0 + by_0 = d$. Hence, by the previous proposition, we only need to check that d is a divisor of both a and b, in order to conclude that $d = c$ is the g.c.d of a, b. By the division algorithm we have $a = dq_0 + r_0$ with $0 \leq r_0 < d$. Now we obtain

$$r_0 = a - dq_0 = a - q_0(ax_0 + by_0) = a(1 - q_0x_0) + b(-q_0y_0).$$

So if $r_0 > 0$ then by definition we have $r_0 \in A$. However, then we would have a contradiction, since $r_0 < d$ while d is the least element of A. Therefore we must have $r_0 = 0$, and thus $d \mid a$. We can similarly show that $d \mid b$. ■

Definition 5.21. A positive integer $p > 1$ is called **prime** if its only divisors are $\pm 1, \pm p$. A positive integer greater than 1 which is not a prime is called **composite**.

Remark. Note that any integer $a \neq 0, \pm 1$ has at least the four divisor $\pm 1, \pm a$, since $a = (\pm 1)(\pm a)$. So a prime number is an integer greater than 1 that does not have any other divisor. On the other hand, a composite number has at least one more divisor.

***Example* 5.4.** $p = 2$ is a prime, since if $r \mid 2$ then $|r| \leq 2$ and we must have $-2 \leq r \leq 2$. But we know that $r \neq 0$, so r can only be $-2, -1, 1, 2$. In fact, 2 is the only even prime. Because any other even integer $n > 1$ satisfies $n = 2k$; thus $2 \mid n$, and hence n is composite if $n \neq 2$. Therefore all the other primes are odd numbers. It can be checked that the first few primes are $2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots$.

Proposition 5.9. *Let $n > 1$ be a positive integer. Then n is composite if and only if there are positive integers $a, b < n$ such that $n = ab$.*

Remark. In fact we have $1 < a, b < n$, as will be shown in the following proof.

Proof. If n is composite then by definition there is $a \mid n$ with $a \neq \pm 1, \pm n$. Then we have $n = ab$. Since we also have $n = (-a)(-b)$, we can assume $a > 0$ (otherwise we can work with $-a$). Then $b > 0$ too, since $n > 0$. We know that $a \neq 1, n$; so $1 < a < n$. Hence $b = 1b < ab = n$ too. Conversely, suppose $n = ab$, where $0 < a, b < n$. Then $a, b > 1$, because otherwise we would have $a = 1$ or $b = 1$, and thus, contrary to our assumption, we get $n = a$ or $n = b$. Hence $1 < a, b < n$. But $a, b \mid n$; so n has divisors other than $\pm 1, \pm n$. Therefore n is composite. ■

Proposition 5.10. *Let $a, b, c \in \mathbb{Z}$ be nonzero. Suppose $a \mid bc$ and the g.c.d of a, b is 1. Then we must have $a \mid c$.*

Proof. We know that there are $x, y \in \mathbb{Z}$ such that $ax + by = 1$. Then we get $cax + cby = c$. Now $a \mid a$ and $a \mid bc$, so $a \mid acx + bcy = c$, as desired. ■

Euclid's Lemma. *Suppose p is a prime number. Then for integers a, b we have*

$$p \mid ab \implies p \mid a \text{ or } p \mid b.$$

Remark. The contrapositive of the above lemma is that

$$p \nmid a \text{ and } p \nmid b \implies p \nmid ab.$$

By an easy induction we can show that if p is prime, then for $a_1, \ldots, a_n \in \mathbb{Z}$ we have

$$p \nmid a_1 \text{ and } p \nmid a_2 \text{ and } \ldots \text{ and } p \nmid a_n \implies p \nmid a_1 \cdots a_n.$$

Equivalently, if $p \mid a_1 \cdots a_n$ then $p \mid a_i$ for some i.

Proof. If $p \nmid a$ then the g.c.d of p, a must be 1. To see this, let the g.c.d of p, a be d. Then we have $d \mid p$, and therefore d has one of the four values $\pm 1, \pm p$. But the g.c.d is positive, so d is either 1 or p. If $d = p$ then $p = d \mid a$, contrary to our assumption. So $d = 1$ as claimed. Hence by the previous proposition we must have $p \mid b$. ■

Proposition 5.11. *Let $a, b, c \in \mathbb{Z}$ be nonzero. Suppose $a \mid c$, $b \mid c$, and the g.c.d of a, b is 1. Then we have $ab \mid c$.*

Proof. We know that $c = br$ for some $r \in \mathbb{Z}$. Now $a \mid c = br$, and the g.c.d of a, b is 1; so we must have $a \mid r$. Thus $ab \mid rb = c$. ■

Fundamental Theorem of Arithmetic. *Every nonzero integer other than ± 1 can be written uniquely as a product of prime integers times ± 1. In other words, for all $a \in \mathbb{Z} - \{0, \pm 1\}$ we have*

(i) *There are prime numbers $p_1, \ldots, p_n$ and $u \in \{\pm 1\}$ such that the integer a has the* ***factorization***

$$a = up_1 \cdots p_n.$$

(ii) *If there is another factorization of a into primes $a = u'q_1 \cdots q_m$, then $m = n$, $u' = u$, and there is a permutation $\sigma \in S_n$ such that $p_i = q_{\sigma(i)}$.*

Proof. The uniqueness of a factorization is a consequence of Euclid's lemma. Suppose

$$up_1 \cdots p_n = u'q_1 \cdots q_m,$$

where p_i, q_j's are primes and $u, u' \in \{\pm 1\}$. First note that since every prime is positive, we must have $u' = u$ as both sides of the above equality must have the same sign. Therefore we get

$$p_1 \cdots p_n = q_1 \cdots q_m.$$

Now we proceed by induction on n. When $n = 1$ we have $p_1 \mid q_1 \cdots q_m$. Thus $p_1 \mid q_k$ for some k. But q_k is prime and $p_1 \neq \pm 1$ is positive, hence $p_1 = q_k$. Therefore

$$1 = q_1 \cdots q_{k-1} q_{k+1} \cdots q_m.$$

Hence if $m > 1$, the other q_j's must be divisors of 1, i.e. ± 1, which is a contradiction. Thus $m = 1$ and $p_1 = q_1$.

Now suppose the uniqueness holds for some n. Consider the case of $n + 1$, and suppose we have

$$p_1 \cdots p_n p_{n+1} = q_1 \cdots q_m.$$

Then $p_{n+1} \mid q_1 \cdots q_m$, and therefore $p_{n+1} \mid q_k$ for some k. We can argue as above and conclude that $p_{n+1} = q_k$, and hence

$$p_1 \cdots p_n = q_1 \cdots q_{k-1} q_{k+1} \cdots q_m.$$

Now by the induction hypothesis we have $n = m - 1$, and there is a permutation $\sigma \in S_n$ such that for $i \leq n$, $p_i = q_{\sigma(i)}$ when $\sigma(i) < k$, and $p_i = q_{\sigma(i)+1}$ when

$\sigma(i) \geq k$ (we can change the index of q_j for $k+1 \leq j \leq m = n+1$ to $j-1$, use the induction hypothesis, and then change the indices back). It is easy to see that

$$\hat{\sigma}(i) := \begin{cases} \sigma(i) & i \leq n,\ \sigma(i) < k \\ k & i = n+1 \\ \sigma(i)+1 & i \leq n,\ \sigma(i) \geq k \end{cases}$$

is a permutation in S_{n+1}. Then $p_i = q_{\hat{\sigma}(i)}$ for all $i \leq n+1$ as desired.

Next, for the existence of a factorization, let a be a nonzero integer other than ± 1. The proof is by strong induction on $|a| \geq 2$. If $|a| = 2$ then $a = \pm 2$ obviously has a factorization into primes. Now suppose every integer with absolute value less than $|a|$ has a factorization into primes. If $|a| = p$ for some prime p, then $a = \pm p$ has a factorization. Otherwise, $|a|$ and hence a, has a divisor other than $\pm 1, \pm a$, which we call b. Then we have $a = bc$, where b, c are nonzero, since $a \neq 0$. In addition, $c \neq \pm 1$ since $b \neq \pm a$; so $|c| > 1$. Hence we have $|b| < |b||c| = |bc| = |a|$. Similarly $|c| < |bc| = |a|$. Therefore by the induction hypothesis b, c have factorizations into primes. Now if we multiply those expressions we obtain a factorization of a into primes, as desired. ∎

Second Proof. Let us give another proof for the existence of a factorization. Suppose to the contrary that there are integers other than $0, \pm 1$ that have no factorization into primes. Consider the set

$$A := \{|a| : a \in \mathbb{Z} - \{0, \pm 1\},\ a \text{ has no factorization into primes}\},$$

which by our assumption is nonempty. Then A is a nonempty set of positive integers, and hence has a least element, which we denote by $|a|$. Now note that $|a| > 1$ cannot be a prime, because otherwise a would have the trivial factorization $a = u|a|$ for some $u \in \{\pm 1\}$. Thus $|a|$ is composite and we must have $|a| = bc$ for some positive integers $1 < b, c < |a|$. Then since a has the least absolute value among the members of A, the positive integers b, c cannot belong to A. Hence b, c must have factorizations into primes. Now if we multiply the factorizations of b, c with a suitable $u \in \{\pm 1\}$ we obtain a factorization of $a = u|a| = ubc$ into primes, contrary to our assumption. Therefore A must be empty, and so every integer other than $0, \pm 1$ has a factorization into primes, as desired. ∎

5.4 Rational Numbers

The next step in constructing numbers, is to build rational numbers. As we explained several times, the inherent nature of numbers is not very important. So, similarly to the case of natural numbers and integers, we just need to find a set whose elements represent the rational numbers, and have the properties that we

expect from the rational numbers. The basic idea in our construction is to consider a rational number as the fraction of two integers. Similarly to the construction of integers, we store the information of the fraction of two integers by the ordered pair of those integers. However, there are many other ordered pairs that can represent the same rational number; for example $\frac{1}{2}$ and $\frac{7}{14}$ represent the same rational number. To overcome this ambiguity, we can identify all ordered pairs of integers that represent the same rational number, by using an equivalence relation. In this construction, we will only use the fact that $\mathbb{Z}$ is an integral domain; and the other properties of integers are not employed. Hence, we will present the construction for an arbitrary integral domain; so that we can apply it to more general integral domains, like the set of polynomials.

Definition 5.22. Let R be an integral domain, and let

$$X = R \times (R - \{0\}) = \{(a, b) : a, b \in R,\ b \neq 0\}.$$

We define the relation $\sim$ on X as follows

$$(a, b) \sim (c, d) \qquad \text{if} \qquad ad = cb.$$

Remark. Note that informally, $ad = cb$ is equivalent to $\frac{a}{b} = \frac{c}{d}$. Thus the two pairs (a, b) and (c, d) are related by $\sim$, if and only if they represent the same fraction.

Theorem 5.24. *The relation $\sim$ is an equivalence relation on X.*

Proof. First note that $(a, b) \sim (a, b)$, since $ab = ab$. Also if $(a, b) \sim (c, d)$ then $ad = cb$; hence we have $cb = ad$, which means $(c, d) \sim (a, b)$. So far we have shown that $\sim$ is reflexive and symmetric. Let us show that it is transitive too. Suppose $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then we have $ad = cb$ and $cf = ed$. If we multiply the first equation by f, and use the second equation, we get

$$adf = cbf = bcf = bed.$$

Hence we have

$$d(af - be) = daf - dbe = adf - bed = 0.$$

But $d \neq 0$, and R is an integral domain, thus $af = be = eb$. Therefore $(a, b) \sim (e, f)$, as desired. ∎

Let F be the set of equivalence classes of $\sim$. We denote the equivalence class of (a, b) by $[(a, b)]$. So, informally, $[(a, b)]$ is the fraction $\frac{a}{b}$. Note that all the other pairs in the equivalence class $[(a, b)]$ represent the same fraction.

Next, we have to define the addition and multiplication on F, and show that they have the expected properties. Let $[(a, b)], [(c, d)] \in F$. Informally, these classes

represent $\frac{a}{b}$ and $\frac{c}{d}$, respectively. Thus if we want to add and multiply them, the results should be

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + cb}{bd},$$
$$\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

The following theorem and Theorem 5.2 show that if we define addition and multiplication as above, they are well defined.

Theorem 5.25. *Let $(a, b), (a', b'), (c, d), (c', d') \in X$. Suppose $(a, b) \sim (a', b')$, and $(c, d) \sim (c', d')$. Then we have*

(i) $(ad + cb, bd) \sim (a'd' + c'b', b'd')$.
(ii) $(ac, bd) \sim (a'c', b'd')$.
(iii) $(-a, b) \sim (-a', b')$.
(iv) *If $a \neq 0$ then $a' \neq 0$, and we have $(b, a) \sim (b', a')$.*

Remark. Notice that $bd, b'd' \neq 0$, since $b, d, b', d' \neq 0$, and R is an integral domain. Thus all the above pairs belong to X, and therefore, they can be related by $\sim$.

Proof. We know that

$$ab' = a'b, \qquad cd' = c'd. \tag{$*$}$$

(i) From the above two equations we can conclude that

$$ab'dd' = a'bdd', \qquad bb'cd' = bb'c'd.$$

If we add these equations we get $ab'dd' + bb'cd' = a'bdd' + bb'c'd$. Hence

$$(ad + cb)b'd' = ab'dd' + bb'cd' = a'bdd' + bb'c'd = (a'd' + b'c')bd,$$

as desired.

(ii) If we multiply the equations of $(*)$ we get $ab'cd' = a'bc'd$. Thus we have $acb'd' = a'c'bd$, as desired.

(iii) We have $(-a)b' = -ab' = -a'b = (-a')b$.

(iv) Since $a \neq 0$ and $b' \neq 0$, we have $ab' \neq 0$. Thus $a'b \neq 0$ too. Hence we must have $a' \neq 0$. Now we have $ba' = a'b = ab' = b'a$, as desired. ■

Definition 5.23. Let $[(a, b)], [(c, d)] \in F$. Then we define their **addition** and **multiplication** as

$$[(a, b)] + [(c, d)] := [(ad + cb, bd)],$$
$$[(a, b)][(c, d)] := [(ac, bd)],$$

respectively. The **zero** and **identity** of F are

$$0 := [(0,1)],$$
$$1 := [(1,1)],$$

respectively. The **opposite** of $[(a,b)]$ is

$$-[(a,b)] := [(-a,b)].$$

And if $a \neq 0$, then the **inverse** of $[(a,b)]$ is

$$[(a,b)]^{-1} := [(b,a)].$$

Remark. It is easy to see that in F we have $0 \neq 1$; because in R we have

$$0 \cdot 1 = 0 \neq 1 = 1 \cdot 1,$$

since R is an integral domain. Also, note that by Theorem 3.21 and the above theorem, the opposite and inverse are well defined.

Theorem 5.26. *F equipped with the above operations is a field, i.e. for every $p, q, r \in F$ we have*

(i) *Associativity* :

$$p + (q + r) = (p + q) + r, \qquad p(qr) = (pq)r.$$

(ii) *Commutativity* :

$$p + q = q + p, \qquad pq = qp.$$

(iii) *Identity elements* :

$$p + 0 = p, \qquad p1 = p.$$

(iv) *Additive and multiplicative inverses* :

$$p + (-p) = 0,$$
$$p \neq 0 \implies pp^{-1} = 1.$$

(v) *Distributivity* :

$$p(q + r) = pq + pr.$$

Proof. Suppose $p = [(a,b)]$, $q = [(c,d)]$, and $r = [(e,f)]$.

(i) We have

$$\begin{aligned} p + (q + r) &= \big[(a,b)\big] + \big[(cf + ed, df)\big] = \big[(adf + (cf + ed)b, bdf)\big] \\ &= \big[(adf + cfb + edb, bdf)\big] = \big[((ad + cb)f + ebd, bdf)\big] \\ &= \big[(ad + cb, bd)\big] + \big[(e,f)\big] = (p + q) + r. \end{aligned}$$

We also have

$$p(qr) = [(a,b)][(ce,df)] = [(ace,bdf)] = [(ac,bd)][(e,f)] = (pq)r.$$

(ii) We have

$$p+q = [(ad+cb,bd)] = [(cb+ad,db)] = q+p.$$

We also have $pq = [(ac,bd)] = [(ca,db)] = qp$.

(iii) We have $p+0 = [(a1+0b,b1)] = [(a,b)] = p$. Also

$$p1 = [(a1,b1)] = [(a,b)] = p.$$

(iv) We have

$$p+(-p) = [(ab+(-a)b,bb)] = [(0,b^2)] = [(0,1)] = 0.$$

Note that $(0,b^2) \sim (0,1)$, since $0 \times 1 = 0 = 0 \times b^2$. Next, suppose $p \neq 0$. Then $a \neq 0$. Now we have

$$pp^{-1} = [(ab,ba)] = [(1,1)] = 1;$$

because $(ab,ba) \sim (1,1)$, since $ab1 = 1ba$.

(v) We have

$$p(q+r) = \big[(a,b)\big]\big[(cf+ed,df)\big] = \big[(a(cf+ed),bdf)\big] = \big[(acf+aed,bdf)\big].$$

On the other hand we have

$$\begin{aligned} pq+pr &= \big[(ac,bd)\big] + \big[(ae,bf)\big] \\ &= \big[(acbf+aebd,bdbf)\big] = \big[((acf+aed)b,(bdf)b)\big]. \end{aligned}$$

However, for every $(g,h) \in X$ we have $(gb,hb) \sim (g,h)$. Because $gbh = hbg$. Also note that $(gb,hb) \in X$, since $b \neq 0$. Thus we have $p(q+r) = pq+pr$, as desired. ■

Definition 5.24. F is called the **field of fractions** of R.

Definition 5.25. The set of **rational numbers**, denoted by $\mathbb{Q}$, is the field of fractions of the integral domain $\mathbb{Z}$.

Remark. Note that the field of fractions of any integral domain is a field, as shown in the above theorem. In particular, $\mathbb{Q}$ is a field.

Next, we want to define the order relation on rational numbers. Let $[(a,b)],[(c,d)] \in \mathbb{Q}$. Informally, these numbers represent $\frac{a}{b}$ and $\frac{c}{d}$, respectively. Thus if we want $[(a,b)]$ to be less than $[(c,d)]$, we must have $\frac{a}{b} < \frac{c}{d}$; or equivalently

$ad < cb$, if we assume that $b, d > 0$. This restriction is not problematic, because it is easy to see that we always have

$$[(a, b)] \sim [(-a, -b)].$$

So we can always assume that the second component of a pair of integers that represents a rational number is positive. Informally, we can always assume that the denominator of a fraction is positive.

Definition 5.26. Let $p, q \in \mathbb{Q}$. Then we say $p < q$ if there are $(a, b) \in p$ and $(c, d) \in q$ such that $b, d > 0$, and

$$ad < cb.$$

Note that every rational number is a set of pairs of integers, i.e. it is the equivalence class of pairs of integers which represent it. Thus the above definition makes sense.

Also, note that $<$ is a relation on $\mathbb{Q}$; and, as we explained after the definition of order of integers, we do not need to check that $<$ is well defined, in order to be sure that it is indeed a relation. However, the next theorem shows that $<$ is compatible with the equivalence relation $\sim$.

Theorem 5.27. *Let $(a, b), (a', b'), (c, d), (c', d') \in \mathbb{Z} \times (\mathbb{Z} - \{0\})$. Suppose $(a, b) \sim (a', b')$, and $(c, d) \sim (c', d')$. Also suppose that $b, b', d, d' > 0$. Then we have*

$$ad < cb \qquad \iff \qquad a'd' < c'b'.$$

Proof. We know that

$$ab' = a'b, \qquad cd' = c'd.$$

Suppose $ad < cb$. Then we have $adb'd' < cbb'd'$, since $b'd' > 0$. Hence by applying the above equations we get

$$a'bdd' = ab'dd' < cd'bb' = c'dbb'.$$

Thus $a'd'bd < c'b'bd$. This implies that $a'd' < c'b'$. Because otherwise we would have $a'd' \geq c'b'$. However, we know that $bd > 0$; so we would get $a'd'bd \geq c'b'bd$, which is a contradiction. Hence we must have $a'd' < c'b'$ as desired. The reverse implication follows similarly. ∎

Theorem 5.28. *The relation $<$ on $\mathbb{Q}$ is a linear order. In addition, $\mathbb{Q}$ does not have a smallest element, nor a largest element. Furthermore, for every $p, q \in \mathbb{Q}$ where $p < q$, there is $r \in \mathbb{Q}$ such that $p < r < q$.*

Proof. Let $p = [(a, b)]$, $q = [(c, d)]$, and $s = [(e, f)]$. Suppose $b, d, f > 0$. First note that $p \not< p$, since $ab \not< ab$. So $<$ is irreflexive. Now suppose $p < q$ and $q < s$. Then we have

$$ad < cb, \qquad cf < ed.$$

Hence we get

$$adf < cbf, \qquad cfb < edb,$$

since $f, b > 0$. Thus $afd = adf < edb = ebd$. This implies that $af < eb$. Because otherwise we would have $af \geq eb$. However, we know that $d > 0$; so we would get $afd \geq ebd$, which is a contradiction. Hence we must have $af < eb$ as desired. So $<$ is also transitive.

Next suppose $p \neq q$. Then $ad \neq cb$. Then since the order of integers is a total order, we either have $ad < cb$, or $ad > cb$. Hence either $p < q$, or $p > q$. So $<$ is a total order on $\mathbb{Q}$. In addition, it is easy to see that when $b > 0$ we have

$$[(a - 1, b)] < [(a, b)] < [(a + 1, b)].$$

Therefore no rational number like $p = [(a, b)]$ can be the smallest element, nor the largest element of $\mathbb{Q}$.

Finally, suppose $p < q$, i.e. $ad < cb$. Let $r := [(a+c, b+d)]$. Note that $b+d > 0$, since $b, d > 0$. Now we have

$$a(b + d) = ab + ad < ab + cb = (a + c)b,$$
$$(a + c)d = ad + cd < cb + cd = c(b + d).$$

Therefore $p < r < q$, as desired. ■

Theorem 5.29. *Let $p, q, r \in \mathbb{Q}$. Then we have*

(i) *If $p < q$ then $p + r < q + r$.*

(ii) *If $r > 0$ and $p < q$, then $pr < qr$.*

Remark. As a consequence, note that if $p, q > 0$ then $pq > 0$. Because we have $pq > p0 = 0$.

Proof. Suppose $p = [(a, b)]$, $q = [(c, d)]$, $r = [(e, f)]$, and $b, d, f > 0$. Also suppose $p < q$, so we have $ad < cb$.

(i) We have $p + r = [(af + eb, bf)]$, and $q + r = [(cf + ed, df)]$. Now we have

$$(af + eb)df = adf^2 + ebdf < cbf^2 + ebdf = (cf + ed)bf,$$

since $f^2 > 0$. Thus $p + r < q + r$, as desired.

(ii) We have $pr = [(ae, bf)]$, and $qr = [(ce, df)]$. We also know that $r > 0$. Thus $e1 > 0f$, i.e. $e > 0$. Now we have

$$aedf = adef < cbef = cebf,$$

since $ef > 0$. Hence $pr < qr$, as desired. ■

Although $\mathbb{Q}$ does not contain $\mathbb{Z}$ as a subset, it has a subset which looks like $\mathbb{Z}$. Informally, we can say that $\mathbb{Q}$ contains a "copy" of $\mathbb{Z}$. This copy of $\mathbb{Z}$ is the set

$$\{[(a,1)] \in \mathbb{Q} : a \in \mathbb{Z}\}.$$

The next theorem shows that the above subset of $\mathbb{Q}$ behaves similarly to $\mathbb{Z}$.

Theorem 5.30. *Let $E : \mathbb{Z} \to \mathbb{Q}$ be defined as $E(a) = [(a,1)]$, for every $a \in \mathbb{Z}$. Then E is a one-to-one function, and for every $a, b \in \mathbb{Z}$ we have*

(i) $E(a+b) = E(a) + E(b)$.
(ii) $E(ab) = E(a)E(b)$.
(iii) $a < b$ *if and only if* $E(a) < E(b)$.
(iv) *If* $b \neq 0$ *we have* $[(a,b)] = E(a)/E(b)$.

Remark. Note that in the relation $E(a+b) = E(a) + E(b)$, a, b are added using the addition of $\mathbb{Z}$, and $E(a), E(b)$ are added using the addition of $\mathbb{Q}$. So in some sense, we can say that the function E transforms the addition of $\mathbb{Z}$ into the addition of $\mathbb{Q}$. Similar remarks apply to the multiplication and order relation.

Remark. Also note that the last part of the theorem confirms our initial intuition that $[(a,b)]$ represents the rational number which is the ratio of the integers a, b. Remember that for two element p, q in a field with $q \neq 0$, we have $p/q := pq^{-1}$. Also, keep in mind that we identify $E(a)$ with a.

Remark. This theorem, except the part (iii) on order relation, is also true if we replace $\mathbb{Z}$ by an integral domain R, and replace $\mathbb{Q}$ by the field of fractions of R. The proof for the general case is the same as the proof for $\mathbb{Z}$.

Proof. First note that E is one-to-one. Because if $E(a) = E(b)$ then we have $[(a,1)] = [(b,1)]$, which means that $a1 = b1$. So we get $a = b$.

(i) We have

$$\begin{aligned} E(a) + E(b) &= [(a,1)] + [(b,1)] \\ &= [(a1 + b1, 1 \times 1)] = [(a+b,1)] = E(a+b). \end{aligned}$$

(ii) We have

$$E(a)E(b) = [(a,1)][(b,1)] = [(ab, 1 \times 1)] = [(ab,1)] = E(ab).$$

(iii) If $a < b$ then $E(a) < E(b)$, since $a1 < b1$. Conversely, if $E(a) < E(b)$ then by definition we must have $a1 < b1$; so $a < b$.

(iv) We have

$$\begin{aligned} [(a,b)] &= [(a,1)][(1,b)] \\ &= [(a,1)][(b,1)]^{-1} = E(a)E(b)^{-1} = E(a)/E(b). \end{aligned}$$

■

5.5 Ordered Fields

Let us first review the notion of field, and its basic properties.

Definition 5.27. A **field** is a nonempty set F equipped with two binary operations

$$\begin{array}{c} F \times F \longrightarrow F \\ (a,b) \mapsto a+b \end{array}, \qquad \begin{array}{c} F \times F \longrightarrow F \\ (a,b) \mapsto ab \end{array},$$

called respectively **addition** and **multiplication**, such that

(i) The operations are *associative* and *commutative*, i.e. for every $a, b, c \in F$

$$a+(b+c) = (a+b)+c, \qquad a(bc) = (ab)c,$$
$$a+b = b+a, \qquad ab = ba.$$

(ii) There exist distinct elements $0, 1 \in F$, called respectively *zero* and *identity* of F, such that for every $a \in F$

$$a+0 = a, \qquad a1 = a.$$

(iii) For every $a \in F$ there exists an element $-a \in F$, called the **opposite** of a, such that

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

(iv) For every $a \in F - \{0\}$ there exists an element $a^{-1} \in F$, called the **inverse** of a, such that

$$aa^{-1} = 1.$$

(v) Multiplication is *distributive* over addition, i.e. for every $a, b, c \in F$

$$a(b+c) = ab + ac.$$

Remark. We have seen that the zero and identity of a field are unique. Also, the opposite of every element, and the inverse of every nonzero element are uniquely determined by that element.

Let a, b be two elements in a field F. Then $a+b$ is called the *sum* of a, b. And a, b are called the *summands* of $a+b$. Also, ab is called the *product* of a, b; and a, b are called the *factors* of ab. We sometimes denote the product of two elements a, b by $a \cdot b$, or $a \times b$. In addition, the *square* and the *cube* of an element a are respectively defined as

$$a^2 := aa, \qquad a^3 := a^2 a = aaa.$$

The inverse of a is also called the *reciprocal* of a.

Furthermore, the **subtraction** and the **division** of two elements $a, b \in F$ are respectively defined as follows

$$a - b := a + (-b), \qquad a/b = \frac{a}{b} := ab^{-1} \text{ when } b \neq 0.$$

The element $a - b$ is called the *difference* of a, b. The element $\frac{a}{b}$ is called the *quotient* or the *ratio* of a, b. We also call $\frac{a}{b}$ a *fraction*. In a fraction $\frac{a}{b}$, a is called the *numerator*, and b is called the *denominator*. The *reciprocal* of the fraction $\frac{a}{b}$ is the fraction $\frac{b}{a}$, provided that a is also nonzero. We will see that $\frac{b}{a}$ is the inverse of $\frac{a}{b}$; so the two uses of the term "reciprocal" are compatible.

Remark. Informally, a field is a structure in which we can perform the four basic arithmetic operations, i.e. addition, subtraction, multiplication, and division.

***Example* 5.5.** $\mathbb{Q}$ is a field.

Theorem 5.31. *Let F be a field. Then for every $a, b, c, d \in F$ we have*

(i) **(Cancellation Laws)**

$$a + c = b + c \implies a = b,$$
$$ac = bc,\ c \neq 0 \implies a = b.$$

(ii) $0a = 0$. *And* $ab = 0 \implies a = 0$ *or* $b = 0$.
(iii) $-(-a) = a$, *and* $-(a + b) = (-a) + (-b) = -a - b$.
(iv) *If* $a \neq 0$ *then* $(a^{-1})^{-1} = a$. *And if* $a, b \neq 0$ *then* $(ab)^{-1} = a^{-1}b^{-1}$.
(v) $(-a)b = -ab = a(-b)$, *and* $(-a)(-b) = ab$.
(vi) $-a = (-1)a$, $-(b - c) = c - b$, *and* $a(b - c) = ab - ac$.
(vii) *If* $a \neq 0$ *then* $a^{-1} = \frac{1}{a}$, *and* $(-a)^{-1} = -a^{-1}$.
(viii) *If* $b, d \neq 0$ *then*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

(ix) *If* $b, c, d \neq 0$ *we have*

$$-\frac{a}{b} = \frac{-a}{b} = \frac{a}{-b}, \qquad \left(\frac{c}{d}\right)^{-1} = \frac{d}{c}, \qquad \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}.$$

Proof. **(i, ii, iii, iv, v)** These are proved in Section 5.2.

