Set theory and topology

An introduction to the foundations of analysis [1]

Part I: Sets, relations, numbers

Felix Nagel

Abstract

We provide a formal introduction into the classic theorems of general topology and its axiomatic foundations in set theory. Starting from ZFC, the exposition in this first part includes relation and order theory as well as a construction of number systems.

[1] For remarks and suggestions please contact: stt.info@t-online.de

The author received his doctoral degree from the University of Heidelberg for his thesis in electroweak gauge theory. He worked for several years as a financial engineer in the financial industry. His fields of interest are probability theory, foundations of analysis, finance, and mathematical physics. He lives in Wales and Lower Saxony.

Preface

This series of articles emerged from the author's personal notes on general topology supplemented by an axiomatic construction of number systems.
At the beginning of the text we introduce our axioms of set theory, from which all results are subsequently derived. In this way the theory is developed *ab ovo* and we do not refer to the literature in any of the proofs.
Needless to say, the presented theory is fundamental to many fields of mathematics like linear analysis, measure theory, probability theory, and theory of partial differential equations. It establishes the notions of relation, function, sequence, net, filter, convergence, pseudo-metric and metric, continuity, uniform continuity etc. To derive the most important classic theorems of general topology is the main goal of the text. Additionally, number systems are studied because, first, important issues in topology are related to real numbers, for instance pseudo-metrics where reals are required at the point of the basic definitions. Second, many interesting examples involve numbers.
In our exposition we particularly put emphasis on the following:

(i) The two advanced concepts of convergence, viz. nets and filters, are treated with almost equal weighting. Most results are presented both in terms of nets and in terms of filters. We use one concept whenever it seems more appropriate than the other.

(ii) We avoid the definition of functions on the ensemble of all sets. Instead we follow a more conservative approach by first choosing an appropriate set in each case on which the respective analysis is based. Notably, this issue occurs in the Recursion theorem for natural numbers (see Theorem 3.13 where, with this restriction, also the Replacement schema is not required), in the Induction principle for ordinal numbers (see Theorem 3.51), and in the Local recursion theorem for ordinal numbers (see Theorem 3.52).

(iii) At many places we try to be as general as possible. In particular, when we analyse relations, many definitions and results are stated in terms of pre-orderings, which we only require to be transitive.

Finally, we would like to warn the reader that for some notions defined in this work there are many differing definitions and notations in the literature, e.g. in the context of relations and orderings. One should always look at the basic definitions before comparing the results.

The text is structured as follows: All definitions occur in the paragraphs explicitly named **Definition**. Important Theorems are named **Theorem**, less important ones **Lemma**, though a distinction seems more or less arbitrary in many cases. In some cases a Lemma and a Definition occur in the same paragraph in order to avoid repetition. Such a paragraph is called **Lemma and Definition**. Claims that lead to a Theorem and are separately stated and proven are called **Proposition**, those derived from Theorems are called **Corollary**. The proofs of most Theorems, Lemmas, Propositions, and Corollaries are given. Within lengthier proofs, intermediate steps are sometimes indented and put in square brackets [...] in order to make the general outline transparent while still explaining every step. Some proofs are left as excercise to the reader. There is no type of paragraph explicitly named as exercise. Paragraphs named **Example** contain specializations of Definitions, Theorems, etc. The analysis of the examples is mostly left to the reader without explicit mention. Statements that do not require extensive proofs and are yet relevant on their own are named **Remark**. Note that Definitions, Theorems etc. are enumerated per Chapter. Some references refer to Chapters that are contained in subsequent parts of this work [Nagel].

Wales, May 2013 *Felix Nagel*

Contents

Part I

Sets, relations, numbers

Chapter 1

Axiomatic foundation

In our exposition, as is the case with every mathematical text, we do not solely use verbal expressions but need a formal mathematical language. To begin with, let us describe which role the formal mathematical language is supposed to play subsequently.

We initially define certain elementary notions of the formal language by means of ordinary language. This is done in Section 1.1. First, we define logical symbols, e.g. the symbol $\Longrightarrow$, which stands for "implies", and the symbol $=$ meaning "equals". Second, we define the meaning of set variables. Every set variable, e.g. the capital letter X of the Latin alphabet, stands for a set. A set has to be interpreted as an abstract mathematical object that has no properties apart from those stated in the theory. Third, we define the symbol $\in$, by which we express that a set is an element of a set. In all three cases the correct interpretation of the symbols comprises, on the one hand, to understand its correct meaning and, on the other hand, not to associate more with it than this pure abstract meaning. A fourth kind of elementary formal component is used occasionally. We sometimes use Greek letters, e.g. φ, as variables that stand for a certain type of formal mathematical expressions called formulae. Such formula variables belong to the elementary parts of our formal language because a formula variable may not only be an abbreviation for a specific formula in order to abridge the exposition