(vi) We have $(-1)a + a = ((-1) + 1)a = 0a = 0$. Therefore $(-1)a = -a$, since the additive inverse is unique. We also have

$$-(b - c) = -(b + (-c)) = -b + (-(-c)) = -b + c = c - b,$$
$$a(b - c) = a(b + (-c)) = ab + a(-c) = ab + (-ac) = ab - ac.$$

(vii) By the definition of division, for $a \neq 0$ we have $\frac{1}{a} = 1a^{-1} = a^{-1}$. Also, $(-a^{-1})(-a) = a^{-1}a = 1$, and we get the desired by uniqueness of inverse.

(viii) By the definition of division we have

$$\begin{aligned}\frac{ad+bc}{bd} &= (ad+bc)(bd)^{-1} = (ad+bc)b^{-1}d^{-1} \\ &= adb^{-1}d^{-1} + bcb^{-1}d^{-1} = ab^{-1} + cd^{-1} = \frac{a}{b} + \frac{c}{d}, \\ \frac{a}{b} \cdot \frac{c}{d} &= (ab^{-1})(cd^{-1}) = acb^{-1}d^{-1} = ac(bd)^{-1} = \frac{ac}{bd}.\end{aligned}$$

(ix) We have

$$\begin{aligned}\frac{-a}{b} &= (-a)b^{-1} = -ab^{-1} = -\frac{a}{b}, \\ \frac{a}{-b} &= a(-b)^{-1} = a(-b^{-1}) = -ab^{-1} = -\frac{a}{b}.\end{aligned}$$

We also have

$$\begin{aligned}(\frac{c}{d})^{-1} &= (cd^{-1})^{-1} = c^{-1}(d^{-1})^{-1} = c^{-1}d = \frac{d}{c}, \\ \frac{\frac{a}{b}}{\frac{c}{d}} &= (\frac{a}{b})(\frac{c}{d})^{-1} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}.\end{aligned}$$

■

Let us also mention the interesting problem of solving *polynomial equations* in a field F. Let $a, b, c \in F$, and suppose $a \neq 0$. Then for any $x \in F$ we have

$$ax + b = 0 \iff ax = -b \iff x = a^{-1}(-b) = -\frac{b}{a}.$$

Hence we can easily solve the *linear equation* $ax + b = 0$. More interestingly, consider the *quadratic equation*

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

Let $\Delta := b^2 - 4ac$. Then Δ is called the **discriminant** of the above quadratic equation. Now we have

$$\begin{aligned}ax^2 + bx + c &= a\big(x^2 + 2\frac{b}{2a}x + \frac{c}{a}\big) \\ &= a\big(x^2 + 2\frac{b}{2a}x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} + \frac{c}{a}\big) \\ &= a\Big(\big(x + \frac{b}{2a}\big)^2 - \frac{b^2 - 4ac}{4a^2}\Big) = a\Big(\big(x + \frac{b}{2a}\big)^2 - \frac{\Delta}{4a^2}\Big).\end{aligned}$$

Hence $ax^2 + bx + c = 0$ if and only if $(x + \frac{b}{2a})^2 = \frac{\Delta}{4a^2}$. Note that for $y, z \in F$ we have

$$y^2 = z^2 \iff (y - z)(y + z) = y^2 - z^2 = 0 \iff y = \pm z.$$

Now let us assume that for some $d \in F$ we have $d^2 = \Delta$. Then we get

$$ax^2 + bx + c = 0 \iff \big(x + \frac{b}{2a}\big)^2 = \big(\frac{d}{2a}\big)^2 \iff x = \frac{-b \pm d}{2a}.$$

If we use the familiar notation $\sqrt{\Delta}$ for d we obtain the famous *quadratic formula*

$$ax^2 + bx + c = 0 \iff x = \frac{-b \pm \sqrt{\Delta}}{2a}.$$

However, note that we do not say that Δ necessarily has a square root in F; and in fact this is not true in arbitrary fields.

Definition 5.28. An **ordered field** is a field F equipped with a linear order $<$, such that for every $a, b, c \in F$ we have

(i) If $a < b$ then $a + c < b + c$.
(ii) If $0 < a$ and $0 < b$, then $0 < ab$.

Notation. As usual, we define $a \leq b$ to mean $a < b$ or $a = b$. Similarly we define $a > b$ to mean $b < a$, and $a \geq b$ to mean $b \leq a$. We say a is **positive** if $a > 0$, and a is **negative** if $a < 0$. We also say a is *nonnegative* or *nonpositive*, if $a \geq 0$ or $a \leq 0$ respectively. The **sign** of a is its property of being positive, negative, or zero.

***Example* 5.6.** $\mathbb{Q}$ is an ordered field.

Theorem 5.32. *Suppose F is an ordered field. Then for every $a, b, c, d \in F$ we have*

(i) *$a < b$ if and only if $b - a > 0$.*
(ii) *$a > 0$ if and only if $-a < 0$.*
(iii) *If $a \neq 0$ then $a^2 > 0$. As a result $1 = 1^2 > 0$.*
(iv) *$a < b$ if and only if $a + c < b + c$.*
(v) *If $a < c$, $b \leq d$ then $a + b < c + d$.*
(vi) *If $c > 0$, then $a < b$ if and only if $ac < bc$.*
(vii) *If $c < 0$, then $a < b$ if and only if $ac > bc$.*
(viii) *If $a, b < 0$ then $ab > 0$.*
(ix) *If $0 \leq a < c$, $0 < b \leq d$ then $ab < cd$.*
(x) *If $a > 0$ then $a^{-1} > 0$.*
(xi) *If $0 < a < b$ then $a^{-1} > b^{-1} > 0$. In particular, $a > 1$ if and only if $0 < a^{-1} < 1$.*
(xii) *If $a \leq b$ and $a \geq b$ then $a = b$.*

(xiii) *If* $a_1, \ldots, a_n \geq 0$ *and* $a_1 + \cdots + a_n = 0$ *then* $a_i = 0$ *for all* i.

Proof. **(i)** We have

$$a < b \implies a + (-a) < b + (-a) \implies 0 < b - a.$$

Similarly, for the converse we can add a to both sides of $0 < b - a$.

(ii) By (i) we have $-a < 0 \iff 0 < 0 - (-a) = a$.

(iii) If $a \neq 0$ then either $0 < a$ or $a < 0$. Hence by (ii) we either have $0 < a$ or $0 < -a$. Thus either

$$0 < aa = a^2, \qquad \text{or} \qquad 0 < (-a)(-a) = a^2.$$

(iv) $a < b$ implies $a + c < b + c$ by definition. On the other hand, if $a + c < b + c$ then we cannot have $a \geq b$. Since in that case we would obtain $a + c \geq b + c$, which is a contradiction. So we must have $a < b$. Alternatively, we can add $-c$ to both sides of $a + c < b + c$ to get $a < b$.

(v) When $b = d$ then the claim holds by the definition of ordered fields. So suppose $b < d$. Then we add b to both sides of $a < c$, and add c to both sides of $b < d$, to obtain

$$a + b < c + b < c + d.$$

(vi) By (i) we have $0 < b - a$. Thus if $c > 0$ then

$$0 < (b - a)c = bc - ac \implies ac < bc.$$

Conversely, suppose $ac < bc$. Then we cannot have $a \geq b$, since in this case we would obtain $ac \geq bc$, which is a contradiction. So we must have $a < b$.

(vii) When $c < 0$ we have $-c > 0$. Hence

$$0 < (b - a)(-c) = -(bc - ac) = ac - bc \implies bc < ac.$$

Conversely, suppose $ac > bc$. Then we cannot have $a \geq b$, since in this case we would obtain $ac \leq bc$, which is a contradiction. So we must have $a < b$.

(viii) We multiply both sides of $a < 0$ by b to obtain $ab > 0b = 0$.

(ix) When $b = d$ then the claim holds by (v). So suppose $b < d$. Then we multiply both sides of $a < c$ by b, and both sides of $b < d$ by c, to obtain

$$ab < cb < cd.$$

(x) Note that $a^{-1} \neq 0$, since otherwise we would have $1 = aa^{-1} = a0 = 0$. Therefore if the claim does not hold we must have $a^{-1} < 0$. But this implies that $1 = aa^{-1} < a0 = 0$, which is a contradiction. So $a^{-1} > 0$.

(xi) Note that $a^{-1} \neq b^{-1}$, since $a \neq b$ and the inverse is unique. Also note that by (x) we have $a^{-1}, b^{-1} > 0$. Therefore if the claim does not hold we must

have $0 < a^{-1} < b^{-1}$. But then by (ix) we get $1 = aa^{-1} < bb^{-1} = 1$, which is a contradiction. The last statement follows easily since $1^{-1} = 1$.

(xii) We either have $a < b$, $a = b$, or $a > b$. But we cannot have $a < b$, since we know that $a \geq b$. Similarly we cannot have $a > b$. Therefore we must have $a = b$.

(xiii) We have $a_i = -\sum_{j \neq i} a_j \leq 0$, hence $a_i = 0$. ■

Remark. The following version of the above theorem can be proved easily from it. We only need to consider some trivial cases in which the inequalities become equalities.

(i) If $a \leq b$, $b \leq c$ then $a \leq c$.
(ii) $a \leq b$ if and only if $b - a \geq 0$.
(iii) $a \geq 0$ if and only if $-a \leq 0$.
(iv) $a^2 \geq 0$.
(v) $a \leq b$ if and only if $a + c \leq b + c$.
(vi) If $a \leq c$, $b \leq d$ then $a + b \leq c + d$.
(vii) If $a \leq b$ then

$$c \geq 0 \implies ac \leq bc,$$
$$c \leq 0 \implies ac \geq bc.$$

(viii) If $a, b \geq 0$, or $a, b \leq 0$, then $ab \geq 0$.
(ix) If $0 \leq a \leq c$, $0 \leq b \leq d$ then $ab \leq cd$.

Theorem 5.33. *The characteristic of an ordered field is zero.*

Proof. Let n be a positive integer. First remember that in a field F, the element n is defined to be

$$n := \overbrace{1 + 1 + \cdots + 1}^{n \text{ times}},$$

where 1 is the identity of F. By the last theorem, in an ordered field we have $n > 0$, since $1 > 0$. Thus in particular, in an ordered field we have $n \neq 0$. ■

Theorem 5.34. *Suppose F is an ordered field, and $a, b, c \in F$. If $a > 0$, then the quadratic expression*

$$f(x) = ax^2 + bx + c$$

is nonnegative for every $x \in F$ if and only if its discriminant $\Delta = b^2 - 4ac$ is nonnegative.

Proof. We have seen that for every $x \in F$ we have

$$ax^2 + bx + c = a\Big(\big(x + \frac{b}{2a}\big)^2 + \frac{-\Delta}{4a^2}\Big).$$

Note that the square of any element in an ordered field is nonnegative. Thus if $\Delta \leq 0$ we have

$$\big(x + \frac{b}{2a}\big)^2 + \frac{-\Delta}{4a^2} = \big(x + \frac{b}{2a}\big)^2 + (-\Delta)(4a^2)^{-1} \geq 0.$$

Hence $f(x) \geq 0$. On the other hand, if $\Delta > 0$ then we have $f(-\frac{b}{2a}) = -\frac{\Delta}{4a} < 0$. Therefore if $f(x) \geq 0$ for every x, then we must have $\Delta \leq 0$, as desired. ■

Theorem 5.35. *Let F be an ordered field. Then there exists a one-to-one function $\varphi : \mathbb{Q} \to F$ such that for every $p, q \in \mathbb{Q}$ we have*

(i) $\varphi(p + q) = \varphi(p) + \varphi(q)$,
(ii) $\varphi(pq) = \varphi(p)\varphi(q)$,
(iii) $p < q$ *if and only if* $\varphi(p) < \varphi(q)$.

Remark. Therefore F has a subset $\varphi(\mathbb{Q})$, which looks like $\mathbb{Q}$. Informally, we can say that F contains a "copy" of $\mathbb{Q}$.

Proof. For every $n \in \mathbb{N}$ let $\varphi(n) := n = \sum_{j \leq n} 1 \in F$, where 1 is the identity of F. Note that $\varphi(n) > 0$, since $1 > 0$ in F. Also, let $\varphi(0) := 0$, and $\varphi(-n) := -\varphi(n)$. Then as shown in Section 5.2, for every $n, m \in \mathbb{Z}$ we have $\varphi(n+m) = \varphi(n)+\varphi(m)$, and $\varphi(nm) = \varphi(n)\varphi(m)$. Now suppose $\varphi(n) = \varphi(m)$. If $n > m$ then we have $n - m = \varphi(n - m) = \varphi(n) - \varphi(m) = 0$, which contradicts the fact that the characteristic of F is zero. Similarly, we cannot have $n < m$; so $n = m$. Thus φ is one-to-one on $\mathbb{Z}$. Next let us define $\varphi(p)$ for $p \in \mathbb{Q}$. We know that $p = \frac{m}{n}$ for some $n, m \in \mathbb{Z}$ with $n \neq 0$. We define

$$\varphi(p) := \frac{\varphi(m)}{\varphi(n)} = \frac{m}{n} \in F.$$

Note that $n \neq 0$ in F, since n is nonzero in $\mathbb{Z}$, and the characteristic of F is nonzero. Now note that the value of $\varphi(p)$ does not depend on the representing fraction $\frac{m}{n}$. Because if we also have $p = \frac{m'}{n'}$ then $mn' = m'n$. Hence

$$\varphi(m)\varphi(n') = \varphi(mn') = \varphi(m'n) = \varphi(m')\varphi(n).$$

Therefore we get $\frac{\varphi(m)}{\varphi(n)} = \frac{\varphi(m')}{\varphi(n')}$. Note that $\varphi(n), \varphi(n') \neq 0$, since $n, n' \neq 0$. So, φ is a well-defined function from $\mathbb{Q}$ to F.

Next let us show that φ is one-to-one. Suppose $\varphi(\frac{m}{n}) = \varphi(\frac{k}{l})$. Then similarly to the above we get

$$\varphi(ml) = \varphi(m)\varphi(l) = \varphi(k)\varphi(n) = \varphi(kn).$$

Hence $ml = kn$, since φ is one-to-one on $\mathbb{Z}$. Thus $\frac{m}{n} = \frac{k}{l}$, as wanted. Now note that φ preserves the addition and multiplication of $\mathbb{Z}$. Therefore we get

$$\begin{aligned}\varphi\big(\frac{m}{n} + \frac{k}{l}\big) &= \varphi\big(\frac{ml + kn}{nl}\big) = \frac{\varphi(ml + kn)}{\varphi(nl)} = \frac{\varphi(ml) + \varphi(kn)}{\varphi(n)\varphi(l)} \\ &= \frac{\varphi(m)\varphi(l) + \varphi(k)\varphi(n)}{\varphi(n)\varphi(l)} = \frac{\varphi(m)}{\varphi(n)} + \frac{\varphi(k)}{\varphi(l)} = \varphi\big(\frac{m}{n}\big) + \varphi\big(\frac{k}{l}\big), \\ \varphi\big(\frac{m}{n} \cdot \frac{k}{l}\big) &= \varphi\big(\frac{mk}{nl}\big) = \frac{\varphi(mk)}{\varphi(nl)} = \frac{\varphi(m)\varphi(k)}{\varphi(n)\varphi(l)} = \frac{\varphi(m)}{\varphi(n)} \cdot \frac{\varphi(k)}{\varphi(l)} = \varphi\big(\frac{m}{n}\big) \cdot \varphi\big(\frac{k}{l}\big).\end{aligned}$$

Hence φ also preserves the addition and multiplication of $\mathbb{Q}$. Finally, suppose $\frac{m}{n} < \frac{k}{l}$, and $n, l > 0$. Then we have $ml < kn$; so $kn - ml > 0$. Therefore

$$0 < \varphi(kn - ml) = \varphi(k)\varphi(n) - \varphi(m)\varphi(l).$$

Also, $\varphi(n), \varphi(l) > 0$. Thus we get

$$\varphi\big(\frac{m}{n}\big) = \frac{\varphi(m)}{\varphi(n)} < \frac{\varphi(k)}{\varphi(l)} = \varphi\big(\frac{k}{l}\big),$$

as desired. Conversely, we can similarly show that $\frac{m}{n} \geq \frac{k}{l}$ implies $\varphi\big(\frac{m}{n}\big) \geq \varphi\big(\frac{k}{l}\big)$. Hence $\varphi\big(\frac{m}{n}\big) < \varphi\big(\frac{k}{l}\big)$ also implies that $\frac{m}{n} < \frac{k}{l}$. ■

5.6 Binary Operations

Definition 5.29. A **binary operation** on a set S is a function

$$\star : S \times S \to S.$$

For two elements $a, b \in S$, we usually write $a \star b$ instead of $\star(a, b)$.

Notation. In the rest of this section we assume that S is a set, and $\star$ is a binary operation on S.

Definition 5.30. A binary operation $\star$ on S is called **associative** if for every $a, b, c \in S$ we have

$$a \star (b \star c) = (a \star b) \star c,$$

and it is called **commutative** if for every $a, b \in S$ we have

$$a \star b = b \star a.$$

We also say two elements a, b commute if $a \star b = b \star a$. An element $e \in S$ is an **identity** if for every $a \in S$ we have

$$a \star e = a = e \star a.$$

Finally, a subset $A \subset S$ is said to be **closed** under $\star$ if for every $a, b \in A$ we have $a \star b \in A$.

Theorem 5.36. *A binary operation has at most one identity.*

Proof. If there exist two identities e, e' we have $e' = e' \star e = e$. Note that the first equality holds because e is an identity, and the second equality holds since e' is an identity. ■

Definition 5.31. Suppose $\star$ is a binary operation on S with identity e, and $a, b \in S$. We say a is **invertible,** and b is an **inverse** of a, if

$$a \star b = e = b \star a.$$

Theorem 5.37. *Suppose $\star$ is an associative operation with identity e, and $a, b \in S$.*

(i) *If a has an inverse, its inverse is unique, and we denote it by a^{-1}.*

(ii) *If a is invertible, then a^{-1} is also invertible and*

$$(a^{-1})^{-1} = a.$$

(iii) *If a and b are invertible, then $a \star b$ is also invertible and we have*

$$(a \star b)^{-1} = b^{-1} \star a^{-1}.$$

Proof. **(i)** Suppose that a has two inverses denoted by b and c, then

$$b = b \star e = b \star (a \star c) = (b \star a) \star c = e \star c = c.$$

(ii) This follows from the definition of a^{-1} and the uniqueness of inverse.

(iii) First note that

$$\begin{aligned}(b^{-1} \star a^{-1}) \star (a \star b) &= b^{-1} \star (a^{-1} \star (a \star b)) \\ &= b^{-1} \star ((a^{-1} \star a) \star b) = b^{-1} \star (e \star b) = b^{-1} \star b = e.\end{aligned}$$

Similarly $(a \star b) \star (b^{-1} \star a^{-1}) = e$. Therefore $(a \star b)$ is invertible. Now the result follows from uniqueness of inverse and the above computations. ■

Cancellation Law. *Suppose $\star$ is an associative operation with identity e. Let $a, b, c \in S$, and suppose a is invertible. Then we have*

(i) *If $a \star b = a \star c$ then $b = c$.*

(ii) *If $b \star a = c \star a$ then $b = c$.*

Proof. We have

$$\begin{aligned} a \star b = a \star c &\implies a^{-1} \star (a \star b) = a^{-1} \star (a \star c) \\ &\implies (a^{-1} \star a) \star b = (a^{-1} \star a) \star c \\ &\implies e \star b = e \star c \implies b = c. \end{aligned}$$

The other one is similar. ∎

Let us recall the notion of product of several elements from Section 4.5.

Definition 5.32. Suppose $\star$ is a binary operation on a set S, and $a_k, \ldots, a_m$ is a finite sequence in S. We inductively define the *(standard) product* of $a_k, \ldots, a_m$ to be

(i) $\prod_{j=k}^{k} a_j := a_k$,

(ii) $\prod_{j=k}^{n+1} a_j := \left(\prod_{j=k}^{n} a_j\right) \star a_{n+1}$ for $k \leq n < m$.

Remark. In the terminology of recursion theorem, we have constructed the function $f(n) = \prod_{j=k}^{n} a_j$, by using the function $F(n, s) = s \star a_{n+1}$. Note that the sequence $j \mapsto a_j$ is given to us; so we know a priori what a_{n+1} is.

Remark. We also denote $\prod_{j=k}^{n} a_j$ by $\prod_{k \leq j \leq n} a_j$. When k is 0 or 1, we may also denote it simply by $\prod_{j \leq n} a_j$. The variable j in the notation $\prod_{j=k}^{n} a_j$ is called a *dummy variable*, because we can change it without causing any harm. For example, we can denote $\prod_{j=k}^{n} a_j$ by $\prod_{i=k}^{n} a_i$, or by $\prod_{l=k}^{n} a_l$.

Remark. When the binary operation is denoted by $a \cdot b$, or simply by ab, then we keep using the notation $\prod$ for the product of several elements. But when the binary operation is denoted by $a + b$, we use the notation $\sum$ instead of $\prod$, and we use the term "sum" instead of "product".

Let $a_k, \ldots, a_m$ be a finite sequence in S. This means that there is a function $f : \{k, \ldots, m\} \to S$ such that $f(j) = a_j$ for every j. Suppose $\tau : \{l, \ldots, n\} \to \{k, \ldots, m\}$ is a bijective map. Then $f \circ \tau : \{l, \ldots, n\} \to S$ is also a sequence, and for every i we have $(f \circ \tau)(i) = f(\tau(i)) = a_{\tau(i)}$. The new sequence $a_{\tau(l)}, \ldots, a_{\tau(n)}$ has the same elements as $a_k, \ldots, a_m$; however, the order of elements in the two sequences might be different. The map τ is called a *change of index*.

Let us for now concentrate on changes of index which do not alter the order of elements in the sequence. Let τ be such a change of index. Then it is obvious that τ must be strictly increasing. Because we want $a_{\tau(i)}$ to appear after $a_{\tau(i')}$ in the sequence, when $i > i'$. In other words, we want $\tau(i) > \tau(i')$ when $i > i'$. It is easy to show by induction that τ must satisfy $\tau(i) = i + k - l$, for every i in its domain. In other words, every strictly increasing change of index is just a shift.

Theorem 5.38. *Let $a_k, \dots, a_m$ be a finite sequence in S, and let $\star$ be a binary operation on S. Also let $\tau : \{l, \dots, n\} \to \{k, \dots, m\}$ be defined by $\tau(i) = i + k - l$, and suppose $\tau(n) = m$. Then τ is a strictly increasing bijective map, and we have*

$$\prod_{i=l}^{n} a_{\tau(i)} = \prod_{j=k}^{m} a_j.$$

Remark. Note that τ is an invertible map, and we have $\tau(l) = k$; so we can write the above equality as

$$\prod_{i=\tau^{-1}(k)}^{\tau^{-1}(m)} a_{\tau(i)} = \prod_{j=k}^{m} a_j.$$

Remark. The most common applications of this theorem are when τ is of the form $i \mapsto i + 1$, in which case we have $k = l + 1$ and $m = n + 1$; or when τ is of the form $i \mapsto i - 1$, in which case we have $k = l - 1$ and $m = n - 1$. In these cases we have

$$\prod_{i=k-1}^{m-1} a_{i+1} = \prod_{j=k}^{m} a_j = \prod_{i=k+1}^{m+1} a_{i-1}.$$

Proof. Note that if $i < i'$ then

$$\tau(i) = i + k - l < i' + k - l = \tau(i').$$

So τ is strictly increasing; hence it is one-to-one. Also, for every $k \le j \le m$ we have $l \le j + l - k \le n$, and $\tau(j + l - k) = j$. Hence τ is onto. Now we prove the desired equality by induction on m. For $m = k$ we have $n = l$, and $\prod_{i=l}^{l} a_{\tau(i)} = a_{\tau(l)} = a_k = \prod_{j=k}^{k} a_j$. Suppose the equality holds for some m. Then for $m + 1$ we have $\tau^{-1}(m + 1) = n + 1$, and

$$\prod_{i=l}^{n+1} a_{\tau(i)} = \Big(\prod_{i=l}^{n} a_{\tau(i)}\Big) \star a_{\tau(n+1)} = \Big(\prod_{j=k}^{m} a_j\Big) \star a_{m+1} = \prod_{j=k}^{m+1} a_j,$$

as desired. ∎

Definition 5.33. Suppose $\star$ is a binary operation on a set S, and $a_k, \dots, a_m$ is a finite sequence in S. We inductively define the set of *all possible products* of $a_k, \dots, a_m$ as follows

(i) $P(a_k) := \{a_k\}$,

(ii) and for $k \le n < m$ we set

$$P(a_k, \dots, a_{n+1}) := \{b \star c : b \in P(a_k, \dots, a_j)\ ,\ c \in P(a_{j+1}, \dots, a_{n+1}) \text{ for some } k \le j \le n\}.$$

Remark. Note that we do not change the order of $a_k, \ldots, a_m$ when we form a possible product of them; we merely rearrange the parentheses.

Remark. Unlike our previous applications of recursion theorem, here, it is not obvious which functions are used for the definition of $P(a_k, \ldots, a_m)$. Let us clearly state how we use the recursion theorem in the above definition. Let Y be the set of all finite sequences in S whose domains are $\{k, \ldots, l\}$ for some $l \leq m$. Let X be the set of all functions from Y to $\mathcal{P}(S)$. We want to define a function $f : \{k, \ldots, m\} \to X$. Let $f(k)$ be the element of X that maps every sequence with length one of the form b_k to the singleton $\{b_k\} \in \mathcal{P}(S)$, and every other sequence to $\emptyset$. Let $F : \{k, \ldots, m\} \times X \to X$ be defined as follows. For every $k \leq n \leq m$ and $g \in X$, $F(n, g)$ is the element of X that maps every sequence $b_k, \ldots, b_l$ in Y to $g(b_k, \ldots, b_l) \in \mathcal{P}(S)$ when $l \neq n + 1$, and when $l = n + 1$ we have

$$F(n, g)(b_k, \ldots, b_{n+1}) = \bigcup_{j=k}^{n} g(b_k, \ldots, b_j) \star g(b_{j+1}, \ldots, b_{n+1}).$$

Here we used the notation $B \star C$ for $B, C \subset S$ to denote $\{b \star c : b \in B, c \in C\}$. Now, recursion theorem tells us that there is a function $f : \{k, \ldots, m\} \to X$ such that $f(k)$ is defined as above, and for every $k \leq n < m$ we have $f(n + 1) = F(n, f(n))$. Finally, for every $k \leq n \leq m$ we define

$$P(a_k, \ldots, a_n) := f(n)(a_k, \ldots, a_n).$$

Intuitively, it should be clear that this definition is the same as the previous one.

Generalized Associativity. *Suppose $\star$ is an associative binary operation, and $a_k, \ldots, a_n \in S$. Then $P(a_k, \ldots, a_n)$ has exactly one element, that is $\prod_{i=k}^{n} a_i$. In addition, for every $k \leq l < n$ we have*

$$\Big(\prod_{i=k}^{l} a_i\Big) \star \Big(\prod_{i=l+1}^{n} a_i\Big) = \prod_{i=k}^{n} a_i.$$

Proof. It is easy to show by induction that $\prod_{i=k}^{n} a_i \in P(a_k, \ldots, a_n)$. For uniqueness we can argue inductively as follows. When $n = k$, $P(a_k)$ has one element by definition. Suppose the theorem is true for $k \leq l < n$, i.e. $P(b_k, \ldots, b_l)$ has exactly one element for every $b_k, \ldots, b_l$. Then the elements of $P(a_k, \ldots, a_n)$ are of the form $b \star c$ where $b \in P(a_k, \ldots, a_l)$ and $c \in P(a_{l+1}, \ldots, a_n)$. Thus by induction hypothesis we have

$$b = \prod_{i=k}^{l} a_i, \qquad c = \prod_{i=l+1}^{n} a_i.$$

When $l = n-1$ we have $c = a_n$, hence $b \star c = \left(\prod_{i=k}^{n-1} a_i\right) \star a_n = \prod_{i=k}^{n} a_i$. Otherwise we have

$$\begin{aligned}\Big(\prod_{i=k}^{l} a_i\Big) \star \Big(\prod_{i=l+1}^{n} a_i\Big) &= \Big(\prod_{i=k}^{l} a_i\Big) \star \Big[\Big(\prod_{i=l+1}^{n-1} a_i\Big) \star a_n\Big] \\ &= \Big[\Big(\prod_{i=k}^{l} a_i\Big) \star \Big(\prod_{i=l+1}^{n-1} a_i\Big)\Big] \star a_n \\ &= \Big(\prod_{i=k}^{n-1} a_i\Big) \star a_n = \prod_{i=k}^{n} a_i.\end{aligned}$$

■

Remark. The above theorem means that when $\star$ is associative, the value of the product of $a_k, \ldots, a_n$ is independent of the arrangement of the parentheses. In this case we sometimes denote $\prod_{i=k}^{n} a_i$ by $a_k \star \cdots \star a_n$.

***Exercise* 5.3.** Suppose $\star$ is an associative binary operation on S, and $a_k, \ldots, a_n \in S$ are invertible. Show that $\prod_{i=k}^{n} a_i$ is also invertible and we have

$$(a_k \star \cdots \star a_n)^{-1} = a_n^{-1} \star \cdots \star a_k^{-1}.$$

We have seen that if τ is a change of index which does not alter the order of elements in the sequence, then we have

$$a_{\tau(l)} \star \cdots \star a_{\tau(n)} = a_k \star \cdots \star a_m.$$

The next theorem shows that if $\star$ is associative and commutative, then the above property holds for every change of index, i.e. the order of $a_k, \ldots, a_m$ does not affect the value of $a_k \star \cdots \star a_m$. Remember that a permutation is a bijective map from a set to itself.

Generalized Commutativity. *Suppose $\star$ is an associative and commutative binary operation, and $a_k, \ldots, a_n \in S$. Then for every permutation $\sigma : \{k, \ldots, n\} \to \{k, \ldots, n\}$ we have*

$$a_{\sigma(k)} \star \cdots \star a_{\sigma(n)} = a_k \star \cdots \star a_n.$$

Proof. We use induction on n. The case $n = k$ is obvious, so suppose the conclusion holds for all permutations on $\{k, \ldots, n-1\}$. Now for the induction step we have

$$\prod_{i=k}^{n} a_{\sigma(i)} = \prod_{i=k}^{n-1} a_{\sigma(i)} \star a_{\sigma(n)}.$$

Suppose $\sigma(j) = n$. Then

$$\begin{aligned}\prod_{i=k}^{n-1} a_{\sigma(i)} &= \prod_{i=k}^{j-1} a_{\sigma(i)} \star a_n \star \prod_{i=j+1}^{n-1} a_{\sigma(i)} \\ &= \prod_{i=k}^{j-1} a_{\sigma(i)} \star \prod_{i=j+1}^{n-1} a_{\sigma(i)} \star a_n = \prod_{i=k, i\neq j}^{n-1} a_{\sigma(i)} \star a_n.\end{aligned}$$

Let $\hat{\sigma}$ be the permutation on $\{k, \ldots, n-1\}$ defined by

$$\hat{\sigma}(i) = \begin{cases} \sigma(i) & i < j \\ \sigma(i+1) & i \geq j. \end{cases}$$

Then we have

$$\begin{aligned}\prod_{i=k}^{n} a_{\sigma(i)} &= \prod_{i=k, i\neq j}^{n-1} a_{\sigma(i)} \star a_n \star a_{\sigma(n)} \\ &= \prod_{i=k, i\neq j}^{n-1} a_{\sigma(i)} \star a_{\sigma(n)} \star a_n \\ &= \prod_{i=k}^{j-1} a_{\sigma(i)} \star \prod_{i=j+1}^{n} a_{\sigma(i)} \star a_n \\ &= \prod_{i=k}^{n-1} a_{\hat{\sigma}(i)} \star a_n = \prod_{i=k}^{n-1} a_i \star a_n = \prod_{i=k}^{n} a_i. \end{aligned}$$

■

Suppose $\star$ is an associative and commutative binary operation. Sometimes we want to compute the product of several elements of S that do not have an order, or are not ordered linearly. Suppose I is a finite set, and $\alpha : I \to S$ is a function. We want to compute the product of all the elements $\alpha(r)$ for $r \in I$. Note that α need not be one-to-one, so some of the $\alpha(r)$'s might be the same. In other words, we may have repetition of factors in our product, as we may have when we multiply the elements of a sequence. Suppose I has n elements, and $f : \{1, \ldots, n\} \to I$ is a one-to-one and onto function. Let us denote $\alpha(f(k))$ by a_k. Now we define

$$\prod_{r \in I} \alpha(r) := \prod_{k \leq n} a_k.$$

We have to check that this definition is independent of f. Let $g : \{1, \ldots, n\} \to I$ be another one-to-one and onto function. Then

$$\sigma := f^{-1} \circ g : \{1, \ldots, n\} \to \{1, \ldots, n\}$$

is also one-to-one and onto, i.e. it is a permutation. Let us denote $\alpha(g(k))$ by $\tilde{a}_k$. Hence by the above theorem we have

$$\prod_{k\le n}\tilde{a}_k = \prod_{k\le n}\alpha(g(k)) = \prod_{k\le n}\alpha(f(\sigma(k))) = \prod_{k\le n}a_{\sigma(k)} = \prod_{k\le n}a_k = \prod_{r\in I}\alpha(r),$$

as desired. The element $\prod_{r\in I}\alpha(r)$ is sometimes called the *unordered product* of the elements $\alpha(r)$ for $r \in I$. A particular case is when $A \subset S$ is a finite set, and $\alpha : A \to S$ is the inclusion map. Then the unordered product of the elements of A is by definition

$$\prod_{a\in A} a := \prod_{a\in A}\alpha(a).$$

Theorem 5.39. *Suppose $\star$ is an associative and commutative binary operation, and $a_i, b_i, a_{ij} \in S$ for $i \le n$, $j \le m$. Then*

$$\prod_{i=1}^{n} a_i \star \prod_{i=1}^{n} b_i = \prod_{i=1}^{n}(a_i \star b_i),$$

and

$$\prod_{j=1}^{m}\prod_{i=1}^{n} a_{ij} = \prod_{i=1}^{n}\prod_{j=1}^{m} a_{ij} = \prod_{k=2}^{m+n}\prod_{i+j=k} a_{ij}.$$

Here $\prod_{i+j=k} a_{ij}$ is a shorthand notation for $\prod_{i=r}^{l} a_{i,k-i}$, where $r = \max\{1, k-m\}$, and $l = \min\{n, k-1\}$.

Remark. Note that we can also rewrite the second equation as follows

$$\prod_{j=1}^{m}(a_{1j} \star \cdots \star a_{nj}) = \Big(\prod_{j=1}^{m} a_{1j}\Big) \star \cdots \star \Big(\prod_{j=1}^{m} a_{nj}\Big).$$

Remark. When we use the additive notation $+$ for $\star$, the above equations take the following more familiar forms

$$\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i = \sum_{i=1}^{n}(a_i + b_i),$$

$$\sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} = \sum_{k=2}^{m+n}\sum_{i+j=k} a_{ij},$$

$$\sum_{j=1}^{m}(a_{1j} + \cdots + a_{nj}) = \Big(\sum_{j=1}^{m} a_{1j}\Big) + \cdots + \Big(\sum_{j=1}^{m} a_{nj}\Big).$$

Proof. The proofs are by induction on n. We only write the induction step. For the first equality we have

$$\begin{aligned}\prod_{i=1}^{n+1} a_i \star \prod_{i=1}^{n+1} b_i &= \Big(\prod_{i=1}^{n} a_i\Big) \star a_{n+1} \star \Big(\prod_{i=1}^{n} b_i\Big) \star b_{n+1} \\ &= \Big(\prod_{i=1}^{n} a_i\Big) \star \Big(\prod_{i=1}^{n} b_i\Big) \star a_{n+1} \star b_{n+1} \\ &= \Big(\prod_{i=1}^{n} (a_i \star b_i)\Big) \star (a_{n+1} \star b_{n+1}) = \prod_{i=1}^{n+1} (a_i \star b_i).\end{aligned}$$

For the second equality we have

$$\begin{aligned}\prod_{i=1}^{n+1}\prod_{j=1}^{m} a_{ij} &= \Big(\prod_{i=1}^{n}\prod_{j=1}^{m} a_{ij}\Big) \star \prod_{j=1}^{m} a_{n+1,j} \\ &= \Big(\prod_{j=1}^{m}\prod_{i=1}^{n} a_{ij}\Big) \star \prod_{j=1}^{m} a_{n+1,j} \\ &= \prod_{j=1}^{m}\Big[\Big(\prod_{i=1}^{n} a_{ij}\Big) \star a_{n+1,j}\Big] = \prod_{j=1}^{m}\prod_{i=1}^{n+1} a_{ij}.\end{aligned}$$

For the third equality let $r := \max\{1, k-m\}$, $l := \min\{n, k-1\}$, and $L := \min\{n+1, k-1\}$. Note that for $k \le n+1$ we have $l = L$, and for $k \ge n+2$ we have $l = n$ and $L = n+1$. Now we have

$$\begin{aligned}\prod_{i=1}^{n+1}\prod_{j=1}^{m} a_{ij} &= \Big(\prod_{i=1}^{n}\prod_{j=1}^{m} a_{ij}\Big) \star \prod_{j=1}^{m} a_{n+1,j} = \Big(\prod_{k=2}^{m+n}\prod_{i=r}^{l} a_{i,k-i}\Big) \star \prod_{j=1}^{m} a_{n+1,j} \\ &= \Big(\prod_{k=2}^{m+n}\prod_{i=r}^{l} a_{i,k-i}\Big) \star \Big(\prod_{k=n+2}^{m+n} a_{n+1,k-n-1}\Big) \star a_{n+1,m} \\ &= \Big(\prod_{k=2}^{n+1}\prod_{i=r}^{l} a_{i,k-i}\Big) \star \Big(\prod_{k=n+2}^{m+n}\prod_{i=r}^{n} a_{i,k-i}\Big) \star \Big(\prod_{k=n+2}^{m+n} a_{n+1,k-n-1}\Big) \star a_{n+1,m} \\ &= \Big(\prod_{k=2}^{n+1}\prod_{i=r}^{l} a_{i,k-i}\Big) \star \Big(\prod_{k=n+2}^{m+n}\Big[\Big(\prod_{i=r}^{n} a_{i,k-i}\Big) \star a_{n+1,k-n-1}\Big]\Big) \star a_{n+1,m}\end{aligned}$$

$$= \Big(\prod_{k=2}^{n+1}\prod_{i=r}^{L} a_{i,k-i}\Big) \star \Big(\prod_{k=n+2}^{m+n}\prod_{i=r}^{n+1} a_{i,k-i}\Big) \star a_{n+1,m}$$
$$= \Big(\prod_{k=2}^{m+n}\prod_{i=r}^{L} a_{i,k-i}\Big) \star a_{n+1,m} = \prod_{k=2}^{m+n+1}\prod_{i=r}^{L} a_{i,k-i} = \prod_{k=2}^{m+n+1}\prod_{i+j=k} a_{ij}.$$

■

Notation. The common element of the second and third equalities in the above theorem is often denoted by

$$\prod_{i,j} a_{ij}.$$

Remark. An interesting consequence of the above theorem is that it enables us to compute the following *telescoping product*. Let $\star$ be an associative and commutative binary operation on S. Then for invertible elements $a_1, \ldots, a_n \in S$ we have

$$\begin{aligned}\prod_{i=2}^{n}(a_i \star a_{i-1}^{-1}) &= \prod_{i=2}^{n} a_i \star \prod_{i=2}^{n} a_{i-1}^{-1} = \prod_{i=2}^{n-1} a_i \star a_n \star \Big(\prod_{i=n}^{2} a_{i-1}\Big)^{-1} \\ &= a_n \star \prod_{i=2}^{n-1} a_i \star \Big(\prod_{i=2}^{n} a_{i-1}\Big)^{-1} = a_n \star \prod_{i=2}^{n-1} a_i \star \Big(\prod_{i=1}^{n-1} a_i\Big)^{-1} \\ &\qquad\qquad \text{(We changed } i-1 \text{ to } i \text{ in the last term.)} \\ &= a_n \star \prod_{i=2}^{n-1} a_i \star \Big(a_1 \star \prod_{i=2}^{n-1} a_i\Big)^{-1} = a_n \star \prod_{i=2}^{n-1} a_i \star \Big(\prod_{i=2}^{n-1} a_i\Big)^{-1} \star a_1^{-1} \\ &= a_n \star a_1^{-1}.\end{aligned}$$

Note that here we have used both the generalized associativity and the generalized commutativity.

Definition 5.34. Suppose $\star$ is a binary operation on S, and $a \in S$. Let n be a positive integer. We inductively define the **powers** of a as follows

(i) $a^1 := a$,

(ii) $a^{n+1} := a^n \star a$.

If there exists an identity e, we define $a^0 := e$. When m is a negative integer, $n := -m$ is positive. In this case, if a has an inverse a^{-1}, we define

$$a^m = a^{-n} := (a^{-1})^n.$$

Remark. In the terminology of recursion theorem, we have constructed the function $f(n) = a^n$ for $n \geq 0$, by using the function $F(n, s) = s \star a$. Although, to be precise, we have to actually define f on ω; and then for a nonnegative integer n we

have to set $a^n = f(E^{-1}(n))$, where E is the one-to-one correspondence between ω and the set of nonnegative integers, constructed in Theorem 5.11.

Remark. When the binary operation is denoted by $a \cdot b$, or simply by ab, then we keep using the notation a^n for powers of a. But when the binary operation is denoted by $a + b$, we use the notation na instead of a^n, and we sometimes use the term "additive power" instead of "power".

Theorem 5.40. *Suppose $\star$ is a binary operation on S. Let $a \in S$, and let n be a positive integer. Then we have*

$$a^n = \prod_{j=1}^{n} a.$$

Proof. Note that the sequence whose product appears in the theorem is the constant sequence whose terms are all a. The proof is by induction on n. When $n = 1$, by definition we have

$$a^1 = a = \prod_{j=1}^{1} a.$$

Suppose the claim holds for some n. Then for $n + 1$ we have

$$a^{n+1} = a^n \star a = \Big(\prod_{j=1}^{n} a\Big) \star a = \prod_{j=1}^{n+1} a,$$

as desired. ■

Theorem 5.41. *Suppose $\star$ is an associative binary operation on S. Let $n, m \in \mathbb{Z}$. Then for every $a, b \in S$ we have*

(i) *If a commutes with b, then a^n commutes with b^m, for all $m, n \geq 0$. If one or both of a, b are invertible, we can allow n and/or m to be negative too.*

(ii) *If a is invertible, then a^n is also invertible for all $n \in \mathbb{Z}$, and*

$$(a^n)^{-1} = a^{-n} = (a^{-1})^n.$$

(iii) *$a^n \star a^m = a^{n+m}$ for all $m, n \geq 0$. If a is invertible, we can allow m, n to be negative too.*

(iv) *$(a^n)^m = a^{nm}$ for all $m, n \geq 0$. If a is invertible, we can allow m, n to be negative too.*

(v) *If a, b commute, we have $a^n \star b^n = (a \star b)^n$ for all $n \geq 0$. If a, b are invertible, we can allow n to be negative too.*

Proof. The proof is the same as of Theorem 5.15. ■

Definition 5.35. Suppose $\star$ and $+$ are binary operations on S. We say $\star$ is **distributive** over $+$ if for all $a, b, c \in S$ we have

$$a \star (b + c) = (a \star b) + (a \star c), \qquad (b + c) \star a = (b \star a) + (c \star a).$$

Generalized Distributivity. *Suppose $\star, +$ are associative binary operations on S, and $\star$ is distributive over $+$. Then for all $a_{ij} \in S$ and $n_j \in \mathbb{N}$ we have*

$$\Big(\sum_{i_1=1}^{n_1} a_{i_1 1}\Big) \star \cdots \star \Big(\sum_{i_k=1}^{n_k} a_{i_k k}\Big) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_k=1}^{n_k} (a_{i_1 1} \star \cdots \star a_{i_k k}).$$

Remark. A particular case of this theorem is when $k = 2$, and n_1 or n_2 is 1. Then we have

$$a \star \Big(\sum_{i=1}^{n} a_i\Big) = \sum_{i=1}^{n} (a \star a_i) \qquad \Big(\sum_{i=1}^{n} a_i\Big) \star a = \sum_{i=1}^{n} (a_i \star a).$$

Proof. First suppose $k = 2$, and n_1 or n_2 is 1. We have to show that the above equations hold. This can be easily done by induction on n. Now for the general case, we proceed by induction on k. Suppose the conclusion holds for $k - 1$. Then we have

$$\begin{aligned}
&\Big(\sum_{i_1=1}^{n_1} a_{i_1 1}\Big) \star \Big(\sum_{i_2=1}^{n_2} a_{i_2 2}\Big) \star \cdots \star \Big(\sum_{i_k=1}^{n_k} a_{i_k k}\Big) \\
&\qquad = \sum_{i_1=1}^{n_1} \left[a_{i_1 1} \star \Big(\sum_{i_2=1}^{n_2} a_{i_2 2}\Big) \star \cdots \star \Big(\sum_{i_k=1}^{n_k} a_{i_k k}\Big) \right] \\
&\qquad = \sum_{i_1=1}^{n_1} \left[\Big(\sum_{i_2=1}^{n_2} a_{i_1 1} \star a_{i_2 2}\Big) \star \cdots \star \Big(\sum_{i_k=1}^{n_k} a_{i_k k}\Big) \right] \\
&\qquad = \sum_{i_1=1}^{n_1} \left[\sum_{i_2=1}^{n_2} \cdots \sum_{i_k=1}^{n_k} \big((a_{i_1 1} \star a_{i_2 2}) \cdots \star a_{i_k k}\big) \right] \\
&\qquad = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_k=1}^{n_k} (a_{i_1 1} \star a_{i_2 2} \cdots \star a_{i_k k}).
\end{aligned}$$

■

Remark. Note that we do not need the commutativity of $+$ in the above theorem. To make this clear let us look at the following special case:

$$\begin{aligned}
(a_{11} + a_{21})(a_{12} + a_{22}) &= a_{11}(a_{12} + a_{22}) + a_{21}(a_{12} + a_{22}) \\
&= a_{11}a_{12} + a_{11}a_{22} + a_{21}a_{12} + a_{21}a_{22} \\
&= \sum_{i_2=1}^{2} a_{11}a_{i_2 2} + \sum_{i_2=1}^{2} a_{21}a_{i_2 2} = \sum_{i_1=1}^{2} \sum_{i_2=1}^{2} a_{i_1 1}a_{i_2 2}.
\end{aligned}$$

***Example* 5.7.** As an application of the above theorems, let us show that

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

By Exercise 4.1, and the results of this section we have

$$\begin{aligned}\sum_{k=1}^{n}(2k-1) = \sum_{k=1}^{n}2k + \sum_{k=1}^{n}(-1) = 2\sum_{k=1}^{n}k - \sum_{k=n}^{1}1 = 2\sum_{k=1}^{n}k - \sum_{k=1}^{n}1 \\ = 2\frac{1}{2}n(n+1) - n = n^2 + n - n = n^2.\end{aligned}$$

Chapter 6

Real Numbers and Complex Numbers

6.1 Real Numbers

In the field of rational numbers, we can perform the four basic arithmetic operations. Therefore rational numbers are adequate for many purposes. However, the field $\mathbb{Q}$ has some deficiencies. For example, the length of the diagonal of a square whose side length is one cannot be expressed by a rational number, i.e. by the ratio of two integers. In other words, there is no rational number p that satisfies $p^2 = 2$ (for the proof, see Theorem 6.19). Hence, in order to fill the gaps of the set of rational numbers, we need to extend it to a larger set, which is known as the set of real numbers. As we explained before, we just need to find a set whose elements represent the real numbers, and have the properties that we expect from the real numbers; and we are not concerned with the inherent nature of real numbers.

The basic idea, due to Dedekind, is to consider a real number as the set of all rational numbers smaller than it. Because, intuitively we know that if two real numbers are different, then there is a rational number between them; so the set of rational numbers less than the larger real number is distinct from the set of rational numbers less than the smaller real number. Thus, a real number is uniquely determined by the set of rational numbers smaller than it.

Definition 6.1. A **Dedekind cut** is a pair (A, B), where $A, B \subset \mathbb{Q}$ satisfy the following conditions:

(i) $A \neq \emptyset$, $B \neq \emptyset$, $A \cap B = \emptyset$, and $A \cup B = \mathbb{Q}$.
(ii) If $a \in A$ and $b \in B$, then $a < b$.
(iii) A does not contain a largest element.

Remark. We denote a Dedekind cut (A, B) by $A|B$. Note that in a Dedekind cut $A|B$ we have $B = \mathbb{Q} - A$; so B is uniquely determined from A, and vice versa.

Remark. Also note that if $a \in A$ and $c \leq a$, then we must have $c \in A$. Since if $c \in B$ then we would have $a < c$, contrary to our assumption. Similarly, if $b \in B$ and $d \geq b$, then we must have $d \in B$.

Remark. Note that B may or may not have a smallest element.

Definition 6.2. A **real number** is a Dedekind cut. The set of all real numbers is

$$\mathbb{R} := \{x \in \mathcal{P}(\mathbb{Q}) \times \mathcal{P}(\mathbb{Q}) : x \text{ is a Dedekind cut}\}.$$

Remark. Intuitively, we think of $\mathbb{R}$ as a line. Thus, a real number $x = A|B$ represents the cut in the line $\mathbb{R}$ at the point x, and A, B are the set of rational numbers in the two pieces of the line that remain after the cut.

Proposition 6.1. *Let $p \in \mathbb{Q}$. Also let*

$$A = \{r \in \mathbb{Q} : r < p\}, \qquad B = \{r \in \mathbb{Q} : r \geq p\}.$$

Then $A|B$ is a real number, i.e. a Dedekind cut.

Proof. First note that A, B are nonempty; because $\mathbb{Q}$ does not have a largest or smallest element. Also $A \cap B = \emptyset$, since we cannot have $r < p$ and $r \geq p$ for any r. In addition, note that $A \cup B = \mathbb{Q}$, since by trichotomy law we have $r < p$ or $r \geq p$ for every r. Furthermore, for $a \in A$ and $b \in B$ we have $a < p \leq b$; so $a < b$. Finally, note that A does not have a largest element. Because for every $a \in A$ we have $a < p$. Hence there is a rational number c such that $a < c < p$. Thus $c \in A$, and $c > a$. Therefore a cannot be the largest element of A. So $A|B$ is a Dedekind cut. ∎

Now let us define the order of real numbers. If $x = A|B$ and $y = C|D$ are two real numbers, then intuitively we know that A consists of rational numbers less than x, and similarly C consists of rational numbers less than y. Hence if $x < y$ we must have $A \subset C$. In addition, we intuitively know that there are rational numbers between x, y; hence we also must have $A \neq C$. Thus we arrive at the following definition of order for real numbers.

Definition 6.3. Let $x = A|B$ and $y = C|D$ be two real numbers. Then we say $x < y$ if $A \subset C$ and $A \neq C$.

Remark. We know that by definition, $x \leq y$ means that $x < y$ or $x = y$. But $x = y$ is equivalent to $A = C$, since B, D are uniquely determined from A, C respectively. Therefore we have $x \leq y$ if and only if $A \subset C$.

Theorem 6.1. *The relation $<$ on $\mathbb{R}$ is a linear order. In addition, $\mathbb{R}$ does not have a smallest element, nor a largest element.*

Proof. Let $x = A|B$, $y = C|D$, and $z = E|F$. First note that $x \not< x$, since $A = A$. So $<$ is irreflexive. Now suppose $x < y$ and $y < z$. Then we have

$$A \subset C,\ A \neq C, \qquad C \subset E,\ C \neq E.$$

Hence we get $A \subset E$. We also have $A \neq E$, since otherwise we would get $C \subset A$, which contradicts the fact that $A \neq C$ and $A \subset C$. Thus $<$ is also transitive.

Next suppose $x \neq y$. Then $A \neq C$. We need to show that either $A \subset C$, or $C \subset A$; in order to conclude that either $x < y$, or $x > y$. Suppose $A \not\subset C$. Then there is $a \in A$ such that $a \notin C$. Hence we have $a \in D$. Thus for every $c \in C$ we must have $c < a$. Now note that if $c \in B$ then we cannot have $a \in A$, since $c < a$. Therefore $c \notin B$. Hence $c \in A$. Thus we have $C \subset A$, as desired. So $<$ is a linear order on $\mathbb{R}$.

Finally, let us show that for every $x = A|B$ there are $y, z \in \mathbb{R}$ such that $y < x < z$. Let $p \in A$ and $q \in B$. Let $y := C|D$ and $z := E|F$, where

$$C = \{r \in \mathbb{Q} : r < p\}, \qquad D = \{r \in \mathbb{Q} : r \geq p\},$$
$$E = \{r \in \mathbb{Q} : r < q + 1\}, \qquad F = \{r \in \mathbb{Q} : r \geq q + 1\}.$$

Then we have $A \subset E$, since for every $a \in A$ we have $a < q < q + 1$. Also, $A \neq E$; because $q \in E - A$. In addition, we have $C \subset A$, since if $r \in B$ then $r > p$; so for every $r < p$ we must have $r \in A$. Furthermore, $C \neq A$; because A must contain an element larger than p, since A does not have a largest element. So we have shown that $y < x < z$. Therefore no real number like x can be the smallest element, nor the largest element of $\mathbb{R}$. ■

Theorem 6.2. *Suppose $S \subset \mathbb{R}$ is nonempty and bounded above. Then S has a least upper bound in $\mathbb{R}$.*

Proof. Let

$$A := \{r \in \mathbb{Q} : r \in C \text{ for some } C|D \in S\}, \qquad B := \mathbb{Q} - A.$$

First note that $A|B$ is a cut. It is obvious that $A \cap B = \emptyset$, and $A \cup B = \mathbb{Q}$. It is also obvious that $A \neq \emptyset$, since S is nonempty. We also have $B \neq \emptyset$. To see this let $E|F$ be an upper bound for S. Then for every $C|D \in S$ we have $C \subset E$; so $A \subset E$. Thus $F \subset B$. Hence $B \neq \emptyset$. Now let $a \in A$ and $b \in B$. Then $a \in C$ for some $C|D \in S$. If $b \in C$ then $b \in A$, which contradicts our assumption. So $b \in D$. Thus we must have $a < b$. Finally, let $a \in A$. Then $a \in C$ for some $C|D \in S$. But we know that there is $c \in C$ such that $a < c$, since C does not have a largest element. However, we also have $c \in A$. Thus no element $a \in A$ can be the largest element of A. Therefore we have shown that $A|B$ is a cut.

Next let us show that $A|B$ is an upper bound for S. Let $C|D \in S$, and let $r \in C$. Then by definition we have $r \in A$. Hence $C \subset A$. Thus $C|D \leq A|B$, as desired. Now let $E|F$ be an upper bound for S. Let $r \in A$. Then there is $C|D \in S$ such that $r \in C$. However, we know that $C \subset E$. Thus we have $r \in E$. Hence $A \subset E$. Therefore $A|B \leq E|F$. So $A|B$ is the least upper bound of S. ■

Next, we have to define the addition and multiplication on $\mathbb{R}$, and show that they have the expected properties. Let $x = A|B, y = C|D \in \mathbb{R}$. Informally, these cuts represent the real numbers which are the supremum of the set of rational numbers in A, C respectively. Thus, at first glance, if we want to add and multiply x, y, the results should be cuts whose first components are respectively

$$\{a + c : a \in A, c \in C\},$$
$$\{ac : a \in A, c \in C\}.$$

Because if $a < x$ and $c < y$, then we have $a + c < x + y$. We also have $ac < xy$, provided that $a, c > 0$. However, if a, c are negative, then ac can be a large positive number greater than xy. Thus the definition of multiplication of real numbers needs more attention, which will be discussed later.

Furthermore, given a real number like $x = A|B$, we need to find its opposite and its inverse (provided that x is nonzero). Intuitively, we know that if the real number $B'|A'$ is the opposite of x, then B' consists of rational numbers s less than $-x$. However $s < -x$ if and only if $-s > x$. Therefore we must have $-s \in B$. In other words, the cut whose first component, i.e. B', consists of the opposite of the elements of B is our candidate for $-x$. But we must note that B can have a least element, and B' cannot have a largest element. Therefore when we construct B' from B we have to exclude the least element of B, if it exists. Hence we arrive at the following definition. The idea for finding the inverse of a nonzero real number is similar, and will be discussed later. First let us check that these proposed definitions for the sum and opposite are actually Dedekind cuts.

Proposition 6.2. *Let $x, y \in \mathbb{R}$, and suppose $x = A|B$, $y = C|D$. Let*

$$E = \{a + c : a \in A, c \in C\}, \qquad F = \mathbb{Q} - E.$$

Also let

$$B' = \{-b : b \in B,\ b \text{ is not the smallest element of } B\}, \qquad A' = \mathbb{Q} - B'.$$

Then $E|F$ and $B'|A'$ are real numbers.

Proof. First note that E is nonempty, since A, C are nonempty. It is also obvious that $E \cap F = \emptyset$, and $E \cup F = \mathbb{Q}$. Now let $b \in B$ and $d \in D$. Then for every $a \in A$

and $c \in C$ we have $a < b$ and $c < d$. Thus $a+c < b+d$. So $b+d \notin E$. Hence $F \neq \emptyset$. Next let $a + c \in E$ and $r \in F$. If $r \leq a + c$ then $r - a \leq c$. Thus $r - a \in C$. Hence we have $r = a + (r - a) \in E$, which is a contradiction. Therefore we must have $a + c < r$. Finally, let $a + c \in E$. Then there are $p \in A$ and $q \in C$ such that $a < p$ and $c < q$; because A, C do not have largest elements. Now we have $p+q \in E$, and $a + c < p + q$. Thus no element $a + c$ can be the largest element of E.

Next lest us consider B', A'. Note that B' is nonempty. Because B is nonempty, and for every $b \in B$ we also have $b + 1 \in B$; so B has elements which are not its smallest element. It is also obvious that $B' \cap A' = \emptyset$, and $B' \cup A' = \mathbb{Q}$. In addition, for $a \in A$ we must have $-a \in A'$. Thus $A' \neq \emptyset$. Now let $b' \in B'$ and $a' \in A'$. Then $-b' \in B$. Also, $-a'$ either belongs to A, or is the smallest element of B (otherwise a' belongs to B'). In either case we have $-a' < -b'$. Hence $b' < a'$.

Finally, note that if $b' \in B'$ is its largest element, then $-b'$ must be less than or equal to every element of B except its smallest element (if B has a smallest element). Now if B does not have a smallest element, then $-b'$ is less than or equal to every element of B. Hence $-b'$ is the smallest element of B, contrary to our assumption. So suppose $p \in B$ is its smallest element. Then we have $p < -b'$. Let q be a rational number such that $p < q < -b'$. Then we have $q \in B$, since $p < q$. However, q is not the smallest element of B. So it cannot be smaller than $-b'$. Thus we have a contradiction, and therefore B' cannot have a largest element. ■

Definition 6.4. Let $x, y \in \mathbb{R}$, and suppose $x = A|B$, $y = C|D$. Then we define their **addition** to be $x + y := E|F$, where

$$E = \{a + c : a \in A, c \in C\}, \qquad F = \mathbb{Q} - E.$$

The **zero** and **identity** of $\mathbb{R}$ are

$$0 := \{r \in \mathbb{Q} : r < 0\} \,|\, \{r \in \mathbb{Q} : r \geq 0\},$$
$$1 := \{r \in \mathbb{Q} : r < 1\} \,|\, \{r \in \mathbb{Q} : r \geq 1\},$$

respectively. The **opposite** of x is $-x := B'|A'$, where

$$B' = \{-b : b \in B,\ b \text{ is not the smallest element of } B\}, \qquad A' = \mathbb{Q} - B'.$$

Remark. It is easy to see that in $\mathbb{R}$ we have $0 < 1$; because $\{r < 0\} \subset \{r < 1\}$, and

$$\frac{1}{2} \in \{r \in \mathbb{Q} : r < 1\} - \{r \in \mathbb{Q} : r < 0\}.$$

So in particular we have $0 \neq 1$.

Remark. Note that if $x = A|B$ and $x > 0$, then $\{r < 0\} \subsetneq A$. Therefore there must be a nonnegative rational number r such that $r \in A$. However, A does not

have a largest element. So 0 cannot be the only nonnegative rational number in A. Hence there must be a positive rational number r such that $r \in A$. In addition, note that every element of B is positive.

Theorem 6.3. *The addition of real numbers has the following properties: for every $x, y, z \in \mathbb{R}$ we have*

(i) *Associativity* :
$$x + (y + z) = (x + y) + z.$$

(ii) *Commutativity* :
$$x + y = y + x.$$

(iii) *Identity element* :
$$x + 0 = x.$$

(iv) *Additive inverse* :
$$x + (-x) = 0.$$

(v) *If $x < y$ then $x + z < y + z$.*

Proof. Let $x = A|B$, $y = C|D$, and $z = E|F$.

(i) Suppose $x + (y + z) = G|H$ and $(x + y) + z = I|J$. Then we have

$$\begin{aligned} r \in G &\iff \exists a \in A\ \exists c \in C\ \exists e \in E\ r = a + (c + e) \\ &\iff \exists a \in A\ \exists c \in C\ \exists e \in E\ r = (a + c) + e \iff r \in I. \end{aligned}$$

Hence $G = I$. Thus $x + (y + z) = (x + y) + z$.

(ii) Suppose $x + y = G|H$ and $y + x = I|J$. Then we have

$$\begin{aligned} r \in G &\iff \exists a \in A\ \exists c \in C\ r = a + c \\ &\iff \exists c \in C\ \exists a \in A\ r = c + a \iff r \in I. \end{aligned}$$

Hence $G = I$. Thus $x + y = y + x$.

(iii) Suppose $x + 0 = G|H$. Let $r \in G$. Then there are $a \in A$ and $c < 0$ such that $r = a + c$. Thus $r < a$. Hence $r \in A$. So $G \subset A$. Conversely, let $p \in A$. Then there is $q > p$ such that $q \in A$. Now we have $p = q + (p - q)$, and $p - q < 0$. Therefore $p \in G$. Hence $A \subset G$. So $A = G$. Thus we have $x + 0 = x$.

(iv) Suppose $-x = B'|A'$, and $x + (-x) = G|H$. Let $r \in G$. Then there are $a \in A$ and $b' \in B'$ such that $r = a + b'$. We know that $-b' \in B$. Therefore we have $a < -b'$. Hence $r = a + b' < 0$. Thus $G \subset \{r < 0\}$.

Conversely, suppose $r < 0$. We claim that there is $a \in A$ such that $a - r \in B$. Because otherwise for every $a \in A$ we would have $a - r \in A$. Hence we would also have $a - 2r = (a - r) - r \in A$. In fact, by induction we can show that $a - nr \in A$ for every $n \in \mathbb{N}$. But this leads to a contradiction, since we can make $a - nr$ larger than

any rational number by taking n large enough (this is known as the Archimedean property of $\mathbb{Q}$). To see this let $b \in B$. Then $b > a$. We know that $-r, b-a$ are positive rational numbers. So we have $-r = m/k$ and $b - a = l/j$, for some $m, k, l, j \in \mathbb{N}$. Now for $n = kl$ we have $-nr = n(-r) = kl \times m/k = lm \geq l/j = b-a$, since $mj \geq 1$. Hence we would get $b \leq a - nr$, which implies that $b \in A$; and this is a contradiction.

Therefore there is $a \in A$ such that $a - r \in B$. Now we need $r - a = -(a - r)$ to be in B'. The only obstruction is that $a - r$ might be the smallest element of B. To solve this problem let $c \in A$ be such that $c > a$. Then we also have $c - r \in B$, since $a - r < c - r$. Let $b := c - r$. Then b is not the smallest element of B. Hence $-b \in B'$. Now we have

$$r = c + r - c = c + (-b) \in G.$$

Thus we have shown that $\{r < 0\} \subset G$. Therefore $G = \{r < 0\}$, and we have $x + (-x) = 0$, as desired.

(v) Suppose $x < y$, $x + z = G|H$, and $y + z = I|J$. Then we know that $A \subset C$. Let $r \in G$. Then we have $r = a + e$, for some $a \in A$ and $e \in E$. However, we also have $a \in C$. Hence $r \in I$ too. Thus $G \subset I$. So we have $x + z \leq y + z$. But $x + z = y + z$ implies that

$$\begin{aligned} x = x + 0 = x + (z + (-z)) &= (x + z) + (-z) \\ &= (y + z) + (-z) = y + (z + (-z)) = y + 0 = y; \end{aligned}$$

which contradicts our assumption. So we must have $x + z \neq y + z$. Therefore $x + z < y + z$, as desired. ■

Remark. As a consequence of the above theorem, for every $x, y \in \mathbb{R}$ we have

$$-(-x) = x, \qquad -(x + y) = (-x) + (-y).$$

These are proved in Section 5.2 for arbitrary rings. (Note that we did not use the multiplication of ring in their proofs. Alternatively, we have proved these results in Section 5.6 for binary operations.) Thus, as a trivial consequence we have $y = -x$ if and only if $-y = x$. In addition, we have

$$x < 0 \qquad \Longleftrightarrow \qquad -x > 0.$$

Because $x < 0 \implies x + (-x) < 0 + (-x) \implies 0 < -x$, and $0 < -x \implies 0 + x < (-x) + x \implies x < 0$.

The next step is to define multiplication and inverse, and to conclude their properties. Let $x = A|B, y = C|D \in \mathbb{R}$. We have seen that a suitable candidate for the product of x, y should be a cut whose first component is

$$\{ac : a \in A, c \in C\}.$$

Because if $a < x$ and $c < y$, then we have $ac < xy$, provided that $a, c > 0$. However as we noted before, a, c can be negative, and consequently ac can be a large positive number greater than xy. In order to overcome this difficulty, we first assume that $x, y > 0$. Then we can represent the positive rational numbers less than xy by ac, where $a, c > 0$. And to construct the first component of the Dedekind cut of xy we also include all nonpositive rational numbers, since they are all less than xy. Finally we can extend the definition of multiplication to all real numbers by taking the opposites, and reducing the general case to the case of positive real numbers, as explained below.

In addition, to find the inverse of a given nonzero real number like $x = A|B$, we first assume $x > 0$. Intuitively, we know that if the real number $\tilde{B}|\tilde{A}$ is the inverse of x, then $\tilde{B}$ consists of rational numbers s less than x^{-1}. However for $s > 0$ we have $s < x^{-1}$ if and only if $s^{-1} > x$. Therefore we must have $s^{-1} \in B$. In other words, the cut whose first component, i.e. $\tilde{B}$, consists of nonpositive rational numbers together with the inverse of the elements of B (which are all positive, since they are not less than x) is our candidate for x^{-1}. But we must note that B can have a least element, and $\tilde{B}$ cannot have a largest element. Therefore when we construct $\tilde{B}$ from B we have to exclude the least element of B, if it exists. Hence we arrive at the following definition. For the inverse of negative real numbers we can take their opposites and use the definition of inverse of positive numbers, as explained below. But first let us check that these proposed definitions for the product and inverse of positive numbers are actually Dedekind cuts.

Proposition 6.3. *Let $x, y \in \mathbb{R}$, and suppose $x = A|B$, $y = C|D$. Suppose $x, y > 0$. Let*

$$G = \{r \in \mathbb{Q} : r \leq 0\} \cup \{ac : a \in A, c \in C,\ a, c > 0\}, \qquad H = \mathbb{Q} - G.$$

Also let

$$\tilde{B} = \{r \in \mathbb{Q} : r \leq 0\} \cup \{b^{-1} : b \in B,\ b \text{ is not the smallest element of } B\},$$
$$\tilde{A} = \mathbb{Q} - \tilde{B}.$$

Then $G|H$ and $\tilde{B}|\tilde{A}$ are real numbers.

Remark. Note that when $x > 0$, A contains a positive rational number; so every element of B is positive, since every element of B is greater than every element of A. Thus the definition of $\tilde{B}$ makes sense.

Proof. It is obvious that G is nonempty, $G \cap H = \emptyset$, and $G \cup H = \mathbb{Q}$. Now let $b \in B$ and $d \in D$. Let $a \in A$ and $c \in C$ be positive. We know that $a < b$ and $c < d$. Hence we have $ac < bd$. So $bd \notin G$. Thus $H \neq \emptyset$. Next let $g \in G$ and $h \in H$. Note that by definition every element of H is positive. Thus if $g \leq 0$ then

$g < h$. So suppose $g = ac > 0$. If $h \le ac$ then $h/a \le c$. Thus $h/a \in C$. Hence we have $h = a \times h/a \in G$, which is a contradiction. Therefore we must have $ac < h$. Finally, let ac be a positive element of G. Then there are $p \in A$ and $q \in C$ such that $a < p$ and $c < q$; because A, C do not have largest elements. Now we have $pq \in G$, and $ac < pq$. Thus no element ac can be the largest element of G. It is also obvious that a nonpositive element of G cannot be its largest element.

Next lest us consider $\tilde{B}, \tilde{A}$. Note that all the elements of B are positive, since $x > 0$. It is obvious that $\tilde{B}$ is nonempty, $\tilde{B} \cap \tilde{A} = \emptyset$, and $\tilde{B} \cup \tilde{A} = \mathbb{Q}$. In addition, for a positive $a \in A$ we have $a^{-1} > 0$; therefore $a^{-1} \in \tilde{A}$. Thus $\tilde{A} \neq \emptyset$. Now let $b' \in \tilde{B}$ and $a' \in \tilde{A}$. Note that by definition every element of $\tilde{A}$ is positive. Thus if $b' \le 0$ then $b' < a'$. So suppose $b' > 0$. Then $b'^{-1} \in B$. Also, a'^{-1} either belongs to A, or is the smallest element of B (otherwise a' belongs to $\tilde{B}$). In either case we have $a'^{-1} < b'^{-1}$. Hence $b' < a'$, since $a', b' > 0$.

Finally, note that if $b' \in \tilde{B}$ is its largest element, then b' must be positive. So $b'^{-1} \in B$, and it must be less than or equal to every element of B except its smallest element (if B has a smallest element). Now if B does not have a smallest element, then b'^{-1} is less than or equal to every element of B. Hence b'^{-1} is the smallest element of B, contrary to our assumption. So suppose $p \in B$ is its smallest element. Then we have $p < b'^{-1}$. Let q be a rational number such that $p < q < b'^{-1}$. Then we have $q \in B$, since $p < q$. However, q is not the smallest element of B. So it cannot be smaller than b'^{-1}. Thus we have a contradiction, and therefore $\tilde{B}$ cannot have a largest element. ■

Definition 6.5. Let $x, y \in \mathbb{R}$, and suppose $x = A|B$, $y = C|D$. When $x, y > 0$, we define the **multiplication** of x, y to be $xy := G|H$, where

$$G = \{r \in \mathbb{Q} : r \le 0\} \cup \{ac : a \in A, c \in C,\ a, c > 0\}, \qquad H = \mathbb{Q} - G.$$

In other cases, we define

(i) When $x < 0$ and $y > 0$, $xy := -((-x)y)$.
(ii) When $x > 0$ and $y < 0$, $xy := -(x(-y))$.
(iii) When $x < 0$ and $y < 0$, $xy := (-x)(-y)$.
(iv) For every $x \in \mathbb{R}$, $x0 := 0$ and $0x := 0$.

Suppose $x \neq 0$. When $x > 0$ the **inverse** of x is $x^{-1} := \tilde{B}|\tilde{A}$, where

$$\tilde{B} = \{r \in \mathbb{Q} : r \le 0\} \cup \{b^{-1} : b \in B,\ b \text{ is not the smallest element of } B\},$$
$$\tilde{A} = \mathbb{Q} - \tilde{B}.$$

And when $x < 0$ we define $x^{-1} := -(-x)^{-1}$.

Remark. Note that when $x, y > 0$ we have $xy > 0$. Hence if one of the x, y is positive, and the other one is negative, we have $xy < 0$; and if both x, y are negative we have $xy > 0$. Also note that when $x > 0$ we have $x^{-1} > 0$. Thus when $x < 0$ we have $x^{-1} < 0$.

Theorem 6.4. *The multiplication of real numbers has the following properties: for every $x, y, z \in \mathbb{R}$ we have*

(i) *Associativity* :
$$x(yz) = (xy)z.$$

(ii) *Commutativity* :
$$xy = yx.$$

(iii) *Identity element* :
$$x1 = x.$$

(iv) *Multiplicative inverse* :
$$x \neq 0 \implies xx^{-1} = 1.$$

(v) *Distributivity* :
$$x(y + z) = xy + xz.$$

(vi) *If $x > 0$ and $y > 0$, then $xy > 0$.*

Remark. This theorem and the previous theorem show that $\mathbb{R}$ is an ordered field.

Proof. Let $x = A|B$, $y = C|D$, and $z = E|F$.

(i) Suppose $x(yz) = G|H$ and $(xy)z = I|J$. If one of x, y, z is zero, then both $x(yz), (xy)z$ are zero; so $x(yz) = (xy)z$. So we can assume that x, y, z are all nonzero. First suppose $x, y, z > 0$. Then we have

$$\begin{aligned} r \in G &\iff r \leq 0 \text{ or } \exists a \in A\ \exists c \in C\ \exists e \in E\ a, c, e > 0 \text{ and } r = a(ce) \\ &\iff r \leq 0 \text{ or } \exists a \in A\ \exists c \in C\ \exists e \in E\ a, c, e > 0 \text{ and } r = (ac)e \\ &\iff r \in I. \end{aligned}$$

Hence $G = I$. Thus $x(yz) = (xy)z$. Now we have (Note that the sign of xy and yz, which are needed in the following computations, can be determined from the sign of x, y, z.)

$$\begin{aligned} x < 0, y > 0, z > 0 \implies x(yz) &= -[(-x)(yz)] = -[((-x)y)z] \\ &= -[(-(xy))z] = (xy)z, \\ x > 0, y < 0, z > 0 \implies x(yz) &= -[x(-(yz))] = -[x((-y)z)] \\ &= -[(x(-y))z] = -[(-(xy))z] = (xy)z, \\ x < 0, y < 0, z > 0 \implies x(yz) &= (-x)(-(yz)) = (-x)((-y)z)) \\ &= ((-x)(-y))z = (xy)z, \\ x > 0, y > 0, z < 0 \implies x(yz) &= -[x(-(yz))] = -[x(y(-z))] \\ &= -[(xy)(-z)] = (xy)z, \end{aligned}$$

$$
\begin{aligned}
x<0, y>0, z<0 \implies x(yz) &= (-x)(-(yz)) = (-x)(y(-z)) \\
&= ((-x)y)(-z) = (-(xy))(-z) = (xy)z, \\
x>0, y<0, z<0 \implies x(yz) &= x((-y)(-z)) = (x(-y))(-z) \\
&= (-(xy))(-z) = (xy)z, \\
x<0, y<0, z<0 \implies x(yz) &= -[(-x)(yz)] = -[(-x)((-y)(-z))] \\
&= -[((-x)(-y))(-z)] = -[(xy)(-z)] = (xy)z.
\end{aligned}
$$

(ii) Suppose $xy = G|H$ and $yx = I|J$. If one of x, y is zero, then both xy, yx are zero; so $xy = yx$. So we can assume that both x, y are nonzero. First suppose $x, y > 0$. Then we have

$$
\begin{aligned}
r \in G &\iff r \le 0 \text{ or } \exists a \in A\ \exists c \in C\ a, c > 0 \text{ and } r = ac \\
&\iff r \le 0 \text{ or } \exists c \in C\ \exists a \in A\ c, a > 0 \text{ and } r = ca \iff r \in I.
\end{aligned}
$$

Hence $G = I$. Thus $xy = yx$. Now we have

$$
\begin{aligned}
x<0, y>0 &\implies xy = -((-x)y) = -(y(-x)) = yx, \\
x>0, y<0 &\implies xy = -(x(-y)) = -((-y)x) = yx, \\
x<0, y>0 &\implies xy = (-x)(-y) = (-y)(-x) = yx.
\end{aligned}
$$

(iii) Suppose $x1 = G|H$. If $x = 0$ we have $x1 = 0 \times 1 = 0 = x$. Let us assume $x > 0$. Let $r \in G$. If $r \le 0$ then we have $r \in A$, since $x > 0$. So suppose $r > 0$. Then there are $a \in A$ and $c < 1$ such that $a, c > 0$, and $r = ac$. Hence we have $r = ac < a1 = a$. Thus $r \in A$. So $G \subset A$. Conversely, let $p \in A$. If $p \le 0$ then $p \in G$ by definition. So suppose $p > 0$. Then there is $q > p$ such that $q \in A$. Now we have $p = q \times p/q$, and $0 < p/q < 1$. Therefore $p \in G$. Hence $A \subset G$. So $A = G$. Thus we have $x1 = x$. Finally, let us assume $x < 0$. Then we have

$$x1 = -((-x)1) = -(-x) = x,$$

since $-x, 1 > 0$.

(iv) Suppose $x \neq 0$, $x^{-1} = \tilde{B}|\tilde{A}$, and $xx^{-1} = G|H$. First let us assume that $x > 0$. Note that in this case all the elements of B are positive. Let $r \in G$. If $r \le 0$ then $r < 1$. So suppose $r > 0$. Then there are $a \in A$ and $b' \in \tilde{B}$ such that $a, b' > 0$, and $r = ab'$. We know that $b'^{-1} \in B$. Therefore we have $a < b'^{-1}$. Hence $r = ab' < 1$, since $b' > 0$. Thus $G \subset \{r < 1\}$.

Conversely, suppose $r < 1$. If $r \le 0$ then $r \in G$ by definition. So suppose $r > 0$. Note that A contains some positive rational numbers, since $x > 0$. We claim that there is $a \in A$ such that $a > 0$, and $a/r \in B$. Because otherwise for every positive $a \in A$ we would have $a/r \in A$. Hence we would also have $a/r^2 = (a/r)/r \in A$, since a/r is positive too. In fact, by induction we can show that $a/r^n \in A$ for every

$n \in \mathbb{N}$. But this leads to a contradiction, since we can make a/r^n larger than any rational number by taking n large enough (this also follows from the Archimedean property of $\mathbb{Q}$, but we have to convert the multiplicative form a/r^n to the additive form $a + nd$ for some positive rational number d). To see this note that $r < 1$, so $s := 1 - r > 0$. It is easy to show by induction that

$$\frac{1}{(1-s)^n} \geq 1 + ns.$$

Because for $n = 0$ both sides are 1. And if the inequality holds for some n, then for $n + 1$ we have

$$\frac{1}{(1-s)^{n+1}} = \frac{1}{1-s}\frac{1}{(1-s)^n} \geq \frac{1+ns}{1-s} \geq 1 + (n+1)s,$$

because we have

$$(1-s)(1+(n+1)s) = 1 - s + (n+1)s - (n+1)s^2 \leq 1 + ns,$$

since $1 - s, (n+1)s^2$ are positive. Therefore we get

$$\frac{a}{r^n} = \frac{a}{(1-s)^n} \geq a(1+ns) = a + nas.$$

Let $b \in B$. Then $b > a$. We know that $as, b - a$ are positive rational numbers. So we have $as = m/k$ and $b - a = l/j$, for some $m, k, l, j \in \mathbb{N}$. Now for $n = kl$ we have $nas = kl \times m/k = lm \geq l/j = b - a$, since $mj \geq 1$. Thus we get

$$b \leq a + nas \leq \frac{a}{r^n}.$$

However, this implies that $b \in A$, which is a contradiction.

Therefore there is a positive $a \in A$ such that $a/r \in B$. Now we need $r/a = (a/r)^{-1}$ to be in $\tilde{B}$. The only obstruction is that a/r might be the smallest element of B. To solve this problem let $c \in A$ be such that $c > a$. Then we also have $c/r \in B$, since $a/r < c/r$. Let $b := c/r$. Then b is not the smallest element of B. Hence $b^{-1} \in \tilde{B}$. Now we have

$$r = c \times \frac{r}{c} = c \times \left(\frac{c}{r}\right)^{-1} = cb^{-1} \in G.$$

Thus we have shown that $\{r < 1\} \subset G$. Therefore $G = \{r < 1\}$, and we have $xx^{-1} = 1$, as desired. Finally, suppose $x < 0$. Then $-x > 0$, and by definition we have $x^{-1} = -(-x)^{-1}$. Hence $x^{-1} < 0$ too, and we have $-x^{-1} = -(-(-x)^{-1}) = (-x)^{-1}$. Thus we get

$$xx^{-1} = (-x)(-x^{-1}) = (-x)(-x)^{-1} = 1,$$

as desired.

(v) Suppose $x(y+z) = G|H$, $xy + xz = I|J$, $y + z = K|K'$, $xy = L|L'$, and $xz = M|M'$. If $x = 0$ then

$$x(y+z) = 0 = 0 + 0 = xy + xz.$$

And if one of y, z is zero, say y is zero, then

$$x(y+z) = x(0+z) = xz = 0 + xz = xy + xz.$$

So we can assume that x, y, z are all nonzero. First suppose $x, y, z > 0$. Note that $y + z > 0 + z = z > 0$. In addition, note that $xy, xz > 0$. Hence we can similarly show that $xy + xz > 0$. Let $r \in G$. If $r \leq 0$ then $r \in I$, since $xy + xz > 0$. So suppose $r > 0$. Then there are $a \in A$ and $k \in K$ such that $a, k > 0$, and $r = ak$. On the other hand, there are $c \in C$ and $e \in E$ such that $k = c + e$. Thus we have $r = ac + ae$. If $c \leq 0$ then $ac \leq 0$; so $ac \in L$, since $xy > 0$. And if $c > 0$ then $ac \in L$ by definition of xy. Similarly, we have $ae \in M$. Therefore $r = ac + ae \in I$. Hence $G \subset I$.

Conversely, let $r \in I$. Then we know that $r = l + m$, for some $l \in L$ and $m \in M$. Suppose $l, m > 0$. Then there are positive numbers $a, a' \in A$, $c \in C$, and $e \in E$ such that $l = ac$, and $m = a'e$. Without loss of generality we can assume $a' \leq a$. Then we have

$$r = l + m = ac + a'e \leq ac + ae = a(c + e).$$

However, $c + e \in K$, and $c + a > 0$; so $a(c + e) \in G$. Therefore we also have $r \in G$. Next suppose one of l, m is positive and the other one is nonpositive, say $l > 0, m \leq 0$. Then there are positive numbers $a \in A$ and $c \in C$ such that $l = ac$. Let $e \in E$ be positive. Then we have

$$r = l + m \leq l = ac \leq ac + ae = a(c + e);$$

which implies $r \in G$. Finally, if $l, m \leq 0$ then $r \leq 0$. Thus by definition of G we have $r \in G$. Therefore we have $I \subset G$. Hence $I = G$, and we get $x(y+z) = xy+xz$, as desired.

Next suppose $x > 0$, $y < 0$, and $z > 0$. Then $-y > 0$. We have to consider two cases. If $y + z \geq 0$ we have

$$-xy + x(y + z) = x(-y) + x(y + z) = x(-y + y + z) = xz.$$

Hence we get $x(y+z) = xy+xz$ by adding xy to both sides. And if $y+z < 0$ then $-(y + z) > 0$. Thus we have

$$\begin{aligned} -x(y + z) + xz &= x(-(y + z)) + xz \\ &= x(-(y + z) + z) = x(-y - z + z) = x(-y) = -xy. \end{aligned}$$

Hence we get $x(y+z) = xy+xz$ by adding $xy, x(y+z)$ to both sides. Now suppose $x > 0$, $y > 0$, and $z < 0$. Then we can switch y, z and use the previous case to obtain

$$x(y+z) = x(z+y) = xz + xy = xy + xz.$$

Next suppose $x > 0$ and $y, z < 0$. Then we have $y + z < 0$. Hence

$$\begin{aligned} x(y+z) = -[x(-(y+z))] &= -[x((-y)+(-z))] = -[x(-y)+x(-z)] \\ &= -[(-xy)+(-xz)] = -[-(xy+xz)] = xy+xz. \end{aligned}$$

Finally, suppose $x < 0$. Let us first show that for every $y \in \mathbb{R}$ we have

$$(-x)y = -xy. \qquad (*)$$

If $y = 0$ then both sides of the above equation are 0. If $y > 0$ then by definition we have $xy = -((-x)y)$. Thus $-xy = (-x)y$. And if $y < 0$ then by definition we have $xy = (-x)(-y)$. We also have $(-x)y = -((-x)(-y))$, since $-x > 0$. Hence we get $(-x)y = -xy$, as desired. Now by the previous paragraph, for every $y, z \in \mathbb{R}$ we have

$$(-x)(y+z) = (-x)y + (-x)z,$$

since $-x > 0$. Hence by $(*)$ we get

$$-(x(y+z)) = (-xy) + (-xz) = -(xy+xz).$$

Thus we obtain $x(y+z) = xy + xz$; because the additive inverse of every number is unique.

(vi) As we have noted before, this is a trivial consequence of the definition of multiplication. ■

Theorem 6.5. *Suppose $S \subset \mathbb{R}$ is nonempty and bounded below. Then S has a greatest lower bound in $\mathbb{R}$.*

Proof. Let

$$T := \{-x : x \in S\}.$$

Then T is nonempty. Also, if z is a lower bound for S then $-z$ is an upper bound for T. So T is bounded above. Hence it has a least upper bound, which we call y. We claim that $-y$ is the greatest lower bound of S. First note that for every $x \in S$ we have $-x \in T$; so $-x \leq y$. Thus $x \geq -y$. So $-y$ is a lower bound for S. Now let w be a lower bound for S. Then $-w$ is an upper bound for T. Hence $-w \geq y$. Therefore $w \leq -y$. Thus $-y$ is the greatest lower bound of S, as desired. ■

Although $\mathbb{R}$ does not contain $\mathbb{Q}$ as a subset, it has a subset which looks like $\mathbb{Q}$. Informally, we can say that $\mathbb{R}$ contains a "copy" of $\mathbb{Q}$. The next theorem shows that this subset of $\mathbb{R}$ behaves similarly to $\mathbb{Q}$.

Theorem 6.6. *Let $E : \mathbb{Q} \to \mathbb{R}$ be defined as*

$$E(p) = \{r \in \mathbb{Q} : r < p\} \,|\, \{r \in \mathbb{Q} : r \geq p\},$$

for every $p \in \mathbb{Q}$. Then E is a one-to-one function, and for every $p, q \in \mathbb{Q}$ we have

(i) $E(p+q) = E(p) + E(q)$.
(ii) $E(pq) = E(p)E(q)$.
(iii) $p < q$ *if and only if* $E(p) < E(q)$.

Remark. Note that in the relation $E(p+q) = E(p) + E(q)$, p, q are added using the addition of $\mathbb{Q}$, and $E(p), E(q)$ are added using the addition of $\mathbb{R}$. So in some sense, we can say that the function E transforms the addition of $\mathbb{Q}$ into the addition of $\mathbb{R}$. Similar remarks apply to the multiplication and order relation.

Proof. First note that E is one-to-one. Because if $E(p) = E(q)$ then we have $\{r \geq p\} = \{r \geq q\}$. But both p, q are the minimum of this set. So we must have $p = q$.

(i) Suppose $E(p) + E(q) = A|B$. Let $r \in A$. Then there are $a < p$ and $b < q$ such that $r = a + b$. Hence we have $r = a + b < p + q$. On the other hand, suppose $r < p + q$. Then $r - p < q$. So there is b such that $r - p < b < q$. Hence we have $r - b < p$. Thus

$$r = r - b + b \in A.$$

Therefore we have $A = \{r < p + q\}$, as desired.

(ii) Note that by definition we have $E(0) = 0$. Thus for every p we have

$$-E(p) = E(-p),$$

since $E(p) + E(-p) = E(p + (-p)) = E(0) = 0$. Now let $E(p)E(q) = C|D$. If one of p, q is zero, then both $E(pq)$ and $E(p)E(q)$ are zero. So suppose that both p, q are nonzero. First let us assume that $p, q > 0$. Then by the next part we have $E(p), E(q) > E(0) = 0$. Let $r \in C$. If $r \leq 0$ then we obviously have $r < pq$. So suppose $r > 0$. Then there are $a < p$ and $b < q$ such that $a, b > 0$, and $r = ab$. Thus we have $r = ab < pq$. On the other hand, suppose $r < pq$. If $r \leq 0$ then by definition we have $r \in C$. So suppose $r > 0$. Then we have $r/p < q$. So there is b such that $r/p < b < q$. Hence we have $r/b < p$, since $b > 0$. Thus

$$r = \frac{r}{b} \times b \in C.$$

Therefore we have $C = \{r < pq\}$, as desired.

Next suppose $p > 0$ and $q < 0$. Then $E(q) < 0$. Hence we have

$$\begin{aligned} E(p)E(q) &= -[E(p)(-E(q))] = -[E(p)E(-q)] \\ &= -E(p(-q)) = -E(-pq) = -(-E(pq)) = E(pq). \end{aligned}$$

Now suppose $p < 0$ and $q > 0$. Then we can switch p, q and use the previous case to obtain

$$E(p)E(q) = E(q)E(p) = E(qp) = E(pq).$$

Finally, suppose $p, q < 0$. Then $E(p), E(q) < 0$. Hence we have

$$E(p)E(q) = (-E(p))(-E(q)) = E(-p)E(-q) = E((-p)(-q)) = E(pq).$$

(iii) If $p < q$ then we clearly have $\{r < p\} \subset \{r < q\}$. In addition, we have $p \in \{r < q\} - \{r < p\}$. Hence $\{r < p\} \subsetneqq \{r < q\}$. Thus $E(p) < E(q)$. Conversely, if $E(p) < E(q)$ then we must have $p < q$. Because if $p \geq q$ then we have shown that $E(p) \geq E(q)$; and this contradicts our assumption. ■

Definition 6.6. A real number x is called **rational** if there is $p \in \mathbb{Q}$ such that $x = E(p)$. In other words, if $x = A|B$ and

$$A = \{r \in \mathbb{Q} : r < p\}, \qquad B = \{r \in \mathbb{Q} : r \geq p\}.$$

A real number which is not rational is called **irrational**. A real number x is an **integer**, if $x = E(p)$ for some $p \in \mathbb{Q}$, and $p = E(n)$ for some $n \in \mathbb{Z}$, where $E(n)$ is defined in Theorem 5.30. If x is an integer corresponding to a nonnegative integer n, we say x is a **natural number**.

Remark. If x is a rational real number as above, we identify it with the rational number p. And if x is an integer as above, we identify it with the integer n. Note that these identifications are mostly harmless, since by Theorems 5.30 and 6.6, the addition, multiplication, and order relation will be preserved under these identifications.

***Example* 6.1.** An example of an irrational number is $x = A|B$ where

$$A = \{r \in \mathbb{Q} : r < 0 \text{ or } r^2 < 2\}, \qquad B = \{r \in \mathbb{Q} : r > 0 \text{ and } r^2 \geq 2\}.$$

The number x is actually $\sqrt{2}$, and we will see later that it is indeed irrational.

Remark. Sometimes we abuse the notation, and denote the set of rational numbers in $\mathbb{R}$ by $\mathbb{Q}$, and the set of integers in $\mathbb{R}$ by $\mathbb{Z}$. We may also denote the set of positive integers in $\mathbb{R}$ by $\mathbb{N}$. However, we will NOT use the notation ω for the set of nonnegative integers in $\mathbb{R}$.

Theorem 6.7. *Let X be one of the sets $\mathbb{N}$, $\mathbb{Z}$, or $\mathbb{Q}$, considered as a subset of $\mathbb{R}$. Then X is closed under addition and multiplication of real numbers, i.e. for every $a, b \in X$ we have $a + b \in X$ and $ab \in X$.*

Proof. Let Y be one of the sets $\mathbb{N}$, $\mathbb{Z}$, or $\mathbb{Q}$ (as we have constructed them, not as subsets of $\mathbb{R}$). We have shown that there is a one-to-one map $E : Y \to \mathbb{R}$ that preserves addition, multiplication, and order. Note that depending on Y, E might be the composition of several of the maps we constructed in Theorems 5.11, 5.30, and 6.6. Now the set X in the theorem is actually the image of E. Let $a, b \in X$. Then there are $r, s \in Y$ such that $E(r) = a$ and $E(s) = b$. Hence we have

$$a + b = E(r) + E(s) = E(r + s) \in E(Y) = X,$$
$$ab = E(r)E(s) = E(rs) \in E(Y) = X,$$

as desired. ∎

The map E considered in the proof of the above theorem can be used to prove that the natural, integer, and rational numbers in $\mathbb{R}$ have the same properties that we have proved before. For example, for any integer $n \in \mathbb{R}$ there cannot be an integer $k \in \mathbb{R}$ such that

$$n < k < n + 1.$$

Because otherwise we can find integers in $\mathbb{Z}$ that have the same property, i.e. there are $m, l, j \in \mathbb{Z}$ such that $E(m) = n$, $E(l) = k$, and $E(j) = n + 1$. We also have $m < l < j$, since E preserves order. However, it is easy to see that $E(1) = 1$. Hence $E(j) = E(m) + E(1) = E(m + 1)$. Thus $j = m + 1$, since E is one-to-one. But now we have $m < l < m + 1$, which is a contradiction.

As another example, suppose $p \in \mathbb{R}$ is rational. Then there are integers $m, n \in \mathbb{R}$ such that $n > 0$, and $p = \frac{m}{n}$. Because we have $p = E(r)$, for some $r \in \mathbb{Q}$. We also have $r = \frac{a}{b}$, where a, b are integers, and $b > 0$. Hence we get

$$p = E(r) = E\big(\frac{a}{b}\big) = \frac{E(a)}{E(b)}.$$

Note that E preserves division too, since we have $E(b)E(\frac{a}{b}) = E(b\frac{a}{b}) = E(a)$. Also note that $E(b) \neq 0$, since $b \neq 0$, E is one-to-one, and we can easily see that $E(0) = 0$.

In addition, it is easy to show that the well-ordering of ω implies the following property for sets of integers in $\mathbb{R}$. However, we will show that this property can be deduced from the fact that $\mathbb{R}$ is a complete ordered field.

Theorem 6.8. *Let A be a set of integers in $\mathbb{R}$, and suppose A is nonempty.*
(i) *If A is bounded below then it has a least element.*
(ii) *If A is bounded above then it has a largest element.*

Proof. **(i)** Since A is nonempty and bounded below, it has a greatest lower bound, which we call m. Then $m + \frac{1}{2}$ cannot be a lower bound for A. Hence there is $n \in A$

such that $n < m + \frac{1}{2}$. Then we have $n - 1 < m - \frac{1}{2} < m$. We also have $m \leq n$. Thus

$$n - 1 < m \leq n.$$

Now if $m \neq n$ then n cannot be a lower bound for A. Therefore there must be $k \in A$ such that $k < n$. However $k \geq m$ too. So we get $n - 1 < k < n$, which contradicts the assumption of k, n being integers. Hence we must have $m = n$; which implies $m \in A$. Therefore m is the least element of A, since it belongs to A, and it is a lower bound for A.

(ii) This part can be proved similarly to the previous part. We can also use the set $B := \{-a : a \in A\}$, and repeat the proof of Theorem 5.12. ■

6.2 More about Real Numbers

Definition 6.7. An ordered field F is **complete**, if every nonempty and bounded above subset $S \subset F$ has a least upper bound in F.

In the last section, we have shown that $\mathbb{R}$ is a complete ordered field. In fact, up to isomorphism, $\mathbb{R}$ is the only complete ordered field, i.e. every complete ordered field is essentially the same as $\mathbb{R}$. Intuitively, this means that every complete ordered field can be obtained from $\mathbb{R}$ by renaming its elements. At the end of this section we will state this fact precisely and prove it.

***Example* 6.2.** $\mathbb{Q}$ is an ordered field which is not complete. For example

$$\{r \in \mathbb{Q} : r \geq 0,\ r^2 < 2\}$$

is a nonempty bounded above subset of $\mathbb{Q}$ which has no least upper bound in $\mathbb{Q}$. Because, as we will show in the proof of Theorem 6.15, if the above set has a supremum, its supremum must be a square root of 2. However, in Theorem 6.19 we will show that no rational number can be a square root of 2.

Although we have constructed real numbers by using Dedekind cuts, in practice we do not think of a real number as a pair of sets of rational numbers. Rather, as it is common, we think of real numbers as the points of a line. And when we want to prove something about real numbers, we will use the fact that $\mathbb{R}$ is a complete ordered field.

The Extended Real Number System. There does not exist a largest or smallest real number, but we consider the so-called "infinite real numbers" $+\infty$ and $-\infty$. They are called **(positive) infinity** and **negative infinity** respectively. We also

denote $+\infty$ simply by ∞. Technically, $\pm\infty$ are two distinct objects that are different from all real numbers. In contrast to infinities, the elements of $\mathbb{R}$ are called "finite real numbers". Also, the set

$$\mathbb{R} \cup \{+\infty, -\infty\}$$

is called "the extended real number system". We extend the order of $\mathbb{R}$ to $\mathbb{R} \cup \{+\infty, -\infty\}$, so that for all $x \in \mathbb{R}$ we have

$$-\infty < x < +\infty.$$

We also define

$$\begin{aligned}
&\pm\infty + x = x + (\pm\infty) = \pm\infty, && \pm\infty + (\pm\infty) = \pm\infty,\\
&\pm\infty \cdot x = x \cdot (\pm\infty) = \pm\infty \text{ if } x > 0, && \pm\infty \cdot (+\infty) = \pm\infty,\\
&\pm\infty \cdot x = x \cdot (\pm\infty) = \mp\infty \text{ if } x < 0, && \pm\infty \cdot (-\infty) = \mp\infty,\\
&-(\pm\infty) = \mp\infty, && \frac{x}{\pm\infty} = 0.
\end{aligned}$$

Note that the following expressions are not defined

$$\pm\infty - (\pm\infty), \qquad \pm\infty + (\mp\infty), \qquad \frac{\pm\infty}{\pm\infty},$$

where in the last expression the infinities in the numerator and denominator can have the same or the opposite signs. Also, $0 \cdot (\pm\infty)$ is sometimes defined to be zero. But we postpone using this convention until needed, and for now consider $0 \cdot (\pm\infty)$ to be undefined. ■

Definition 6.8. Let $A \subset \mathbb{R}$. When A is nonempty and bounded above, we denote its least upper bound by $\sup A$, and we call it the **supremum** of A. When A is nonempty and has no upper bound we define

$$\sup A := +\infty.$$

When A is nonempty and bounded below, we denote its greatest lower bound by $\inf A$, and we call it the **infimum** of A. When A is nonempty and has no lower bound we define

$$\inf A := -\infty.$$

Remark. It is obvious that for a nonempty set $A \subset \mathbb{R}$ we always have $\inf A \leq \sup A$.

Definition 6.9. Suppose $a, b \in \mathbb{R}$, and $a < b$. The **closed interval** and the **open interval** with **endpoints** a, b are respectively

$$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\},$$
$$(a, b) := \{x \in \mathbb{R} : a < x < b\}.$$

We can similarly define the *half-open* (or *half-closed*) intervals

$$[a, b) := \{x \in \mathbb{R} : a \leq x < b\},$$
$$(a, b] := \{x \in \mathbb{R} : a < x \leq b\}.$$

All these intervals are called **bounded**. The **length** of these bounded intervals is the positive real number $b - a$. We also define the **unbounded intervals**, whose endpoints can be $\pm\infty$, as follows

$$(-\infty, a) := \{x \in \mathbb{R} : x < a\},$$
$$(-\infty, a] := \{x \in \mathbb{R} : x \leq a\},$$
$$(b, +\infty) := \{x \in \mathbb{R} : x > b\},$$
$$[b, +\infty) := \{x \in \mathbb{R} : x \geq b\},$$
$$(-\infty, +\infty) := \mathbb{R}.$$

An unbounded interval that contains its finite endpoint is called closed, and an unbounded interval that does not contain its finite endpoint is called open. We consider $\mathbb{R}$ to be both an open interval and a closed interval.

Archimedean Property. *For every $x \in \mathbb{R}$ there exists $n \in \mathbb{N}$ such that $x < n$. Also, for every $x \in (0, \infty)$ there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < x$.*

Proof. Suppose to the contrary that there is $x_0 \in \mathbb{R}$ such that $x_0 \geq n$ for all $n \in \mathbb{N}$. Let

$$A := \{x \in \mathbb{R} : x \geq n \ \forall n \in \mathbb{N}\}.$$

Then A is nonempty by our assumption. A is also bounded below, since for example 1 is a lower bound for it. Therefore $s := \inf A$ is a finite number. Now $s + \frac{1}{2}$ is not a lower bound for A, since s is the greatest lower bound for A. Hence there is $x_1 \in A$ such that $x_1 < s + \frac{1}{2}$. On the other hand $s - \frac{1}{2}$ does not belong to A, since s is a lower bound for A. Thus there is $n_0 \in \mathbb{N}$ such that $s - \frac{1}{2} < n_0$. But this implies that

$$x_1 < s + \frac{1}{2} < n_0 + 1 \in \mathbb{N},$$

which is a contradiction.

For the second statement, note that there is $n \in \mathbb{N}$ such that $\frac{1}{x} < n$. Thus as $\frac{1}{x} > 0$ we can deduce that $x > \frac{1}{n}$. ■

The meaning of the Archimedean property is that $\mathbb{R}$ does not contain infinitely large or infinitely small elements, since intuitively we consider $n = 1 + \cdots + 1$ to be finite, no matter how large n is. Similarly, we consider $\frac{1}{n}$ to be finite, even though it might be very small.

There is an alternative way to formulate the Archimedean property, which also states the intuition that $\mathbb{R}$ does not have infinitely large or small elements. Let $x, y > 0$ be real numbers. Then there is $n \in \mathbb{N}$ such that $nx > y$. In other words, no matter how small x is, and how large y is, in finitely many steps of length x we can surpass y. For the proof note that there is n such that $\frac{y}{x} < n$; so we have $y < nx$, since x is positive. Conversely, if this formulation of the Archimedean property holds, then by setting $x = 1$ we obtain the previous formulation.

Remark. The Archimedean property of $\mathbb{R}$ is a consequence of its least upper bound property, but the two properties are not equivalent. For example $\mathbb{Q}$ is not complete but it is Archimedean, since for every $\frac{p}{q} \in \mathbb{Q}$ we have $\frac{p}{q} < |p| + 1$. Finally we mention that there are ordered fields that are not Archimedean.

Integer Part. Let $x \in \mathbb{R}$. Then by the Archimedean property there is $n_1 \in \mathbb{N}$ such that $x < n_1$. Thus the set

$$A := \{n \in \mathbb{Z} : x < n\}$$

is nonempty and bounded below. Hence A has a smallest element, which we call $n_0 + 1$. So $n_0 \notin A$. Therefore we have

$$n_0 \leq x < n_0 + 1.$$

We set $\lfloor x \rfloor := n_0$, and call $\lfloor x \rfloor$ the **integer part** of x. Note that $\lfloor x \rfloor$ is the greatest integer less than or equal to x, and we have

$$\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1, \qquad x - 1 < \lfloor x \rfloor \leq x.$$

The function $\lfloor\ \rfloor : \mathbb{R} \to \mathbb{Z}$ is called the greatest integer function or the **floor** function. ■

Theorem 6.9. *Between any two real numbers there is a rational and an irrational number.*

Proof. Suppose $a < b$. Let $c = \frac{a+b}{2}$. It is easy to see that $a < c < b$. Then $c - a > 0$, and there is $n \in \mathbb{N}$ such that $c - a > \frac{1}{n}$. On the other hand, we have $nc - 1 < \lfloor nc \rfloor \leq nc$. Hence $a < c - \frac{1}{n} < \frac{\lfloor nc \rfloor}{n} \leq c < b$, and $\frac{\lfloor nc \rfloor}{n} \in \mathbb{Q}$.

Next, let r be a rational number between $a + \sqrt{2}$ and $b + \sqrt{2}$. Then $r - \sqrt{2}$ is an irrational number between a, b. ■

Theorem 6.10. *A subset $I \subset \mathbb{R}$ is an interval, if and only if it has more than one element; and for every $a, b \in I$ with $a < b$, and every c where $a < c < b$, we have $c \in I$.*

Proof. First suppose I is an interval with endpoints $\alpha < \beta$, where the endpoints can be $\pm\infty$. Then by Theorem 6.9, I has at least two elements. Let $a, b \in I$ with $a < b$, and let $a < c < b$. Then we have $\alpha \leq a$, and $b \leq \beta$. Hence $\alpha < c < \beta$, and therefore $c \in I$ by the definition of intervals.

Now suppose conversely that I has the specified property. Then in particular I is nonempty. Let

$$\alpha := \inf I, \qquad \beta := \sup I.$$

Then $\alpha < \beta$, since otherwise I cannot have more than one element. We claim that $(\alpha, \beta) \subset I$. Let $\alpha < c < \beta$. Then there is $c < x < \beta$. But x cannot be an upper bound for I. Hence there is $x < b \leq \beta$ such that $b \in I$. Similarly there is $\alpha \leq a < c$ such that $a \in I$. Therefore $c \in I$. Thus $(\alpha, \beta) \subset I$.

On the other hand, if $c > \beta$ then $c \notin I$, since β is an upper bound for I. Similarly I cannot contain any number less than α. Hence $I \subset [\alpha, \beta]$. Therefore I equals one of the sets

$$(\alpha, \beta), \quad [\alpha, \beta), \quad (\alpha, \beta], \quad [\alpha, \beta].$$

Thus I is an interval. Note that if one of the α, β is infinite, then we have to eliminate the sets containing it. ■

Definition 6.10. Let $x \in \mathbb{R}$. Then the **absolute value** of x is

$$|x| := \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

Theorem 6.11. *For all $x, y \in \mathbb{R}$ we have*

(i) $|x| \geq 0$, *and for $x \neq 0$ we have $|x| > 0$.*
(ii) $|-x| = |x|$.
(iii) $|x| = |y|$ *if and only if $x = \pm y$.*
(iv) $|xy| = |x||y|$.
(v) $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$ *when $y \neq 0$.*
(vi) *For $r > 0$ we have $|x| < r$ if and only if $-r < x < r$; and for $r \geq 0$ we have $|x| \leq r$ if and only if $-r \leq x \leq r$.*
(vii) $-|x| \leq x \leq |x|$.
(viii) **Triangle Inequality:**

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

(ix) $\big||x| - |y|\big| \leq |x - y|$.

Proof. **(i)** For $x > 0$ we have $|x| = x > 0$, and for $x < 0$ we have $|x| = -x > 0$. Also, $|0| = 0$.

(ii) We have

$$|-x| = \begin{cases} -x & \text{if } -x \geq 0 \\ -(-x) & \text{if } -x < 0 \end{cases} = \begin{cases} -x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ x & \text{if } x > 0 \end{cases} = |x|.$$

(iii) If $x = \pm y$ then we have $|x| = |\pm y| = |y|$. Conversely, suppose $|x| = |y|$. We need to consider four cases. If $x, y \geq 0$ then we have $x = |x| = |y| = y$. If $x, y < 0$ then $-x = |x| = |y| = -y$, so $x = y$. If $x \geq 0$ and $y < 0$ then $x = |x| = |y| = -y$. And finally if $x < 0$ and $y \geq 0$ then $-x = |x| = |y| = y$.

(iv) When $x, y \geq 0$ we have $xy \geq 0$, hence $|xy| = xy = |x||y|$. When one of the x or y is negative, we consider its opposite and apply the same argument (using part (ii)). For example when $x < 0$, and $y \geq 0$ we have

$$|xy| = |-(-x)y| = |(-x)y| = |-x||y| = |x||y|.$$

The other two cases are similar.

(v) We have $y\frac{1}{y} = 1$. Thus $|y||\frac{1}{y}| = |1| = 1$, since $1 > 0$. Therefore $|\frac{1}{y}| = \frac{1}{|y|}$. Then we have

$$|\frac{x}{y}| = |x||\frac{1}{y}| = |x|\frac{1}{|y|} = \frac{|x|}{|y|}.$$

(vi) We have

$$|x| \leq r \iff \begin{cases} x \leq r & \text{if } x \geq 0 \\ -x \leq r & \text{if } x < 0 \end{cases} \iff 0 \leq x \leq r, \quad \text{or} \quad -r \leq x < 0.$$

The other one is similar.

(vii) Let $r = |x| \geq 0$. Then since $|x| \leq |x| = r$, we get the desired by the previous part.

(viii) By the previous part we have

$$-|x| \leq x \leq |x|, \qquad -|y| \leq y \leq |y|.$$

If we add these two inequalities we get

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

Therefore by part (vi) we obtain $|x + y| \leq |x| + |y|$.

(ix) Note that by the triangle inequality we have

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

Therefore $|x| - |y| \le |x - y|$. By switching x, y we get

$$|y| - |x| \le |y - x| = |x - y| \implies |x| - |y| \ge -|x - y|.$$

Therefore $\big||x| - |y|\big| \le |x - y|$. ■

Theorem 6.12. *If for all $\epsilon > 0$ we have $|x - y| \le \epsilon$ then $x = y$.*

Proof. Suppose to the contrary that $x \ne y$. Then $|x - y| > 0$. Let $\epsilon = \frac{1}{2}|x - y|$. Then we have $|x - y| \le \frac{1}{2}|x - y|$, so $2|x - y| \le |x - y|$. Therefore we get $|x - y| \le 0$, which is a contradiction. ■

Remark. Similarly if $x \le y + \epsilon$ for all $\epsilon > 0$ then $x \le y$.

Finally let us prove that $\mathbb{R}$ is the unique complete ordered field, i.e. every complete ordered field can essentially be obtained from $\mathbb{R}$ by renaming its elements.

Theorem 6.13. *Let F be a complete ordered field. Then F is **isomorphic** to $\mathbb{R}$, i.e. there exists a one-to-one and onto function $\varphi : \mathbb{R} \to F$ such that for every $x, y \in \mathbb{R}$ we have*

(i) $\varphi(x + y) = \varphi(x) + \varphi(y)$,
(ii) $\varphi(xy) = \varphi(x)\varphi(y)$,
(iii) $x < y$ *if and only if* $\varphi(x) < \varphi(y)$.

Proof. In Theorem 5.35 we have shown that there exists a one-to-one function $\varphi : \mathbb{Q} \to F$ that satisfies the above three conditions. Also remember that the image of φ is the set of "rationals" $\frac{m}{n}$ in F, where m, n are the "integers" in F, i.e. they are constructed by adding the identity of F to itself. Furthermore, since F is complete, we can show that between any two of its elements there is a "rational" element. The proof is similar to the proof of the same property for $\mathbb{R}$, as has been demonstrated in this section. In this proof, r, s, p denote rational numbers.

Let us start by extending φ to all of $\mathbb{R}$. For $x \in \mathbb{R}$ let

$$A_x := \{\varphi(r) : r \in \mathbb{Q} \text{ and } r < x\}.$$

Then A_x is a nonempty and bounded above subset of F, since for any rational number $s > x$ we have $\varphi(s) > \varphi(r)$ for every rational number $r < x$. Now let $\hat{\varphi}(x)$ be the supremum of A_x in F, i.e.

$$\hat{\varphi}(x) := \sup A_x.$$

Note that $\sup A_x$ exists because F is complete. First let us show that $\hat{\varphi}(r) = \varphi(r)$ for every $r \in \mathbb{Q}$. Note that for every rational number $p < r$ we have $\varphi(p) < \varphi(r)$; so $\varphi(r)$ is an upper bound for A_r. Hence $\hat{\varphi}(r) \le \varphi(r)$. Suppose to the contrary that $\hat{\varphi}(r) < \varphi(r)$. Then there is a "rational" element $\varphi(s) \in F$ such that $\hat{\varphi}(r) < \varphi(s) <$

$\varphi(r)$. But then we must have $s < r$; so $s \in A_r$. However, then we get $\varphi(s) \leq \hat{\varphi}(r)$, which is a contradiction. Therefore $\hat{\varphi}(r) = \varphi(r)$. Hence $\hat{\varphi}$ is an extension of φ. So in the rest of this proof we will simply denote $\hat{\varphi}$ by φ.

Now suppose $x < y$. Then we have $A_x \subset A_y$. Hence

$$\varphi(x) = \sup A_x \leq \sup A_y = \varphi(y).$$

We know that there are rational numbers $x < p < s < y$. So $\varphi(s) \in A_y$. Also, $\varphi(p)$ is an upper bound for A_x, since for $r \in A_x$ we have $r < p$ and hence $\varphi(r) < \varphi(p)$ (as φ is increasing on $\mathbb{Q}$). Therefore

$$\varphi(x) = \sup A_x \leq \varphi(p) < \varphi(s) \leq \sup A_y = \varphi(y).$$

Thus $\varphi(x) < \varphi(y)$. Hence φ is strictly increasing on $\mathbb{R}$, and, in particular, it is one-to-one. In addition note that $\varphi(x) < \varphi(y)$ implies $x < y$ too; because otherwise we would have $x \geq y$, which implies $\varphi(x) \geq \varphi(y)$.

Next let us show that φ is onto. Let $a \in F$, and let

$$\tilde{A}_a := \{r : r \in \mathbb{Q} \text{ and } \varphi(r) < a\}.$$

Note that there is a "rational" element $a - 1 < \varphi(p) < a$, so $\varphi(p) \in \tilde{A}_a$. Also, for any "rational" element $\varphi(s) \geq a$ and every $r \in \tilde{A}_a$ we have $\varphi(r) < \varphi(s)$ and thus $r < s$. Hence $\tilde{A}_a$ is a nonempty and bounded above subset of $\mathbb{R}$. Let $x := \sup \tilde{A}_a$. We claim that $a = \varphi(x)$. First note that if $\varphi(s) \geq a$ then, as we have just shown, s is an upper bound for $\tilde{A}_a$, so $x = \sup \tilde{A}_a \leq s$. Therefore if $s < x$ then $\varphi(s) < a$. Thus a is an upper bound for A_x. So $\varphi(x) = \sup A_x \leq a$. Suppose to the contrary that $\varphi(x) < a$. Then there is $\varphi(x) < \varphi(r) < a$. Hence $r \in \tilde{A}_a$. So we must have $r \leq \sup \tilde{A}_a = x$. But $\varphi(x) < \varphi(r)$ implies that $x < r$, which is a contradiction. Therefore $\varphi(x) = a$. So φ is onto.

Next we show that

$$\varphi(x + y) = \varphi(x) + \varphi(y).$$

If this does not happen, then we either have $\varphi(x+y) < \varphi(x)+\varphi(y)$, or $\varphi(x+y) > \varphi(x) + \varphi(y)$. Suppose the latter inequality holds; the other case can be treated similarly. Then there is a "rational" element $\varphi(r)$ such that

$$\varphi(x + y) > \varphi(r) > \varphi(x) + \varphi(y).$$

Hence we get $r < x + y$; so $r - y < x$. Thus there is also a rational number s such that $r - y < s < x$. Then we have $r - s < y$. Hence we have $\varphi(s) < \varphi(x)$, and $\varphi(r - s) < \varphi(y)$. Therefore since φ preserves addition over $\mathbb{Q}$ we obtain

$$\varphi(x) + \varphi(y) > \varphi(s) + \varphi(r - s) = \varphi(s + r - s) = \varphi(r),$$

which is a contradiction.

Finally, let us show that $\varphi(xy) = \varphi(x)\varphi(y)$. Note that since φ preserves addition we have $\varphi(-x) = -\varphi(x)$. We also know that $\varphi(0) = 0$. Hence it suffices to check the equality $\varphi(xy) = \varphi(x)\varphi(y)$ for $x, y > 0$. If the equality does not hold, then either $\varphi(xy) < \varphi(x)\varphi(y)$, or $\varphi(xy) > \varphi(x)\varphi(y)$. Suppose the former inequality holds; the other case can be treated similarly. Then there is a "rational" element $\varphi(r)$ such that

$$\varphi(xy) < \varphi(r) < \varphi(x)\varphi(y).$$

Hence we get $0 < xy < r$; so $x < \frac{r}{y}$. Thus there is also a rational number s such that $0 < x < s < \frac{r}{y}$. Then we have $y < \frac{r}{s}$. Hence we have $\varphi(x) < \varphi(s)$, and $\varphi(y) < \varphi(\frac{r}{s})$. Therefore since φ preserves multiplication over $\mathbb{Q}$ we obtain

$$\varphi(x)\varphi(y) < \varphi(s)\varphi(\tfrac{r}{s}) = \varphi(s\tfrac{r}{s}) = \varphi(r),$$

which is a contradiction. ■

6.3 Powers and Roots

Let us first review the notion of power.

Definition 6.11. We define the **powers** of $a \in \mathbb{R}$ as follows. For a positive integer n we inductively define

$$a^1 := a, \ \ldots \ a^n := a^{n-1}a.$$

Here a is called the **base**, and n is called the **exponent**. If $a \neq 0$, we define

$$a^0 := 1, \qquad a^{-n} := (a^{-1})^n.$$

Remark. We also use the convention that $0^0 = 1$. This is useful in some algebraic manipulations, but we must be careful that 0^0 does not have a definite value when we deal with limits.

Theorem 6.14. *Suppose $a, b \in \mathbb{R}$, and $n, m \in \mathbb{Z}$. In all of the following statements, when the base is zero the exponent must be nonnegative.*

(i) *If $a \neq 0$ then $(a^n)^{-1} = a^{-n} = (a^{-1})^n$.*
(ii) $a^n a^m = a^{n+m}$.
(iii) $(a^n)^m = a^{nm}$.
(iv) $a^n b^n = (ab)^n$.
(v) *For $n \geq m \geq 0$ and $a \neq 1$ we have*

$$\sum_{k=m}^{n} a^k = \frac{a^{n+1} - a^m}{a - 1}.$$

(vi) *Suppose* $n \geq 0$. *If* $a > 1$ *then* $a^{n-1} < a^n$, *and if* $0 < a < 1$ *then* $a^{n-1} > a^n$.
(vii) *If* $a > 0$ *then* $a^n > 0$.
(viii) *If* n *is even then* $(-a)^n = a^n$, *and if* n *is odd then* $(-a)^n = -a^n$.
(ix) *If* $n > 0$ *and* $0 \leq a < b$ *then* $a^n < b^n$.
(x) *If* $n > 0$ *is odd and* $a < b$ *then* $a^n < b^n$.
(xi) $|a^n| = |a|^n$.

Proof. Almost all of the proofs are by induction. We will only write the induction steps below, since the base of inductions can be checked easily.

(i, ii, iii, iv) These are proved in Theorem 5.15.

(v) We have

$$\begin{aligned}(a-1)\Big(\sum_{k=m}^{n} a^k\Big) &= \sum_{k=m}^{n}(a^{k+1} - a^k) \\ &= a^{n+1} - a^n + a^n - a^{n-1} + \cdots + a^{m+1} - a^m = a^{n+1} - a^m.\end{aligned}$$

(vi) For $a > 1$ we have $a^{n-1} < a^n$; hence

$$a^n = aa^{n-1} < aa^n = a^{n+1}.$$

The case of $0 < a < 1$ is similar.

(vii) For $n > 0$ we multiply both sides of $a^n > 0$ by a to get $a^{n+1} > 0$. When $n = 0$ we have $a^0 = 1 > 0$. And when $n = -m < 0$ we have $a^n = (a^{-1})^m > 0$, since $a^{-1} > 0$.

(viii) Since $(-a)^n = ((-1)a)^n = (-1)^n a^n$, we only need to compute $(-1)^n$. Now if $(-1)^{2k} = 1$ then

$$(-1)^{2(k+1)} = (-1)^{2k}(-1)^2 = 1 \cdot 1 = 1.$$

For negative powers we have the same result since $(-1)^{-1} = -1$. Finally for odd powers we have $(-1)^{2k+1} = (-1)^{2k}(-1) = -1$.

(ix) First suppose $0 < a < b$. Then by induction hypothesis and (vii) we know that $0 < a^n < b^n$. Now we multiply these two inequalities to get

$$0 < a^{n+1} < b^{n+1}.$$

Since $0^n = 0$, we can allow $a = 0$ too by (vii).

(x) If $0 \leq a < b$ then the claim holds by last part. If $a < b \leq 0$ then $0 \leq -b < -a$, hence

$$-b^n = (-b)^n < (-a)^n = -a^n.$$

Thus $a^n < b^n$. Finally if $a < 0 < b$, then $b^n > 0$. Also $-a^n = (-a)^n > 0$, since $-a > 0$. Hence $a^n < 0 < b^n$.

(xi) For $n \geq 0$ we have

$$|a^{n+1}| = |a^n a| = |a^n||a| = |a|^n|a| = |a|^{n+1}.$$

When $n = -m < 0$ we have

$$|a^{-m}| = |(a^{-1})^m| = |a^{-1}|^m = (|a|^{-1})^m = |a|^{-m}.$$ ■

***Exercise* 6.1.** Show that for $a_1, \ldots, a_m \in \mathbb{R}$ we have

$$(a_1 + \cdots + a_m)^2 = \sum_{i \leq m} a_i^2 + 2 \sum_{j \leq m} \sum_{i<j} a_i a_j.$$

Theorem 6.15. *For every real number $x \geq 0$ and every $n \in \mathbb{N}$ there is a unique real number $y \geq 0$ such that $y^n = x$. We denote y by $\sqrt[n]{x}$ or $x^{\frac{1}{n}}$, and call it the* ***nth root*** *of x.*

Notation. We denote $\sqrt[2]{x}$ by $\sqrt{x}$, and we call it the **square root** of x.

Proof. The uniqueness of y is obvious, since if $0 \leq y_1 < y_2$ then $y_1^n < y_2^n$. For the existence, consider the set

$$A := \{z \geq 0 : z^n \leq x\}.$$

Note that $0 \in A$, so A is nonempty. Also if $z \geq 1 + x \geq 1$, then $z^n \geq (1 + x)^n \geq 1 + x > x$. Thus $1 + x$ is an upper bound for A. Let $y := \sup A$. We need to show that $y^n = x$.

Suppose to the contrary that $y^n < x$. Then for $0 < \epsilon < 1$ and $k \in \mathbb{N}$ we have $\epsilon^k \leq \epsilon$. Thus by using the binomial theorem we get

$$(y + \epsilon)^n = \sum_{k=0}^{n} \binom{n}{k} y^{n-k} \epsilon^k \leq y^n + \sum_{k=1}^{n} \binom{n}{k} y^{n-k} \epsilon = y^n + M\epsilon,$$

where $M := \sum_{k=1}^{n} \binom{n}{k} y^{n-k} > 0$. Now for $0 < \epsilon < \min\{\frac{x-y^n}{M}, 1\}$ we have

$$(y + \epsilon)^n \leq y^n + M\epsilon < x,$$

which contradicts the fact that y is an upper bound for A.

Next, suppose to the contrary that $y^n > x \geq 0$. Suppose $0 < \epsilon < \min\{1, y\}$. Let $z := y - \epsilon$. Then as $0 < z < y$ we have

$$y^n = (z + \epsilon)^n = \sum_{k=0}^{n} \binom{n}{k} z^{n-k} \epsilon^k < z^n + \sum_{k=1}^{n} \binom{n}{k} y^{n-k} \epsilon = z^n + M\epsilon,$$

where $M := \sum_{k=1}^{n} \binom{n}{k} y^{n-k} > 0$. Now for $0 < \epsilon < \min\{\frac{y^n - x}{M}, 1, y\}$ we have

$$(y-\epsilon)^n = z^n > y^n - M\epsilon > x.$$

But y is the supremum of A, so $y - \epsilon$ is not an upper bound for A. Hence there is $a \in A$ such that $a > y - \epsilon$. Therefore $a^n > (y-\epsilon)^n > x$, which is a contradiction. ■

Remark. When $n \in \mathbb{N}$ is odd we have $(-1)^n = -1$. Thus for $x < 0$ we have

$$(-\sqrt[n]{-x})^n = (-1)^n(\sqrt[n]{-x})^n = (-1)(-x) = x.$$

Since for $y_1 < y_2$ we have $y_1^n < y_2^n$, there is no other real number whose nth power is x. So we define

$$\sqrt[n]{x} := -\sqrt[n]{-x}.$$

Theorem 6.16. *For any two nonnegative real numbers x, y, and any $n \in \mathbb{N}$ we have*

$$\sqrt[n]{xy} = \sqrt[n]{x}\,\sqrt[n]{y}.$$

When n is odd we can allow x and/or y to be negative too.

Proof. We have

$$(\sqrt[n]{x}\,\sqrt[n]{y})^n = (\sqrt[n]{x})^n(\sqrt[n]{y})^n = xy.$$

Thus when $x, y \geq 0$ we get the desired result since $\sqrt[n]{x}\,\sqrt[n]{y} \geq 0$. When n is odd, we do not need to assume anything about the sign of x, y, since there is only one real number whose nth power is xy. ■

Theorem 6.17. *For any two nonnegative real numbers x, y, and any $n \in \mathbb{N}$ we have*

$$x < y \implies \sqrt[n]{x} < \sqrt[n]{y}.$$

When n is odd we can allow x and/or y to be negative too.

Proof. Suppose to the contrary that $\sqrt[n]{x} \geq \sqrt[n]{y}$. Then as $\sqrt[n]{x}, \sqrt[n]{y} \geq 0$ we have

$$x = (\sqrt[n]{x})^n \geq (\sqrt[n]{y})^n = y,$$

which is a contradiction. When n is odd the same argument works, except that we do not need to assume $\sqrt[n]{x}, \sqrt[n]{y} \geq 0$, so we can allow x, y to be negative too. ■

Theorem 6.18. *For any real number x we have $\sqrt{x^2} = |x|$.*

Proof. We have $|x|^2 = |x^2| = x^2$, since $x^2 \geq 0$. Hence we get the desired result because $|x| \geq 0$. ■

Theorem 6.19. *$\sqrt{2}$ is irrational.*

Proof. Suppose to the contrary that $\sqrt{2} \in \mathbb{Q}$. Thus we have $\sqrt{2} = \frac{a}{b}$ for some $a, b \in \mathbb{N}$, since $\sqrt{2} > 0$. Let q be the smallest element of the set

$$\{b \in \mathbb{N} \subset \mathbb{R} : \sqrt{2} = \frac{a}{b} \text{ for some } a \in \mathbb{N}\},$$

and let p be such that $\sqrt{2} = \frac{p}{q}$. Note that both p, q cannot be even. Because otherwise we would have $p = 2p'$ and $q = 2q'$ for some $p', q' \in \mathbb{N}$. Then we would get $\sqrt{2} = \frac{2p'}{2q'} = \frac{p'}{q'}$. However $q' < q$; which contradicts the fact that q is the smallest element of the above set. Therefore at least one of p, q is odd.

Now we know that $p^2 = 2q^2$. Suppose p is even; so $p = 2k$ for some $k \in \mathbb{N}$. Then q is odd; thus q^2 is odd too. But we have $4k^2 = p^2 = 2q^2$. Hence $q^2 = 2k^2$ is even, which is a contradiction. Next suppose p is odd. Then p^2 is odd too. But $p^2 = 2q^2$ must be even, which is again a contradiction. Hence $\sqrt{2}$ cannot be rational. ■

Rational and Real Exponents.

Suppose $p \in \mathbb{Q}$ and $x \in \mathbb{R}$ are positive. Then there are $n, k \in \mathbb{N}$ with no common factor such that $p = \frac{k}{n}$. Furthermore, $p = \frac{km}{nm}$ for every $m \in \mathbb{N}$, and these are all the representations of p as a fraction with positive denominator. Now we have $(\sqrt[n]{x})^k = \sqrt[n]{x^k}$. We also have

$$((\sqrt[nm]{x})^{km})^n = ((\sqrt[nm]{x})^{nm})^k = x^k \implies (\sqrt[nm]{x})^{km} = \sqrt[n]{x^k} = (\sqrt[n]{x})^k.$$

So if we set

$$x^p = x^{\frac{k}{n}} := (\sqrt[n]{x})^k,$$

then x^p is well defined. We also set

$$x^{-p} := \frac{1}{x^p} = (\sqrt[n]{x})^{-k}.$$

Also remember that $x^0 := 1$. Hence we have defined x^p for all $p \in \mathbb{Q}$ and all $x > 0$.

Remark. It is obvious from the definition that for all $p \in \mathbb{Q}$ we have $1^p = 1$, and $x^p > 0$ for all $x > 0$.

Proposition 6.4. *Suppose $p, q \in \mathbb{Q}$ and $p < q$. Then for $x > 1$ we have $x^p < x^q$, and for $0 < x < 1$ we have $x^p > x^q$.*

Proof. First suppose $x > 1$. Let $p = \frac{k}{n}$ and $q = \frac{l}{m}$, where $k, l \in \mathbb{Z}$, and $n, m \in \mathbb{N}$. Then we have $\frac{k}{n} < \frac{l}{m}$, so $mk < nl$. If $0 \leq p < q$ then $0 \leq mk < nl$. Thus $x^{mk} < x^{nl}$. Hence

$$x^p = \sqrt[n]{x^k} = \sqrt[nm]{x^{mk}} < \sqrt[nm]{x^{nl}} = \sqrt[m]{x^l} = x^q.$$

If $p < q \leq 0$ then $0 \leq -q < -p$. Therefore $0 < x^{-q} < x^{-p}$. Hence

$$x^p = \frac{1}{x^{-p}} < \frac{1}{x^{-q}} = x^q.$$

And if $p < 0 < q$ then $x^p < x^0 < x^q$. The case of $0 < x < 1$ is similar. ■

Now we can define x^r for all $x > 0$ and all $r \in \mathbb{R}$. We set

$$x^r := \begin{cases} \sup\{x^p : p \in \mathbb{Q}, p \leq r\} & x \geq 1, \\ \inf\{x^p : p \in \mathbb{Q}, p \leq r\} & 0 < x < 1. \end{cases}$$

First we have to check that the above supremum and infimum are finite. But this is easy since if $q \geq r$ is a rational number, then for all rational numbers $p \leq r$ we have

$$\begin{cases} x^p \leq x^q & x > 1, \\ x^p \geq x^q & 0 < x < 1. \end{cases} \qquad (*)$$

Hence the set $\{x^p : p \leq r\}$, which is obviously nonempty, has the appropriate bound to ensure the finiteness of the above supremum and infimum. The inequalities in $(*)$ also show that when r is rational, the new definition of x^r agrees with the old definition. Because in this case we have $x^r \in \{x^p : p \leq r\}$, so x^r is the required supremum, or infimum, of this set.

Remark. Note that for all $r \in \mathbb{R}$ we have

$$1^r = \sup\{1^p : p \leq r\} = \sup\{1\} = 1.$$

Also, for every $x > 0$ and $r \in \mathbb{R}$ we have $x^r > 0$. This is obvious when $x > 1$, since in this case x^r is the supremum of a set of positive numbers. When $0 < x < 1$, the second inequality in $(*)$ implies that

$$x^r = \inf\{x^p : p \leq r\} \geq x^q > 0,$$

for some rational number $q \geq r$.

Remark. We can develop the properties of powers with real exponents using the tools that we already have, but the proofs are cumbersome and lengthy. So we will not pursue this direction here.

Remark. There are other equivalent ways to define x^r. For example we can use the notion of limit of sequences, and set

$$x^r := \lim x^{p_i},$$

where p_i is a sequence of rational numbers converging to r. We have to check that the limit exists, and does not depend on the particular sequence p_i, i.e. if p_i, q_i are two sequences of rational numbers converging to r then $|x^{p_i} - x^{q_i}| \to 0$.

6.4 Complex Numbers

Our final construction is the set of complex numbers. Although real numbers can be used to develop considerable portions of mathematics, and mathematical analysis, there is still a need for a larger set of numbers. As an example, consider the polynomial equation $x^2 + 1 = 0$. This equation does not have a solution in $\mathbb{R}$, because for every $x \in \mathbb{R}$ we have $x^2 + 1 \geq 0 + 1 = 1 > 0$. However, there is a complex number that satisfies the equation $x^2 + 1 = 0$.

Definition 6.12. The set $\mathbb{C}$ of **complex numbers** is the set $\mathbb{R}^2$ equipped with the following addition and multiplication

$$\begin{aligned}(a, b) + (c, d) &:= (a + c, b + d),\\ (a, b)(c, d) &:= (ac - bd, ad + bc).\end{aligned}$$

Theorem 6.20. *$\mathbb{C}$ is a field, whose zero and identity are respectively*

$$(0, 0), \text{ and } (1, 0).$$

Also the opposite of a complex number $z = (a, b)$ is

$$-z := (-a, -b),$$

and when z is nonzero its inverse is

$$z^{-1} := \Big(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\Big).$$

Proof. Let $z = (a, b)$, $w = (c, d)$, and $u = (e, f)$ be complex numbers. It is easy to check that addition is associative and commutative:

$$\begin{aligned}z + (w + u) &= \big(a + (c + e), b + (d + f)\big)\\ &= \big((a + c) + e, (b + d) + f\big) = (z + w) + u,\\ z + w &= (a + c, b + d) = (c + a, d + b) = w + z.\end{aligned}$$

We can also easily check that

$$\begin{aligned}(a, b) + (0, 0) &= (a + 0, b + 0) = (a, b),\\ z + (-z) &= (a + (-a), b + (-b)) = (0, 0).\end{aligned}$$

It is obvious that $(1, 0) \neq (0, 0)$. In addition, we have

$$(a, b)(1, 0) = (a1 - b0, a0 + b1) = (a, b).$$

Now let us check that multiplication is associative, commutative, and distributive over addition. We have

$$\begin{aligned} zw &= (ac-bd, ad+bc) = (ca-db, cb+da) = wz, \\ z(wu) &= (a,b)(ce-df, cf+de) \\ &= \big(ace-adf-bcf-bde, acf+ade+bce-bdf\big) \\ &= \big((ac-bd)e-(ad+bc)f, (ac-bd)f+(ad+bc)e\big) \\ &= (ac-bd, ad+bc)(e,f) = (zw)u, \\ z(w+u) &= (a,b)(c+e, d+f) \\ &= (ac+ae-bd-bf, ad+af+bc+be) \\ &= (ac-bd, ad+bc)+(ae-bf, af+be) = zw+zu. \end{aligned}$$

Finally, suppose $z \neq (0,0)$. Then $a \neq 0$ or $b \neq 0$. Hence we must have $a^2+b^2 > 0$. To simplify the notation let $r := a^2+b^2$. Then we have

$$zz^{-1} = (a,b)\big(\frac{a}{r}, \frac{-b}{r}\big) = \big(\frac{a^2}{r} - \frac{-b^2}{r}, \frac{-ab}{r} + \frac{ba}{r}\big) = \big(\frac{a^2+b^2}{r}, 0\big) = (1,0),$$

as desired. ■

Remark. The map $a \mapsto (a,0)$ from $\mathbb{R}$ into $\mathbb{C}$ is a one-to-one map that preserves addition and multiplication, i.e.

$$(a,0)+(b,0) = (a+b,0), \qquad (a,0)(b,0) = (ab,0).$$

Thus $\mathbb{C}$ contains a copy of the field $\mathbb{R}$. We will abuse the notation and denote the element $(a,0)$ by a. We also define $i := (0,1)$. Then any complex number $z = (a,b)$ can be written as

$$z = (a,b) = (a,0)+(0,b) = (a,0)+(0,1)(b,0) = a+ib.$$

Note that we have

$$i^2 = (0,1)^2 = (-1,0) = -1,$$

i.e. i is a square root of -1.

Definition 6.13. Let $z = (a,b) = a+ib$ be a complex number. The real numbers a, b are called the **real part** and the **imaginary part** of z, respectively, and we will denote them by

$$a = \operatorname{Re} z, \qquad b = \operatorname{Im} z.$$

The **conjugate** of z is the complex number

$$\bar{z} := (a,-b) = a-ib.$$

The **modulus** or the **absolute value** of z is the nonnegative real number

$$|z| := \sqrt{a^2+b^2}.$$

Remark. Note that for $a \in \mathbb{R}$ we have $|(a,0)| = \sqrt{a^2} = |a|$. Thus the absolute value of complex numbers is compatible with the absolute value of real numbers.

Theorem 6.21. *Let $z, w \in \mathbb{C}$. Then we have*

(i) $|z| \geq 0$, *and* $|z| = 0 \iff z = 0$.

(ii) **Triangle Inequality:** $|z + w| \leq |z| + |w|$.

(iii) $\big||z| - |w|\big| \leq |z - w|$.

Proof. Let $z = a + ib$ and $w = c + id$, where $a, b, c, d \in \mathbb{R}$.

(i) Obviously we have $|z| \geq 0$, and $|0| = 0$. Now if $|z| = 0$ then $a^2 + b^2 = 0$. But $a^2, b^2 \geq 0$. Hence $a^2, b^2 = 0$; so $a, b = 0$. Thus $z = 0$.

(ii) First note that $a^2d^2 - 2acbd + b^2c^2 = (ad - bc)^2 \geq 0$. So we have $2acbd \leq a^2d^2 + b^2c^2$. Therefore

$$\begin{aligned}(ac + bd)^2 &= a^2c^2 + 2acbd + b^2d^2 \\ &\leq a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 = (a^2 + b^2)(c^2 + d^2) = |z|^2|w|^2.\end{aligned}$$

Thus we get

$$ac + bd \leq |ac + bd| \leq |z||w|,$$

since the square root is an increasing function. Hence we obtain

$$\begin{aligned}|z + w|^2 &= (a + c)^2 + (b + d)^2 \\ &= a^2 + 2ac + c^2 + b^2 + 2bd + d^2 \\ &= |z|^2 + |w|^2 + 2(ac + bd) \\ &\leq |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2.\end{aligned}$$

Thus $|z + w| \leq |z| + |w|$, as desired.

(iii) We have $|z| = |z - w + w| \leq |z - w| + |w|$. Hence we get

$$|z| - |w| \leq |z - w|.$$

Similarly we have $|w| \leq |w - z| + |z| = |z - w| + |z|$, since we can easily see that the modulus of a complex number is the same as the modulus of its opposite. Thus we get

$$-|z - w| \leq |z| - |w|.$$

Therefore we must have $\big||z| - |w|\big| \leq |z - w|$, as desired. ■

Remark. Note that $\mathbb{C}$ does not have a natural order. In fact, there is no order on $\mathbb{C}$ that makes it into an ordered field. The reason is that in any ordered field the square of any nonzero element is positive. In particular $1 = 1^2 > 0$. Thus $-1 < 0$. But in $\mathbb{C}$ we have $i^2 = -1$. Hence if $\mathbb{C}$ was an ordered field, then -1 must have been simultaneously positive and negative, which is impossible.

Remark. We can define the integer powers of complex numbers, similarly to the case of real numbers. Then all the basic properties of powers expressed in Theorem 6.14 also hold for powers of complex numbers, except obviously those properties that are related to the order structure.

Theorem 6.22. *For every $z, w \in \mathbb{C}$ and $n \in \mathbb{Z}$ we have*

(i) $\overline{z+w} = \bar{z} + \bar{w}$.
(ii) $\overline{zw} = \bar{z}\bar{w}$.
(iii) $\bar{\bar{z}} = z$.
(iv) $z\bar{z} = |z|^2$, *hence* $z^{-1} = |z|^{-2}\bar{z}$.
(v) $|\bar{z}| = |z|$.
(vi) $|zw| = |z||w|$.
(vii) $|\operatorname{Re} z| \le |z|$, *and* $|\operatorname{Im} z| \le |z|$.
(viii) $z = \bar{z}$ *if and only if* $z \in \mathbb{R}$.
(ix) $z + \bar{z} = 2\operatorname{Re} z$, *and* $z - \bar{z} = 2i \operatorname{Im} z$.
(x) $|z^n| = |z|^n$ *(when $z = 0$ we assume $n > 0$).*
(xi) $\overline{z^n} = \bar{z}^n$ *(when $z = 0$ we assume $n > 0$).*

Proof. Let $z = a + ib$ and $w = c + id$, where $a, b, c, d \in \mathbb{R}$.

(i) We have

$$\begin{aligned}\overline{z+w} &= \overline{(a+c)+i(b+d)} \\ &= (a+c) - i(b+d) = a - ib + c - id = \bar{z} + \bar{w}.\end{aligned}$$

(ii) We have

$$\begin{aligned}\overline{zw} &= \overline{(ac-bd)+i(ad+bc)} = (ac-bd) - i(ad+bc) \\ &= (ac - (-b)(-d)) + i(a(-d) + (-b)c) = (a-ib)(c-id) = \bar{z}\bar{w}.\end{aligned}$$

(iii) $\bar{\bar{z}} = \overline{a - ib} = a - (-ib) = a + ib = z$.

(iv) We have

$$\begin{aligned}z\bar{z} &= (a+ib)(a-ib) \\ &= (a^2 - b(-b)) + i(a(-b) + ba) = a^2 + b^2 = |z|^2.\end{aligned}$$

Thus we have $(|z|^{-2}\bar{z})z = |z|^{-2}|z|^2 = 1$. Hence $|z|^{-2}\bar{z} = z^{-1}$, since the inverse is unique.

(v) $|\bar{z}| = \sqrt{a^2 + (-b)^2} = \sqrt{a^2 + b^2} = |z|$.

(vi) We have

$$\begin{aligned}|zw| &= \sqrt{(ac-bd)^2 + (ad+bc)^2} \\ &= \sqrt{a^2c^2 - 2acbd + b^2d^2 + a^2d^2 + 2adbc + b^2c^2} \\ &= \sqrt{(a^2+b^2)(c^2+d^2)} = \sqrt{a^2+b^2}\sqrt{c^2+d^2} = |z||w|.\end{aligned}$$

(vii) $|\operatorname{Re} z| = |a| = \sqrt{a^2} \leq \sqrt{a^2+b^2} \leq |z|$. Note that we have used the monotonicity of the square root function over nonnegative real numbers. The other inequality can be proved similarly.

(viii) $z = \bar{z} \iff b = -b \iff b = 0 \iff z = a \in \mathbb{R}$.

(ix) $z + \bar{z} = a + ib + a - ib = 2a = 2\operatorname{Re} z$. The other one is similar.

(x) The proof is by induction on n, when $n > 0$. The claim holds obviously for $n = 1$, so suppose it also holds for n. Then for $n+1$ we have

$$|z^{n+1}| = |z^n z| = |z^n||z| = |z|^n|z| = |z|^{n+1}.$$

Now suppose $z \neq 0$. When $n = 0$ both sides of the equation are one. For $n = -1$ we have

$$|z^{-1}| = ||z|^{-2}\bar{z}| = ||z|^{-2}||\bar{z}| = |z|^{-2}|z| = |z|^{-1}.$$

Note that $|z|^{-2}$ is a positive real number, therefore its modulus equals its absolute value as a real number, which is itself. Finally for $n = -m < 0$ we have

$$|z^n| = |(z^{-1})^m| = |z^{-1}|^m = (|z|^{-1})^m = |z|^n.$$

(xi) The proof is by induction on n, when $n > 0$. The claim holds obviously for $n = 1$, so suppose it also holds for n. Then for $n+1$ we have

$$\overline{z^{n+1}} = \overline{z^n z} = \overline{z^n}\bar{z} = \bar{z}^n\bar{z} = \bar{z}^{n+1}.$$

Now suppose $z \neq 0$. When $n = 0$ both sides of the equation are one. For $n = -1$ we have

$$\overline{z^{-1}} = \overline{|z|^{-2}\bar{z}} = \overline{|z|^{-2}}\,\overline{\bar{z}} = |z|^{-2}\overline{\bar{z}} = |\bar{z}|^{-2}\overline{\bar{z}} = (\bar{z})^{-1}.$$

Note that $|z|^{-2}$ is a real number, therefore its conjugate is itself. Finally for $n = -m < 0$ we have

$$\overline{z^n} = \overline{(z^{-1})^m} = \big(\overline{(z^{-1})}\big)^m = ((\bar{z})^{-1})^m = (\bar{z})^{-m} = \bar{z}^n.$$ ■

Remark. Suppose $z \in \mathbb{C}$ is nonzero. Then $r := |z| > 0$. Now $\frac{z}{r}$ has modulus one, so it belongs to the unit circle in $\mathbb{C}$. It can be shown that there is a unique $\theta \in [0, 2\pi)$ such that $\frac{z}{r} = e^{i\theta} = \cos\theta + i\sin\theta$. Therefore

$$z = re^{i\theta} = r(\cos\theta + i\sin\theta).$$

This is called the **polar representation** of z. The number θ is called the **argument** of z, and is denoted by $\arg z$. In fact θ is the signed angle between the segment connecting z and 0, and the half line of nonnegative real numbers.

Remark. Suppose $z = re^{i\theta}$ and $w = se^{i\phi}$. Then it can be shown that

$$zw = rse^{i(\theta+\phi)}.$$

The interpretation of this formula is that when you multiply a complex number w by a complex number z, you scale the modulus of w by the modulus of z, and you rotate w around the origin by the angle $\arg z$.

Chapter 7

The Axiom of Choice and Countable Sets

7.1 The Axiom of Choice

The axiom of choice informally says that if we have a family of nonempty sets, then we can choose an element from each set in the family. It does not provide us a concrete method for choosing the elements; it merely says that such a choice can be made. Let us first state the axiom formally. To simplify the notation, let

$$\mathrm{Fun}(f, A, B)$$

denote the formula which says that f is a function from A to B.

Axiom of Choice.

$$\vdash \forall X \big[\big(\forall x \in X \; x \neq \emptyset \big) \to \exists f \big(\mathrm{Fun}(f, X, \bigcup X) \wedge \forall x \in X \; f(x) \in x \big) \big].$$

In other words, the axiom of choice says that if X is a set whose members are nonempty sets, then there is a function $f : X \to \bigcup X$ that maps each member of X to an element of its own, i.e. $f(x) \in x$ for every $x \in X$. Hence the function f chooses an element of every member of X. A function that has this property is called a **choice function**. Thus the axiom of choice says that every set whose elements are nonempty sets has a choice function.

Note that we do not need the axiom of choice when X is finite, i.e. we do not need the axiom of choice to choose from finitely many nonempty sets. This fact can be proved by induction, using the other axioms.

Theorem 7.1. *Let X be a finite set whose elements are nonempty. Then X has a choice function.*

Proof. When X is empty then $\emptyset$ is vacuously a choice function for X. So suppose X is nonempty. Let $n = |X|$. The proof is by induction on n. If $n = 1$ then X is a singleton. Hence $X = \{a\}$ for some a. In addition, by our assumption a is nonempty, i.e. $\exists b(b \in a)$. Therefore there exist a, b such that $(a, b) \in X \times \bigcup X$. Let $f := \{(a, b)\}$. Then f is a function from X to $\bigcup X$. In addition, for every $x \in X$ we have $x = a$, since X is a singleton. Thus $f(x) = f(a) = b \in a = x$. Hence f is a choice function for X.

Now suppose the claim holds for some n. Let X be a finite set whose elements are nonempty, such that $|X| = n + 1$. Then X is nonempty, since $|X| > 0$. Let $a \in X$, and $Y := X - \{a\}$. Then Y is also a finite set whose elements are nonempty, and $|Y| = n$. Then by induction hypothesis Y has a choice function g. Also, as we showed above, the singleton set $\{a\}$ has a choice function h. Note that the codomains of g, h are $\bigcup Y, \bigcup\{a\}$ respectively; and they are both contained in $\bigcup X$. So we can consider g, h as functions into $\bigcup X$. Let $f := g \cup h$, i.e.

$$f(x) := \begin{cases} g(x) & x \in Y, \\ h(a) & x = a. \end{cases}$$

Then by Theorem 3.20, f is a function from X to $\bigcup X$, since $Y \cap \{a\} = \emptyset$. Now for every $x \in X$ if $x \in Y$ we have $f(x) = g(x) \in x$, and if $x = a$ we have $f(x) = h(a) \in a = x$. Hence f is a choice function for X, as desired. ■

Next let us formulate an equivalent version of the axiom of choice.

Theorem 7.2. *The axiom of choice is equivalent to the following statement:*

Let X be a set whose elements are nonempty and disjoint, then there exists a set C such that for every $x \in X$, $x \cap C$ is a singleton.

Proof. ■

Remark. As a consequence of the previous two theorems, without the axiom of choice we can show that if X is a finite set whose elements are nonempty and disjoint, then there exists a set C such that for every $x \in X$, $x \cap C$ is a singleton.

7.2 Countable Sets

Definition 7.1. A set A is **countably infinite** or **denumerable** if there exists a bijective function from $\mathbb{N}$ to A. A set is **countable** if it is finite or countably infinite. A set that is not countable is **uncountable**.

Theorem 7.3. *We have*

(i) *A set A is countable if and only if there exists a surjective function from $\mathbb{N}$ to A.*

(ii) *A set A is countable if and only if there exists an injective function from A to $\mathbb{N}$.*
(iii) *A subset of a countable set is countable.*
(iv) *The Cartesian product of finitely many countable sets is countable.*
(v) *The union of countably many countable sets is countable.*

Proof. **(i)**
(ii)
(iii)
(iv)
(v) ■

Theorem 7.4. $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ *are countably infinite.*

Proof. $\mathbb{N}$ is countably infinite, since the identity map is a bijection from $\mathbb{N}$ onto $\mathbb{N}$. For $\mathbb{Z}$ we have the bijection

$$f(n) := \begin{cases} 2n & n > 0 \\ -2n+1 & n \leq 0. \end{cases}$$

Finally, $\mathbb{Q}$ is infinite and can be written as the union of countably many countable sets as follows

$$\mathbb{Q} = \bigcup_{q \in \mathbb{N}} \{\frac{p}{q} : p \in \mathbb{Z}\}.$$ ■

Theorem 7.5. *The set of all sequences of $0, 1$, i.e. the set of all functions from $\mathbb{N}$ to $\{0, 1\}$, is uncountable.*

Proof. Let f be an injective map from $\mathbb{N}$ into the set of all sequences of $0, 1$. We show that f cannot be surjective. We use the diagonal method due to Cantor. Consider

$$\begin{aligned} f(1) &= a_{11}a_{12}a_{13}\cdots \\ f(2) &= a_{21}a_{22}a_{23}\cdots \\ &\vdots \end{aligned}$$

where $a_{ij} \in \{0, 1\}$. Now we define the sequence $b = b_1 b_2 b_3 \cdots$ as follows

$$b_i := \begin{cases} 1 & \text{if } a_{ii} = 0, \\ 0 & \text{if } a_{ii} = 1. \end{cases}$$

Then $b \neq f(n)$ for each $n \in \mathbb{N}$, since $b_n \neq a_{nn}$. Hence b is not in the image of f, and f is not onto. ■

Theorem 7.6. $\mathbb{R}$ *is uncountable.*

Proof. The proof is similar to the last theorem. We only need to use the decimal expansion of real numbers instead of the sequences. ■

Theorem 7.7. *Suppose* $a, b \in \mathbb{R}$*, and* $a < b$*. Then the intervals* $(a, b), [a, b]$ *are uncountable.*

Proof. We will show that there are bijections from the $\mathbb{R}$ onto (a, b). Then it follows that the (a, b) cannot be countable, since otherwise $\mathbb{R}$ would be countable too. It also follows that $[a, b]$ is uncountable, since $(a, b) \subset [a, b]$.

Now the function

$$x \mapsto \frac{x}{1 + |x|}$$

is a bijection from $\mathbb{R}$ onto $(-1, 1)$ (why?), and $x \mapsto a + \frac{x+1}{2}(b - a)$ is a bijection from $(-1, 1)$ onto (a, b). So their composition is the required bijection from $\mathbb{R}$ onto (a, b). ■

Appendix A

Factorization

A.1 Euclidean Domains

Definition A.1. Let R be an integral domain, and $a, b \in R$. If there exists $r \in R$ such that $b = ra$, then we say a **divides** b, or a is a **divisor** of b, or b is a **multiple** of a; and we write $a \mid b$. If a is not a divisor of b we write $a \nmid b$.

An element $u \in R$ is called a **unit** if $u \mid 1$. Two elements a, b are **associates** if $a \mid b$ and $b \mid a$.

***Example* A.1.** In the integral domain domain $\mathbb{Z}$ the units are ± 1. Also, two integers a, b are associate if and only if $a = \pm b$. (These facts have been shown in Proposition 5.7.)

Proposition A.1. *Suppose R is an integral domain, and $a, b, c \in R$. Then*

(i) *The units of R are exactly the invertible elements of R.*
(ii) *For every $a \in R$ we have $1 \mid a$, $a \mid a$, and $a \mid 0$.*
(iii) *a, b are associates if and only if there is a unit $u \in R$ such that $a = ub$.*
(iv) *If $a \mid b$ and $b \mid c$ then $a \mid c$.*
(v) *$a \mid b$ implies $a \mid bc$ and $ac \mid bc$.*
(vi) *If $a \mid b$ and $a \mid c$ then $a \mid bx + cy$ for every $x, y \in R$, in particular $a \mid b + c$.*

Remark. Note that if $0 \mid a$ then for some r we have $a = r0 = 0$. Hence 0 is the only element that has 0 as a divisor.

Proof. **(i)** $a \mid 1$ if and only if there is $r \in R$ such that $ra = 1$, i.e. a is invertible with $a^{-1} = r$.

(iii) If a, b are associates, then $a \mid b$ and $b \mid a$. Thus there are $r, s \in R$ such that $ra = b$ and $sb = a$. If $a = 0$ then $b = 0$, hence $a = 1b$. So suppose $a \neq 0$. Now we have

$$1a = a = sb = sra.$$

Thus $sr = 1$, since R is an integral domain. Therefore s is a unit, and $a = sb$. Conversely, if $a = ub$ for some unit u, then $b = u^{-1}a$. Hence $a \mid b$ and $b \mid a$.

(ii, iv, v, vi) The proofs of these parts are the same as of analogous parts of Proposition 5.7. ■

***Exercise* A.1.** Suppose R is an integral domain and $a, b, c \in R$. Show that if a, b are associates, and b, c are associates, then a, c are associates too.

Definition A.2. A **Euclidean domain** is an integral domain R on which there exists a **degree function**

$$d : R - \{0\} \to \{n \in \mathbb{Z} : n \geq 0\},$$

such that

(i) For nonzero $a, b \in R$ if $a \mid b$ then $d(a) \leq d(b)$.

(ii) **(Division Algorithm)** For every $a \in R$ and every nonzero $b \in R$, there are $q, r \in R$ such that

$$a = bq + r, \qquad \text{where either } r = 0, \text{ or } d(r) < d(b).$$

Here q is called the **quotient** and r is called the **remainder**.

***Example* A.2.** As we have seen in Section 5.3, $\mathbb{Z}$ is a Euclidean domain with the degree function $d(n) = |n|$. In this particular case, the quotient and the remainder in the division algorithm are unique provided that we require the remainder to be nonnegative.

Remark. Note that q, r are not necessarily unique in general. Even in $\mathbb{Z}$, if we do not restrict the remainder to be nonnegative, the quotient and remainder are not unique. For example we have $7 = 3 \times 2 + 1 = 3 \times 3 - 2$.

Proposition A.2. *Suppose R is a Euclidean domain, and $a, b \in R$. If b is nonzero and it is not a unit, then for all nonzero a we have*

$$d(a) < d(ab).$$

Proof. We know that $d(b) \leq d(ab)$, since $b \mid ab$. Suppose to the contrary that $d(ab) = d(a)$. Then we have $a = abq + r$ where $d(r) < d(ab) = d(a)$ or $r = 0$. If $r = 0$ we have $a(1 - bq) = 0$ which is a contradiction, since b is not a unit and a is nonzero. Hence we have $r \neq 0$ and so $d(r) < d(a)$. Now

$$a(1 - bq) = a - abq = r.$$

Thus $a \mid r$ and we must have $d(a) \leq d(r)$. This contradiction proves the result. ■

A.2 Polynomials

Definition A.3. Let R be a commutative ring. The ring of **polynomials** with coefficients in R is the set of all sequences in R that terminate eventually, i.e.

$$R[x] := \{f : \mathbb{N} \cup \{0\} \to R \ : \ \text{there is } N \geq 0 \text{ such that } f_n = 0 \text{ for } n \geq N\}.$$

The elements f_n are called the **coefficients** of f. The **zero polynomial** is the polynomial whose coefficients are all zero. For a nonzero polynomial f, the largest nonnegative integer n for which $f_n \neq 0$ is called the **degree** of f and is denoted by

$$\deg f.$$

We also define the degree of the zero polynomial to be $-\infty$, with the understanding that for all $n \in \mathbb{Z}$ we have

$$-\infty < n, \qquad -\infty + n = -\infty.$$

The addition and multiplication of two polynomials f, g are defined as follows

$$(f+g)_n := f_n + g_n, \qquad (fg)_n := \sum_{k \leq n} f_k g_{n-k}.$$

Remark. Note that as polynomials are sequences of elements of R, it suffices to define them by specifying their nth terms for every n. Also, when we want to show that two polynomials are equal, it is enough to check the equality of their nth terms for each n.

Remark. It is easy to see that $R[x]$ is closed under the addition and multiplication defined above. Because $(f+g)_n$ and $(fg)_n$ are zero for all large values of n. In fact for nonnegative $n > \max\{\deg f, \deg g\}$ we have

$$(f+g)_n = f_n + g_n = 0 + 0 = 0,$$

and for nonnegative $n > \deg f + \deg g$ we have

$$\begin{aligned}(fg)_n &= \sum_{k=0}^{n} f_k g_{n-k} = \sum_{k \leq \deg f} f_k g_{n-k} + \sum_{k > \deg f} f_k g_{n-k} \\ &= \sum_{k \leq \deg f} f_k 0 + \sum_{k > \deg f} 0 g_{n-k} = 0.\end{aligned}$$

Note that for $k \leq \deg f$ we have $n - k > \deg g$, hence $g_{n-k} = 0$. Also, note that as a result we have

$$\begin{aligned}&\deg(f+g) \leq \max\{\deg f, \deg g\}, \\ &\deg(fg) \leq \deg f + \deg g.\end{aligned}$$

Theorem A.1. *Let R be a commutative ring. Then the ring of polynomials $R[x]$ is also a commutative ring.*

Proof. Let $f, g, h \in R[x]$. It is easy to check that addition of polynomials is commutative and associative, and the zero polynomial

$$0 := (0, 0, 0, \ldots, 0, \ldots)$$

is an additive identity. We have

$$\begin{aligned}(f+g)_n &= f_n + g_n = g_n + f_n = (g+f)_n,\\ (f+(g+h))_n &= f_n + (g+h)_n = f_n + (g_n + h_n)\\ &= (f_n + g_n) + h_n = (f+g)_n + h_n = ((f+g)+h)_n,\\ (f+0)_n &= f_n + 0_n = f_n + 0 = f_n.\end{aligned}$$

Also, for any f, the polynomial defined by $(-f)_n := -f_n$ is its opposite, since

$$(f + (-f))_n = f_n + (-f)_n = f_n + (-f_n) = 0 = 0_n.$$

Note that $-f$ is a polynomial as $(-f)_n = 0$ for $n > \deg f$.

The sequence

$$1 := (1, 0, 0, \ldots, 0, \ldots)$$

is the multiplicative identity of $R[x]$, since

$$(1f)_n = 1f_n + 0f_{n-1} + \cdots + 0f_1 + 0f_0 = f_n.$$

The commutativity of the multiplication is also easy to show

$$\begin{aligned}(fg)_n &= \sum_{k=0}^{n} f_k g_{n-k} = \sum_{k=0}^{n} g_{n-k} f_k\\ &= \sum_{l=n}^{0} g_l f_{n-l} = \sum_{l=0}^{n} g_l f_{n-l} = (gf)_n. \qquad (l := n-k)\end{aligned}$$

(Note that we have used both the commutativity of the multiplication of R, and the generalized commutativity of the addition of R.) To see that multiplication is distributive over addition, we have

$$\begin{aligned}(f(g+h))_n &= \sum_{k=0}^{n} f_k (g+h)_{n-k} = \sum_{k=0}^{n} f_k (g_{n-k} + h_{n-k})\\ &= \sum_{k=0}^{n} (f_k g_{n-k} + f_k h_{n-k}) = \sum_{k=0}^{n} f_k g_{n-k} + \sum_{k=0}^{n} f_k h_{n-k} = (fg)_n + (fh)_n.\end{aligned}$$

It remains to show that multiplication is associative. Let

$$\Delta_{l,k} := \begin{cases} 1 & l \le k, \\ 0 & l > k, \end{cases}$$

where l, k are nonnegative integers and $0, 1 \in R$. Then we have

$$\begin{aligned} \big((fg)h\big)_n &= \sum_{k\le n}(fg)_k h_{n-k} = \sum_{k\le n}\Big[\sum_{l\le k} f_l g_{k-l}\Big] h_{n-k} \\ &= \sum_{k\le n}\sum_{l\le n}\Delta_{l,k} f_l g_{k-l} h_{n-k} = \sum_{l\le n}\sum_{k\le n}\Delta_{l,k} f_l g_{k-l} h_{n-k} \\ &= \sum_{l\le n} f_l \sum_{l\le k\le n} g_{k-l} h_{n-k} = \sum_{l\le n} f_l \sum_{0\le j\le n-l} g_j h_{n-l-j} \qquad (j := k-l) \\ &= \sum_{l\le n} f_l (gh)_{n-l} = \big(f(gh)\big)_n, \end{aligned}$$

as desired. ∎

Notation. We use the abbreviations

$$\begin{aligned} r &:= (r, 0, 0, \ldots, 0, \ldots), \\ x &:= (0, 1, 0, \ldots, 0, \ldots), \end{aligned}$$

where $r \in R$. Then we have

$$rx^n = x^n r = (0, 0, \ldots, \overset{n\text{-th}}{\underset{\downarrow}{r}}, \ldots, 0, \ldots).$$

Thus any polynomial can be written as

$$f = f_0 + f_1 x + \cdots + f_m x^m,$$

where $m = \deg f$ for nonzero f. Note that the coefficients of f in this representation are exactly the coefficients of f, so they are uniquely determined by f. We sometimes write $f(x)$ instead of f.

The zero polynomial, and polynomials of degree zero, are called **constant polynomials**. So constant polynomials are polynomials of degree less than one. Also, polynomials of degree one, two, and three are respectively called *linear*, *quadratic*, and *cubic* polynomials.

Remark. By identifying $r \in R$ with the constant polynomial $r \in R[x]$, we can consider R as a subring of $R[x]$.

Remark. Every polynomial defines a function on R by evaluation. That is for

$$f(x) = f_0 + f_1 x + \cdots + f_m x^m,$$

we have $f(r) := f_0 + f_1 r + \cdots + f_m r^m$, where $r \in R$. Note that $f(r)$ is uniquely determined by f and r, since the coefficients of f are uniquely determined by f. Also note that in the above, m need not be the $\deg f$. Because for $i > \deg f$ we have $f_i = 0$, hence the terms with $i > \deg f$ do not change the value of $f(r)$.

Theorem A.2. *For any two polynomials $f, g \in R[x]$ and all $r \in R$ we have*

$$(f+g)(r) = f(r) + g(r), \qquad (fg)(r) = f(r)g(r).$$

Remark. Note that the multiplication of polynomials is defined in a way that makes the above theorem valid, and this is one of the reasons behind its definition. The significance of this theorem is that the addition and multiplication of polynomials convert to the addition and multiplication of R via the map $f \mapsto f(r)$.

Proof. Let $m = \deg f$, and $n = \deg g$. Then $f_i = 0$ for $i > m$, and $g_j = 0$ for $j > n$. Let $l = \max\{m, n\}$, then $\deg(f+g) \leq l$. Now we have

$$\begin{aligned} f(r) + g(r) &= \sum_{i \leq m} f_i r^i + \sum_{i \leq n} g_i r^i = \sum_{i \leq l} f_i r^i + \sum_{i \leq l} g_i r^i \\ &= \sum_{i \leq l} (f_i r^i + g_i r^i) = \sum_{i \leq l} (f_i + g_i) r^i = (f+g)(r). \end{aligned}$$

Next, remember that $\deg(fg) \leq m+n$. For $0 \leq k \leq m+n$ let $a = \max\{0, k-m\}$, and $b = \min\{n, k\}$. Then by the generalized distributivity and Theorem 5.39 we have

$$\begin{aligned} f(r)g(r) &= \Big(\sum_{i \leq m} f_i r^i\Big)\Big(\sum_{j \leq n} g_j r^j\Big) = \sum_{i \leq m} \sum_{j \leq n} (f_i r^i)(g_j r^j) \\ &= \sum_{i \leq m} \sum_{j \leq n} (f_i g_j)(r^i r^j) = \sum_{i \leq m} \sum_{j \leq n} (f_i g_j) r^{i+j} \\ &= \sum_{k \leq m+n} \sum_{i+j=k} (f_i g_j) r^k = \sum_{k \leq m+n} \Big(\sum_{a \leq i \leq b} f_i g_{k-i}\Big) r^k \\ &= \sum_{k \leq m+n} \Big(\sum_{i \leq k} f_i g_{k-i}\Big) r^k = \sum_{k \leq m+n} (fg)_k r^k = (fg)(r). \end{aligned}$$

In the last line of the above formula we used the fact that $\sum_{a \leq i \leq b} f_i g_{k-i} = \sum_{i \leq k} f_i g_{k-i}$. The reason is that for $i > b \geq n$ we have $f_i = 0$, and for $i < a \leq k-m$ we have $g_{k-i} = 0$ since $k - i > m$.

Finally, to prove the last statement of the theorem, note that we have

$$f(r)g(r) = (fg)(r) = (gf)(r) = g(r)f(r),$$

since $R[x]$ is a commutative ring. ■

Remark. As a consequence of the above theorem, we can easily show by induction that if $p_1, \ldots, p_k \in R[x]$ then we have

$$\begin{aligned}(p_1 + \cdots + p_k)(r) &= p_1(r) + \cdots + p_k(r),\\ (p_1 p_2 \cdots p_k)(r) &= p_1(r)p_2(r)\cdots p_k(r).\end{aligned}$$

Definition A.4. Let f be a polynomial with coefficients in a ring R, and suppose $r \in R$. Then when $f(r) = 0$ we say r is a **root** of f.

Theorem A.3. *Suppose R is an integral domain. Then $R[x]$ is an integral domain too. In addition, for any two polynomials f, g we have*

$$\deg(fg) = \deg f + \deg g.$$

Proof. We have already shown that $R[x]$ is a commutative ring. If f or g, for example f, is zero, then $fg = 0$. Hence

$$\deg fg = -\infty = -\infty + \deg g = \deg f + \deg g.$$

Now suppose f, g are nonzero polynomials. Let $\deg f = n$ and $\deg g = m$. Then f_m, g_n are nonzero elements of R. Hence $(fg)_{m+n} = f_m g_n \neq 0$. Thus in particular fg is nonzero. Therefore $R[x]$ is an integral domain. Furthermore we have $\deg(fg) \geq m + n$. On the other hand we know that in general $\deg(fg) \leq m + n$, so $\deg(fg)$ is exactly $m + n$. ■

Theorem A.4. *Suppose F is a field, and $f, g \in F[x]$ with $g \neq 0$. Then there are unique $q, r \in F[x]$ such that*

$$f = gq + r, \qquad \textit{and} \quad \deg r < \deg g.$$

Proof. If there is h such that $f = gh$, then we put $q = h$ and $r = 0$. Now suppose no such h exists. Then all the polynomials in

$$\{f - gp : p \in F[x]\}$$

are nonzero. Let q be an element of this set for which $f - gq$ has the least degree. This is possible due to the well-ordering of nonnegative integers. Then set

$$r := f - gq.$$

We must show that $\deg r < \deg g$. Suppose to the contrary that $\deg r \geq \deg g$. Let

$$r(x) = r_n x^n + \cdots + r_0, \qquad g(x) = g_m x^m + \cdots + g_0.$$

Note that $r_n, g_m \neq 0$. Set $s(x) := r(x) - \frac{r_n}{g_m} x^{n-m} g(x)$. If $s = 0$ then we have

$$f(x) = g(x)\big(q(x) + \frac{r_n}{g_m} x^{n-m}\big),$$

which is in contradiction with our assumption. Thus $s \neq 0$ and we have $\deg s < \deg r$, since we have eliminated the x^n term. But this implies

$$f(x) = g(x)\big(q(x) + \frac{r_n}{g_m} x^{n-m}\big) + s(x),$$

which is in contradiction with the choice of q. Hence we have $\deg r < \deg g$ as desired.

For the uniqueness, suppose we have

$$gq_1 + r_1 = f = gq_2 + r_2.$$

Then $g(q_1 - q_2) = r_2 - r_1$. Since $g \neq 0$, we have $r_2 - r_1 = 0$ if and only if $q_1 - q_2 = 0$. Now if $r_1 \neq r_2$ and $q_1 \neq q_2$ then we get

$$\deg g + \deg(q_1 - q_2) = \deg(r_2 - r_1),$$

which is in contradiction with the fact that

$$\deg(r_2 - r_1) \leq \max\{\deg r_2, \deg r_1\} < \deg g. \qquad \blacksquare$$

Theorem A.5. *Suppose F is a field, $a \in F$, and $f \in F[x]$. Then*

$$f(x) = (x - a)g(x) + f(a).$$

Thus $f(a) = 0$ if and only if there is $g \in F[x]$ such that

$$f(x) = (x - a)g(x).$$

As a result, the number of distinct roots of a nonzero polynomial f is at most $\deg f$.

Proof. We divide f by $x - a$ to get $f(x) = (x - a)g(x) + r(x)$. But $\deg r < \deg(x - a) = 1$, so r is a constant. By evaluating the above equality at a we get $r = f(a)$. Thus

$$f(x) = (x - a)g(x) + f(a).$$

Now the second statement follows easily.

The last statement can be proved by induction on $\deg f$. Nonzero polynomials of degree zero are constant polynomials which have no root. Suppose the claim holds for all polynomials with degree less than $\deg f$. If f has no root, then there is nothing to prove. So let a be a root of f. Then $f = (x-a)g$. If $g = 0$ then $f = 0$, which is contrary to our assumption. So $g \neq 0$, and we have $\deg g = \deg f - 1$. Now if b is another root of f we must have $g(b) = 0$. But g has at most $\deg g$ distinct roots, hence f has at most $\deg g + 1 = \deg f$ distinct roots. ■

Remark. When the field F has infinitely many elements, the function defined by a polynomial uniquely determines the polynomial. Since if there are two distinct polynomials $f, g \in F[x]$ such that $f(a) = g(a)$ for all $a \in F$, then the nonzero polynomial $f - g$ has infinitely many roots, which is in contradiction with the above theorem.

***Example* A.3.** The above theorems show that when F is a field, $F[x]$ is a Euclidean domain with the degree function $d(f) = \deg f$. Since, in addition, for nonzero polynomials f, g we have

$$g \mid f \implies \deg g \leq \deg f.$$

The reason is that if $f = gh$ then $\deg f = \deg g + \deg h$. As a result, the units of $F[x]$ are precisely the nonzero constant polynomials. Because if $u \in F[x]$ is unit then it is invertible, hence it is nonzero. Thus we have $\deg u \leq \deg 1 = 0$ as $u \mid 1$. Therefore $\deg u = 0$, and u is constant. On the other hand, nonzero constant polynomials are units, since they are invertible as F is a field.

Theorem A.6. *Let F be a field. Let $f, g \in F[x]$ be nonzero polynomials, and suppose $\deg g \geq 1$. Then there is a unique integer $m \geq 0$, and unique polynomials $r_0, r_1, \ldots, r_m \in F[x]$ with $r_m \neq 0$, such that*

$$f = r_m g^m + r_{m-1} g^{m-1} + \cdots + r_1 g + r_0,$$

and $\deg r_i < \deg g$ for each i.

Proof. First we prove the existence. The proof is by induction on $\deg f$. If $0 \leq \deg f < \deg g$, then we can put $m = 0$ and $r_0 = f \neq 0$. Suppose the conclusion holds for all polynomials with degree less than $\deg f$. We can assume that $\deg f \geq \deg g$. Now we have $f = gq + r$ where $\deg r < \deg g$. Then $gq = f - r$, so $\deg q + \deg g = \deg(f - r)$. But $\deg r < \deg f$. Thus $\deg(f - r) = \deg f$, since subtracting r does not change the coefficient of the highest degree term in f. Hence

$$\deg q = \deg f - \deg g < \deg f.$$

Also $q \neq 0$, since otherwise we would have $f = r$, which implies $\deg f = \deg r < \deg g$. Therefore by the induction hypothesis we have

$$q = s_m g^m + \cdots + s_1 g + s_0,$$

for some $s_i \in F[x]$, with $s_m \neq 0$, and $\deg s_i < \deg g$. Then we have

$$f = gq + r = s_m g^{m+1} + \cdots + s_1 g^2 + s_0 g + r,$$

as desired.

For the uniqueness, again the proof is by induction on $\deg f$. If $0 \leq \deg f < \deg g$, then the representation is unique. Because if

$$f = r_m g^m + r_{m-1} g^{m-1} + \cdots + r_1 g + r_0$$

for some $m > 0$, where $\deg r_i < \deg g$ and $r_m \neq 0$, then

$$\begin{aligned} \deg(r_{m-1} g^{m-1} + \cdots + r_1 g + r_0) &\leq \max_{j \leq m-1} \big(\deg(r_j g^j)\big) \\ &= \max_{j \leq m-1} (\deg r_j + j \deg g) \\ &< m \deg g \leq \deg(r_m g^m). \end{aligned}$$

Hence $\deg f = \deg(r_m g^m) \geq \deg g$, which is contrary to our assumption. Thus $m = 0$, and therefore $r_0 = f$.

Now suppose the uniqueness holds for all polynomials with degree less than $\deg f$. We can assume that $\deg f \geq \deg g$. Suppose that

$$s_k g^k + \cdots + s_1 g + s_0 = f = r_m g^m + \cdots + r_1 g + r_0,$$

where $\deg r_i, \deg s_j < \deg g$, and $r_m, s_k \neq 0$. Then $m, k > 0$, since $\deg f \geq \deg g > \deg r_0, \deg s_0$. Now we have

$$(s_k g^{k-1} + \cdots + s_1) g + s_0 = f = (r_m g^{m-1} + \cdots + r_1) g + r_0.$$

Therefore r_0, s_0 are the remainder in the division of f by g. Hence $r_0 = s_0$, since the remainder and the quotient in the division of polynomials are unique. Thus we also have

$$s_k g^{k-1} + \cdots + s_1 = r_m g^{m-1} + \cdots + r_1.$$

But, in the first part of this proof we showed that when $\deg f \geq \deg g$ then the quotient is a nonzero polynomial whose degree is strictly less than $\deg f$. Therefore by the induction hypothesis we have $m - 1 = k - 1$, and $r_i = s_i$ for $1 \leq i \leq m$. Hence $m = k$, and the representation of f is unique. ∎

Remark. As a special case of the above theorem we set $g(x) = x - a$, for some $a \in F$. Then $\deg r_i < \deg g = 1$, so each r_i is a constant. Hence for every $f \in F[x]$ there are unique $a_1, \ldots, a_m \in F$, where $a_m \neq 0$ when f is not constant, such that

$$f(x) = a_m(x-a)^m + \cdots + a_1(x-a) + f(a).$$

The fact that $r_0 = f(a)$ follows easily by evaluating both sides of the identity at a. Note that we can allow f to be zero by setting each $a_i = 0$.

Definition A.5. A field F is called **algebraically closed**, if every nonconstant polynomial with coefficients in F has at least one root in F.

Theorem A.7. *Let f be a polynomial with coefficients in an algebraically closed field F. Suppose $\deg f = n \geq 1$. Then there are (not necessarily distinct) elements $a_1, \ldots, a_n \in F$, and $c \in F - \{0\}$, such that*

$$f(x) = c(x-a_1)\cdots(x-a_n).$$

Proof. The proof is by induction on n. For $n = 1$ the claim holds trivially. Now suppose it also holds for polynomials of degree $n-1$. Then we know that f has at least one root a_1. Hence there is a polynomial g of degree $n-1$ such that

$$f(x) = (x-a_1)g(x).$$

Now by the induction hypothesis g has a factorization

$$g(x) = c(x-a_2)\cdots(x-a_n).$$

Thus we get the desired factorization for f. Finally note that if $c = 0$ then $f = 0$, which contradicts the fact that $\deg f \geq 1$. ∎

Definition A.6. Suppose R is a commutative ring. We inductively define the ring of **polynomials in n variables** with coefficients in R to be the commutative ring

$$R[x_1, \ldots, x_n] := R[x_1, \ldots, x_{n-1}][x_n].$$

Remark. Intuitively we consider a polynomial p in n variables to be a formal sum of finitely many expressions of the form $rx_1^{m_1}\cdots x_n^{m_n}$, where $r \in R$. Then we can collect all terms in which x_n has power m_n, and factor $x_n^{m_n}$ out, to get $q(x_1, \ldots, x_{n-1})x_n^{m_n}$, where q is (what we intuitively consider) a polynomial in $n-1$ variables. Thus we can write p as a sum of terms of this form, with different powers of x_n. In other words, p is a polynomial in x_n whose coefficients are polynomials in $n-1$ variables. Continuing inductively we can see that our intuitive notion of polynomials in n variables is the same notion described rigorously in the above definition.

Remark. When R is an integral domain, $R[x_1]$ is also an integral domain. Hence $R[x_1, x_2] = R[x_1][x_2]$ is an integral domain too. By an easy induction it follows that for any n, $R[x_1, \ldots, x_n]$ is also an integral domain.

Remark. Every polynomial in n variables defines a function on R^n by evaluation. Let $f(x_1, \ldots, x_n)$ be a polynomial in n variables. Then by definition we have

$$f(x_1, \ldots, x_n) = \\ f_0(x_1, \ldots, x_{n-1}) + f_1(x_1, \ldots, x_{n-1})x_n + \cdots + f_m(x_1, \ldots, x_{n-1})x_n^m,$$

where f_j's are polynomials in $n-1$ variables which are uniquely determined by f. Then we can inductively define the value of f at $(a_1, \ldots, a_n) \in R^n$ to be

$$f(a_1, \ldots, a_n) := \\ f_0(a_1, \ldots, a_{n-1}) + f_1(a_1, \ldots, a_{n-1})a_n + \cdots + f_m(a_1, \ldots, a_{n-1})a_n^m.$$

It can also be proved inductively that the value of f at a point is uniquely determined by f and that point, since the same is true for each f_j, and f_j's are uniquely determined by f.

Theorem A.8. *Suppose R is a commutative ring. Then every polynomial $f \in R[x_1, \ldots, x_n]$ can be written as a sum of finitely many* ***monomials****, i.e.*

$$f(x_1, \ldots, x_n) = \sum_{m_1 \le k_1} \cdots \sum_{m_n \le k_n} r_{m_1 \ldots m_n} x_1^{m_1} \cdots x_n^{m_n},$$

where $r_{m_1 \ldots m_n} \in R$ and $k_1, \ldots, k_n$ are nonnegative integers. Furthermore, this representation of f is unique, and for every $(a_1, \ldots, a_n) \in R^n$ we have

$$f(a_1, \ldots, a_n) = \sum_{m_1 \le k_1} \cdots \sum_{m_n \le k_n} r_{m_1 \ldots m_n} a_1^{m_1} \cdots a_n^{m_n}.$$

Proof. The proof is by induction on n. The case of $n = 1$ is obvious. So suppose the theorem is true for $n-1$. First note that the monomials are actually polynomials in n variables, since if $r x_1^{m_1} \cdots x_{n-1}^{m_{n-1}} \in R[x_1, \ldots, x_{n-1}]$ then by definition we have $r x_1^{m_1} \cdots x_{n-1}^{m_{n-1}} x_n^{m_n} \in R[x_1, \ldots, x_n]$. Now we know that

$$f(x_1, \ldots, x_n) = \\ f_0(x_1, \ldots, x_{n-1}) + f_1(x_1, \ldots, x_{n-1})x_n + \cdots + f_m(x_1, \ldots, x_{n-1})x_n^m,$$

where f_j's are polynomials in $n-1$ variables. By the induction hypothesis each f_j can be written as a sum of finitely many monomials in $n-1$ variables. If we substitute those expansions into the above formula for f, and multiply them by x_n^j, then we get an expansion of f into a sum of finitely many monomials in n variables.

Now let us prove the second statement. We have

$$\begin{aligned} f(x_1,\ldots,x_n) &= \sum_{m_1\le k_1}\cdots\sum_{m_n\le k_n} r_{m_1\ldots m_n} x_1^{m_1}\cdots x_n^{m_n} \\ &= \sum_{j=0}^{k_n}\Bigg(\sum_{m_1\le k_1}\cdots\sum_{m_{n-1}\le k_{n-1}} r_{m_1\ldots m_{n-1}j} x_1^{m_1}\cdots x_{n-1}^{m_{n-1}}\Bigg)x_n^j \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(We replaced } m_n \text{ with } j.\text{)} \\ &= \sum_{j\le k_n} f_j(x_1,\ldots,x_{n-1})x_n^j, \end{aligned}$$

where $f_j := \sum_{m_1\le k_1}\cdots\sum_{m_{n-1}\le k_{n-1}} r_{m_1\ldots m_{n-1}j}x_1^{m_1}\cdots x_{n-1}^{m_{n-1}}$ is a polynomial in $n-1$ variables. Hence by the induction hypothesis, $r_{m_1\ldots m_{n-1}j}$'s are uniquely determined by f_j, and we have

$$f_j(a_1,\ldots,a_{n-1}) = \sum_{m_1\le k_1}\cdots\sum_{m_{n-1}\le k_{n-1}} r_{m_1\ldots m_{n-1}j}a_1^{m_1}\cdots a_{n-1}^{m_{n-1}}.$$

On the other hand, f_j's must be the coefficients of $f \in R[x_1,\ldots,x_{n-1}][x_n]$, since the coefficients of f are uniquely determined by f. Thus $r_{m_1\ldots m_n}$'s are uniquely determined by f, and we have

$$\begin{aligned} f(a_1,\ldots,a_n) &= \sum_{j\le k_n} f_j(a_1,\ldots,a_{n-1})a_n^j \\ &= \sum_{j\le k_n}\Bigg(\sum_{m_1\le k_1}\cdots\sum_{m_{n-1}\le k_{n-1}} r_{m_1\ldots m_{n-1}j}a_1^{m_1}\cdots a_{n-1}^{m_{n-1}}\Bigg)a_n^j \\ &= \sum_{m_1\le k_1}\cdots\sum_{m_n\le k_n} r_{m_1\ldots m_n}a_1^{m_1}\cdots a_n^{m_n}. \quad \text{(We replaced } j \text{ with } m_n.\text{)} \end{aligned}$$

Note that we have used Theorem 5.39 several times, to change the order of summations. ■

Remark. When we expand the polynomial $f \in R[x_1,\ldots,x_n]$ into a sum of monomials as described in the above theorem, the elements $r_{m_1\ldots m_n} \in R$ are referred to as the **coefficients** of f. Note that we sometimes consider the coefficients of f to be polynomials of $n-1$ variables, but usually we consider the coefficients of f to be $r_{m_1\ldots m_n}$'s.

Definition A.7. The **elementary symmetric polynomials** in n variables

$x_1, \dots, x_n$ are

$$\begin{aligned}
s_1 &:= x_1 + x_2 + \cdots + x_n, \\
s_2 &:= x_1x_2 + x_1x_3 + \cdots + x_2x_3 + \cdots + x_{n-1}x_n, \\
&\vdots \\
s_k &:= \sum_{i_1=1}^{n-k+1} \sum_{i_2=i_1+1}^{n-k+2} \cdots \sum_{i_k=i_{k-1}+1}^{n} x_{i_1}x_{i_2}\cdots x_{i_k}, \\
&\vdots \\
s_n &:= x_1x_2\cdots x_n.
\end{aligned}$$

Theorem A.9. *Suppose R is a commutative ring and $a_1, \dots, a_n$ are (not necessarily distinct) elements of R. Then*

$$(x - a_1)\cdots(x - a_n) = x^n - b_1x^{n-1} + b_2x^{n-2} - \cdots + (-1)^n b_n,$$

where $b_k = s_k(a_1, \dots, a_n)$.

Proof. The proof is by induction on n. The case of $n = 1$ is trivial. Suppose the claim holds for $n - 1$. Let $\tilde{s}_1, \dots, \tilde{s}_n$ be the elementary symmetric polynomials in $n - 1$ variables. Then we have

$$(x - a_1)\cdots(x - a_{n-1}) = x^{n-1} - c_1x^{n-2} + \cdots + (-1)^{n-1}c_{n-1},$$

where $c_k = \tilde{s}_k(a_1, \dots, a_{n-1})$. Now we have

$$\begin{aligned}
(x - a_1)\cdots(x - a_n) &= (x^{n-1} - c_1x^{n-2} + \cdots + (-1)^{n-1}c_{n-1})(x - a_n) \\
&= x^n - (c_1 + a_n)x^{n-1} + (c_2 + c_1a_n)x^{n-2} \\
&\quad - (c_3 + c_2a_n)x^{n-3} + \cdots + (-1)^n c_{n-1}a_n.
\end{aligned}$$

But for $1 < k < n$ we have

$$\begin{aligned}
b_k &= \sum_{i_1=1}^{n-k+1} \sum_{i_2=i_1+1}^{n-k+2} \cdots \sum_{i_k=i_{k-1}+1}^{n} a_{i_1}a_{i_2}\cdots a_{i_k} \\
&= \sum_{i_1=1}^{n-k} \sum_{i_2=i_1+1}^{n-k+1} \cdots \sum_{i_{k-1}=i_{k-2}+1}^{n-2} \sum_{i_k=i_{k-1}+1}^{n-1} a_{i_1}a_{i_2}\cdots a_{i_k} \\
&\quad + \Big(\sum_{i_1=1}^{n-k+1} \sum_{i_2=i_1+1}^{n-k+2} \cdots \sum_{i_{k-1}=i_{k-2}+1}^{n-1} a_{i_1}a_{i_2}\cdots a_{i_{k-1}} \Big) a_n.
\end{aligned}$$

Hence $b_k = c_k + c_{k-1}a_n$. It is also obvious that $b_1 = c_1 + a_n$, and $b_n = c_{n-1}a_n$. Therefore we get the desired formula. ∎

***Example* A.4.** Let F be a field. The field of fractions of the ring of polynomials $F[x_1, \ldots, x_n]$ is denoted by $F(x_1, \ldots, x_n)$, and is called the field of **rational functions in n-variables** over F. The elements of $F(x_1, \ldots, x_n)$ are of the form

$$\frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n)},$$

where $f, g \in F[x_1, \ldots, x_n]$ are polynomials in n-variables. These elements are called rational functions. For $(a_1, \ldots, a_n) \in F$ if $g(a_1, \ldots, a_n) \neq 0$, we can compute the value of the rational function $\frac{f}{g}$ at $(a_1, \ldots, a_n)$ to be

$$\frac{f(a_1, \ldots, a_n)}{g(a_1, \ldots, a_n)} \in F.$$

A.3 Principal Ideal Domains

Definition A.8. Let R be a commutative ring. An **ideal** is a nonempty subset $I \subset R$ such that

(i) For every $a, b \in I$ we have $a + b \in I$.
(ii) For every $a \in I$ and $r \in R$ we have $ra \in I$.

***Example* A.5.** Let $a_1, \ldots, a_n \in R$. Then it is easy to see that the set

$$(a_1, \ldots, a_n) := \{r_1 a_1 + \cdots + r_n a_n : \text{ for all } r_i \in R\}$$

is an ideal. It is called the *ideal generated by* $a_1, \ldots, a_n$.

Definition A.9. An ideal I is called **principal** if

$$I = (a) = \{ra : r \in R\}$$

for some $a \in R$.

Proposition A.3. *Suppose R is an integral domain, and $a, b \in R$. Then*

(i) $(a) = R$ *if and only if* a *is a unit.*
(ii) $(b) \subset (a)$ *if and only if* $a \mid b$*, if and only if* $b \in (a)$.
(iii) $(b) = (a)$ *if and only if* a, b *are associates.*

Proof. **(i)** If a is unit then $1 = a^{-1}a \in (a)$. Hence for all $r \in R$ we have $r = r1 \in (a)$. Thus $R = (a)$. Conversely, if $R = (a)$ then $1 \in (a)$. Hence there is $s \in R$ such that $sa = 1$, i.e. a is a unit.

(ii) If $a \mid b$ then $b = sa$ for some $s \in R$. Hence $b \in (a)$. Thus for all $r \in R$ we have $rb \in (a)$, i.e. $(b) \subset (a)$. Conversely, if $(b) \subset (a)$ then $b = 1b \in (b) \subset (a)$. Therefore $b = sa$ for some $s \in R$, hence $a \mid b$.

(iii) $(b) = (a)$ is equivalent to $(b) \subset (a)$ and $(a) \subset (b)$. Thus $(b) = (a)$ if and only if $a \mid b$ and $b \mid a$, i.e. if and only if a, b are associates. ■

Definition A.10. An integral domain is called a **principal ideal domain (PID)**, if all of its ideals are principal.

Theorem A.10. *Every Euclidean domain is a PID.*

Proof. Let I be an ideal in R. Let $a \in I$ be a nonzero element that has the least degree among all nonzero elements of I. This is possible due to the well-ordering of nonnegative integers. We claim that $I = (a)$. It is obvious that $(a) \subset I$. For the other inclusion, let b be an arbitrary element of I. We have $b = aq + r$ where either $r = 0$ or $d(r) < d(a)$. Since $r = b + (-q)a \in I$, we cannot have $d(r) < d(a)$. Thus $r = 0$, and therefore $b \in (a)$. Hence $I \subset (a)$ as desired. ■

Definition A.11. A **greatest common divisor (g.c.d)** of two nonzero elements a, b in an integral domain R, is an element $c \in R$ such that

(i) $c \mid a$ and $c \mid b$.

(ii) If $r \in R$ is a common divisor of a, b, i.e. if $r \mid a$ and $r \mid b$, then $r \mid c$.

Remark. Note that this definition is a bit different from the definition of g.c.d given for $R = \mathbb{Z}$ in Section 5.3. There we required $r \leq c$. However, in an arbitrary integral domain we do not have a notion of order, so we need to work with the above more general definition. As we have seen, if c is the g.c.d of two nonzero integers a, b, then it satisfies the above definition too. But $-c$ also satisfies the above definition. So, with the above more general definition, $\pm c$ are both g.c.d of a, b. The following proposition shows that these are the only g.c.d of the integers a, b.

Proposition A.4. *Suppose R is an integral domain, and $a, b \in R$ are nonzero. Then any two greatest common divisors of a, b are associates.*

Proof. Suppose c_1, c_2 are greatest common divisors of a, b. Then c_1, c_2 are both common divisors of a, b. Hence we must have $c_1 \mid c_2$ and $c_2 \mid c_1$, since c_1, c_2 are both greatest common divisors of a, b. Thus c_1, c_2 are associates. ■

Euclidean Algorithm. *Suppose R is a Euclidean domain, and $a, b \in R$ are nonzero. Consider the following sequence of divisions*

$$\begin{aligned} a &= bq_1 + r_1, \\ b &= r_1 q_2 + r_2, \\ r_1 &= r_2 q_3 + r_3, \\ &\vdots \end{aligned}$$

In this sequence, after finitely many divisions the remainder becomes zero, i.e. for some $n \geq 0$ we have (we set $r_0 = b$ and $r_{-1} = a$)

$$r_{n-1} = r_n q_{n+1}.$$

Then r_n is a greatest common divisor of a, b.

Proof. We provide a brief sketch of the proof, which is essentially the same as of the Euclidean algorithm for integers, given in Section 5.3. First note that $d(b) > d(r_1) > d(r_2) > \cdots \geq 0$. Hence the process must stop at some point, since the d-values are nonnegative integers. Therefore the division must be impossible at some step, which means that the remainder in the last step is zero. Now let us show that the remainder from one step before the last step, i.e. r_n, is a g.c.d of a, b. First note that $r_n \mid r_{n-1}$. Thus from $r_{n-2} = r_{n-1}q_n + r_n$ we see that $r_n \mid r_{n-2}$. If we continue inductively we get $r_n \mid b$ and $r_n \mid a$. Hence r_n is a common divisor of a, b. Next suppose $c \mid a$ and $c \mid b$. Then $c \mid a - bq_0 = r_0$. Again we can show inductively that $c \mid r_n$. Thus r_n is a g.c.d of a, b. ■

Theorem A.11. *Suppose R is a PID, and $a, b \in R$ are nonzero. Then a, b have a greatest common divisor $c \in R$. Furthermore we have*

$$c = ra + sb,$$

for some $r, s \in R$.

Proof. The ideal generated by a, b, i.e. (a, b) is principal, since R is a PID. Thus there is $c \in R$ such that

$$(a, b) = (c).$$

We claim that c is a greatest common divisor of a, b. Since $a, b \in (a, b) = (c)$, we have $c \mid a$ and $c \mid b$. On the other hand $c \in (c) = (a, b)$, so $c = ra + sb$ for some $r, s \in R$. Thus if an element $q \in R$ divides both a, b, then it will also divide both ra, sb. Hence q divides $ra + sb = c$ too. ■

Remark. Recall that when R is a Euclidean domain, the generator of an ideal is a nonzero element with the least degree in that ideal. Hence in this case, we can say that a g.c.d of a, b is a nonzero element with the least degree in the set

$$\{ra + sb : r, s \in R\}.$$

Remark. The above two theorems are in particular true in $F[x]$ when F is a field. (They are also true in $\mathbb{Z}$ as we have seen in Section 5.3.)

A.4 Unique Factorization Domains

Definition A.12. Let R be an integral domain. A nonzero element $r \in R$ is called **irreducible** if it is not a unit, and its only divisors are units or associates of itself, i.e.

$$r = ab \implies a, b \text{ are either unit, or associates of } r.$$

A nonzero element $p \in R$ is called **prime**, if it is not a unit, and for all $a, b \in R$ we have

$$p \mid ab \implies p \mid a \text{ or } p \mid b.$$

Remark. The contrapositive of the definition of a prime element p is that

$$p \nmid a \text{ and } p \nmid b \implies p \nmid ab.$$

By an easy induction we can show that if p is prime, then for $a_1, \ldots, a_n \in R$ we have

$$p \nmid a_1 \text{ and } p \nmid a_2 \text{ and } \ldots \text{ and } p \nmid a_n \implies p \nmid a_1 \cdots a_n.$$

Equivalently, if $p \mid a_1 \cdots a_n$ then $p \mid a_i$ for some i.

Remark. Note that when $R = \mathbb{Z}$ we used irreducibility as the definition of prime numbers. However, Euclid's lemma shows that prime numbers are also prime elements as defined above.

***Exercise* A.2.** Let R be an integral domain, and suppose $a, b \in R$ are associates. Show that if a is irreducible then b is also irreducible, and if a is prime then b is also prime.

Theorem A.12. *Every prime element of an integral domain is irreducible.*

Proof. Suppose p is a prime, and we have a factorization $p = ab$. Then $p \mid ab$ so either $p \mid a$ or $p \mid b$. On the other hand, both a, b divide p. Thus either a is an associate of p, or b is. Suppose for instance that a, p are associates. Then $a = pu$ for some unit u. Hence

$$0 = p - ab = p - pub = p(1 - ub) \implies ub = 1.$$

Therefore b is a unit. ■

Theorem A.13. *Every irreducible element of a PID is prime.*

Proof. Let r be an irreducible element of the PID R. Suppose for $a, b \in R$ we have $r \mid ab$, and $r \nmid a$. We must show that $r \mid b$. We claim that the greatest common divisor of r, a is a unit. The reason is that if the g.c.d of r, a is u, then $u \mid r$. Hence u is either a unit or an associate of r, since r is irreducible. But $u \mid a$ too, so it cannot be an associate of r. Because otherwise we would have $u = rv$ for some unit element v, and therefore $r \mid a$, which is a contradiction.

Now we know that for some $x, y \in R$ we have $u = xr + ya$. Therefore we have

$$b = u^{-1}xbr + u^{-1}yab.$$

Since r divides the right hand side of the above equation, we get $r \mid b$ as desired. ■

Definition A.13. An integral domain R is called a **unique factorization domain (UFD)**, if every nonzero element of R that is not a unit, can be written as a product of irreducible elements in a unique way. In other words, for all nonzero $a \in R$ which is not a unit we have

(i) There are irreducible elements $p_1, \dots, p_n \in R$ such that a has the factorization

$$a = p_1 \cdots p_n.$$

(ii) If there is another factorization of a into irreducible elements $a = q_1 \cdots q_m$, then $m = n$, and there is a permutation $\sigma \in S_n$ such that $p_i, q_{\sigma(i)}$ are associates.

***Exercise* A.3.** Show that every irreducible element of a UFD is prime.

Theorem A.14. *Every PID is a UFD.*

Proof. The uniqueness of a factorization is a consequence of the fact that in a PID irreducible elements are prime. Suppose

$$p_1 \cdots p_n = q_1 \cdots q_m,$$

where p_i, q_j's are primes. We proceed by induction on n. When $n = 1$ we have $p_1 \mid q_1 \cdots q_m$. Thus $p_1 \mid q_k$ for some k. But q_k is irreducible and p_1 is not a unit, hence $p_1 = uq_k$ for some unit u. Therefore

$$1 = q_1 \cdots q_{k-1} u^{-1} q_{k+1} \cdots q_m,$$

since the cancellation law holds in integral domains. Hence if $m > 1$, the other q_j's must be units, which is a contradiction. Thus $m = 1$ and $p_1 = uq_1$.

Now suppose the uniqueness holds for some n. Consider the case of $n + 1$, and suppose we have

$$p_1 \cdots p_n p_{n+1} = q_1 \cdots q_m.$$

Then $p_{n+1} \mid q_1 \cdots q_m$, and therefore $p_{n+1} \mid q_k$ for some k. We can argue as above and conclude that for some unit u we have $p_{n+1} = uq_k$, and hence

$$p_1 \cdots p_n = q_1 \cdots q_{k-1} u^{-1} q_{k+1} \cdots q_m = q_1 \cdots q_{k-1} (u^{-1} q_{k+1}) \cdots q_m.$$

Note that $u^{-1}q_{k+1}$ is also irreducible. Now by the induction hypothesis we have $n = m - 1$, and there is a permutation $\sigma \in S_n$ such that for $i \leq n$, $p_i, q_{\sigma(i)}$ are associates when $\sigma(i) < k$, $p_i, u^{-1}q_{k+1}$ are associates when $\sigma(i) = k$, and $p_i, q_{\sigma(i)+1}$ are associates when $\sigma(i) > k$. Let

$$\hat{\sigma}(i) := \begin{cases} \sigma(i) & i \leq n,\ \sigma(i) < k \\ k & i = n + 1 \\ \sigma(i) + 1 & i \leq n,\ \sigma(i) \geq k \end{cases}$$

be a permutation in S_{n+1}. Then $p_i, q_{\tilde{\sigma}(i)}$ are associates for all $i \leq n+1$ as desired. Note that when $p_i, u^{-1}q_{k+1}$ are associates, then p_i, q_{k+1} are associates too.

Next for the existence of a factorization, let a be a nonzero element of our PID, which is not a unit. Suppose to the contrary that a does not have a factorization into irreducible elements. Then a cannot be irreducible itself, since then we have the factorization $a = a$. Thus $a = bc$ where b, c are not unit nor an associate of a. If both b, c can be factorized into irreducible elements, then a can be factorized too. Hence at least one of them, which we call it a_1, does not have a factorization. As a_1 is not an associate of a and divides a, we have

$$(a) \subsetneq (a_1).$$

Since a_1 does not have a factorization into irreducible elements, we can argue as above and find another element a_2 such that

$$(a) \subsetneq (a_1) \subsetneq (a_2).$$

We can continue this process inductively and get

$$(a) \subsetneq (a_1) \subsetneq \cdots \subsetneq (a_n) \subsetneq \cdots .$$

Now, let $I := \bigcup_{n\geq 1}(a_n)$. It is easy to see that I is an ideal. Because if $b, c \in I$ then $b \in (a_i)$ and $c \in (a_j)$ for some i, j. Let $n = \max\{i, j\}$. Then $b, c \in (a_n)$. Hence $b + c \in (a_n) \subset I$, and also $sb \in (a_n) \subset I$ for all elements s. Thus we have $I = (r)$ for some element r. Then $r \in I$, so $r \in (a_m)$ for some m. But this means that $(r) \subset (a_m)$. In particular we have $(a_{m+1}) \subset (a_m)$, which is a contradiction. Therefore a must have a factorization into irreducible elements. ■

Second Proof. Here we give another proof for the existence of the factorization when our PID is a Euclidean domain. Let a be a nonzero element which is not a unit. The proof is by strong induction on $d(a)$, the degree of a. If a has the least d-value among the nonzero elements that are not units, then a must be irreducible. To see this suppose that $a = bc$ and c is not a unit. Note that $b, c \neq 0$ since $a \neq 0$. Then by Proposition A.2 we have $d(b) < d(bc) = d(a)$. But a has the least d-value among nonzero elements which are not units, so b must be unit. Therefore a is irreducible. In particular, a has a factorization into irreducible elements.

Now suppose every element with degree less than $d(a)$ has a factorization into irreducible elements. If a is irreducible, then it has a factorization. Otherwise we have $a = bc$, where b, c are nonzero and they are not units. Hence again by Proposition A.2 we have $d(b) < d(bc) = d(a)$. Similarly $d(c) < d(bc) = d(a)$. Therefore by the induction hypothesis b, c can be written as a product of irreducible elements. Now if we multiply those expressions we obtain a factorization of a into irreducible elements, as desired. ■

***Example* A.6.** As we have seen in Section 5.3, $\mathbb{Z}$ is a Euclidean domain, so it is a PID, hence it is a UFD. Note that ± 1 are the only units of $\mathbb{Z}$, since they are the only invertible elements of the ring $\mathbb{Z}$.

Theorem A.15. *Suppose F is a field. Then $F[x]$ is a UFD, i.e. every nonconstant polynomial with coefficients in a field can be written uniquely as a product of irreducible polynomials.*

Proof. $F[x]$ is a Euclidean domain, so it is a PID, hence it is a UFD. Note that the nonzero constant polynomials are precisely the units of $F[x]$, as we showed in Example A.3. ■

Appendix B

Decimal Expansion

Theorem B.1. *Consider the set of sequences of the form*

$$a_0.a_1a_2a_3\ldots,$$

where $a_0 \in \mathbb{N}\cup\{0\}$, $a_j \in \{0,1,2,3,4,5,6,7,8,9\}$ *for* $j>0$, *and for every* j *there is* $k>j$ *such that* $a_k \neq 9$. *Then there exists a one-to-one and onto map from this set of sequences to* $[0,\infty)\subset\mathbb{R}$, *that maps the sequence* $a_0.a_1a_2a_3\ldots$ *to the real number*

$$x := \sup\{x_n : n\in\mathbb{N}\},$$

where $x_n := a_0 + \sum_{j=1}^{n}\frac{a_j}{10^j}$.

Definition B.1. The unique sequence $a_0.a_1a_2a_3\ldots$ whose existence is proved in the above theorem is called the **decimal expansion** of x.

Remark. By using the concepts of *limit* and *series*, we can easily conclude from the following proof that

$$x = \lim_{n\to\infty} x_n = a_0 + \sum_{j=1}^{\infty}\frac{a_j}{10^j}.$$

Remark. In the following proof, we actually provide a concrete method for finding the decimal expansion of x.

Proof. First note that $x_{n+1} = x_n + \frac{a_{n+1}}{10^{n+1}} \geq x_n$. So x_n forms an increasing sequence, i.e. it is an increasing function of n. Next note that the set $\{x_n\}$ is nonempty and bounded above, since

$$a_0 + \sum_{j\leq n}\frac{a_j}{10^j} \leq a_0 + \sum_{j\leq n}\frac{9}{10^j} = a_0 + 9\frac{\frac{1}{10}-\frac{1}{10^{n+1}}}{1-\frac{1}{10}} = a_0 + 1 - \frac{1}{10^n} < a_0 + 1.$$

Therefore $x = \sup\{x_n\}$ exists. In addition, for some n we have $x \geq x_n \geq a_0 \geq 0$; thus $x \in [0, \infty)$.

Now let us show that the map defined in the theorem is one-to-one. Suppose $a_0.a_1a_2a_3\ldots$ and $b_0.b_1b_2b_3\ldots$ are distinct sequences, which are mapped to x, y respectively. We want to show that $x \neq y$. Suppose l is the smallest index for which we have $a_l \neq b_l$. We will show that if $a_l < b_l$ then $x < y$. We know that there is $k > l$ such that $a_k \neq 9$. Hence $a_k \leq 8 = 9 - 1$. We also have $a_l + 1 \leq b_l$. Then for $n > k$ we have

$$\begin{aligned}
x_n = a_0 + \sum_{j\leq n} \frac{a_j}{10^j} &= a_0 + \sum_{j<l} \frac{a_j}{10^j} + \frac{a_l}{10^l} + \sum_{l<j<k} \frac{a_j}{10^j} + \frac{a_k}{10^k} + \sum_{k<j\leq n} \frac{a_j}{10^j} \\
&\leq a_0 + \sum_{j<l} \frac{a_j}{10^j} + \frac{a_l}{10^l} + \sum_{l<j<k} \frac{9}{10^j} + \frac{9-1}{10^k} + \sum_{k<j\leq n} \frac{9}{10^j} \\
&= a_0 + \sum_{j<l} \frac{a_j}{10^j} + \frac{a_l}{10^l} - \frac{1}{10^k} + 9\frac{\frac{1}{10^{l+1}} - \frac{1}{10^{n+1}}}{1 - \frac{1}{10}} \\
&= b_0 + \sum_{j<l} \frac{b_j}{10^j} + \frac{a_l}{10^l} - \frac{1}{10^k} + \frac{1}{10^l} - \frac{1}{10^n} \\
&\leq b_0 + \sum_{j<l} \frac{b_j}{10^j} + \frac{b_l}{10^l} - \frac{1}{10^k} - \frac{1}{10^n} \\
&= y_l - \frac{1}{10^k} - \frac{1}{10^n} < y_l - \frac{1}{10^k} \leq y_n - \frac{1}{10^k}.
\end{aligned}$$

Thus we get $x_n + \frac{1}{10^k} < y_n \leq y$, since y is the supremum of $\{y_n\}$. Also note that for $n \leq k$ we have $x_n + \frac{1}{10^k} \leq x_{k+1} + \frac{1}{10^k} < y$, because x_n is an increasing sequence. Therefore we get $x \leq y - \frac{1}{10^k}$, since x is the supremum of $\{x_n\}$. Hence we obtain $x < y$, as desired.

Next let us prove that the map defined in the theorem is onto. Let x be a nonnegative real number. Set a_0 to be the integer part of x, i.e.

$$a_0 := \lfloor x \rfloor.$$

We know that $a_0 \leq x < a_0 + 1$. Hence $0 \leq x - a_0 < 1$. Also note that since $x \geq 0$ we have $a_0 > -1$; so $a_0 \geq 0$, since it is an integer. Next let

$$a_1 := \lfloor 10(x - a_0) \rfloor.$$

Note that $0 \leq 10(x - a_0) < 10$. Thus a_1 is a nonnegative integer less than 10, i.e. it belongs to $\{0, 1, \ldots, 9\}$. In addition we have

$$0 \leq 10(x - a_0) - a_1 < 1 \implies 0 \leq x - \big(a_0 + \frac{a_1}{10}\big) < \frac{1}{10}.$$

We continue this process to inductively define a_n. More precisely, we define

(i) $x_0 := \lfloor x \rfloor$,
(ii) $x_{n+1} := x_n + \frac{\lfloor 10^{n+1}(x-x_n)\rfloor}{10^{n+1}}$.

Remark. In the terminology of recursion theorem, we have constructed x_n by using the function $F(n,s) = s + \frac{\lfloor 10^{n+1}(x-s)\rfloor}{10^{n+1}}$.

Then for $n \geq 0$ we define

$$a_{n+1} := 10^{n+1}(x_{n+1} - x_n) = \lfloor 10^{n+1}(x - x_n)\rfloor.$$

It is easy to show that $x_n = a_0 + \sum_{j=1}^{n} \frac{a_j}{10^j}$. Because for $n = 0$ we have $x_0 = \lfloor x \rfloor = a_0$. And if the equality holds for some n, then for $n+1$ we get

$$x_{n+1} = x_n + \frac{a_{n+1}}{10^{n+1}} = a_0 + \sum_{j=1}^{n} \frac{a_j}{10^j} + \frac{a_{n+1}}{10^{n+1}} = a_0 + \sum_{j=1}^{n+1} \frac{a_j}{10^j},$$

as desired.

Now let us show that

$$0 \leq x - x_n < \frac{1}{10^n}. \qquad (*)$$

Note that as a consequence we get $0 \leq 10^{n+1}(x - x_n) < 10$. Therefore $a_{n+1} = \lfloor 10^{n+1}(x-x_n)\rfloor$ is a nonnegative integer less than 10, i.e. it belongs to $\{0, 1, \ldots, 9\}$. The proof of the inequality $(*)$ is by induction on n. For $n = 0$ we have $x_0 = \lfloor x \rfloor$, and the inequality holds due to the properties of the integer part, as we have seen above. Suppose the inequality holds for some n. We know that

$$0 \leq 10^{n+1}(x - x_n) - \lfloor 10^{n+1}(x - x_n)\rfloor < 1,$$

due to the properties of the integer part. Hence we get

$$0 \leq x - \big(x_n + \frac{\lfloor 10^{n+1}(x - x_n)\rfloor}{10^{n+1}}\big) < \frac{1}{10^{n+1}},$$

which is the desired inequality for $n+1$.

Next, note that we have

$$a_n = \lfloor 10^n(x - x_{n-1})\rfloor \leq 10^n(x - x_{n-1}) < a_n + 1.$$

Hence for small positive ϵ we have $10^n(x - x_{n-1}) < a_n + 1 - \epsilon$. But by Archimedean property there is k such that $k > \frac{1}{10\epsilon}$; so $10^k \geq 10k > \frac{1}{\epsilon}$. Thus we get

$$x - x_{n-1} < \frac{a_n}{10^n} + \frac{1}{10^n} - \frac{1}{10^{n+k}}.$$

So $x - x_n < \frac{1}{10^n} - \frac{1}{10^{n+k}}$. Now suppose to the contrary that $a_{n+1} = \cdots = a_{n+k} = 9$. Then we have

$$x_{n+k} - x_n = \sum_{j=n+1}^{n+k} \frac{a_j}{10^j} = \sum_{j=n+1}^{n+k} \frac{9}{10^j} = 9\frac{\frac{1}{10^{n+1}} - \frac{1}{10^{n+k+1}}}{1 - \frac{1}{10}} = \frac{1}{10^n} - \frac{1}{10^{n+k}}.$$

Hence we get

$$x - x_{n+k} = x - x_n + x_n - x_{n+k} < 0,$$

which contradicts the inequality $(*)$. Therefore one of the $a_{n+1}, \ldots, a_{n+k}$ must be less than 9. So the sequence $a_0.a_1a_2a_3\ldots$ satisfies all the properties required in the theorem.

Finally let us show that $a_0.a_1a_2a_3\ldots$ is mapped to x, i.e. x is the supremum of $\{x_n\}$. First note that by inequality $(*)$ we have $x_n \leq x$ for every n. Thus x is an upper bound for $\{x_n\}$. On the other hand, suppose y is an upper bound for $\{x_n\}$. Then $(*)$ implies that $x - \frac{1}{10^n} < x_n \leq y$. So $x - y < \frac{1}{10^n}$ for every n. However, as we have shown above, for every $\epsilon > 0$ there is n such that $\epsilon > \frac{1}{10^n}$, due to the Archimedean property. Hence $x - y < \epsilon$ for every $\epsilon > 0$; so $x - y \leq 0$. Thus $x = \sup\{x_n\}$, as desired. ■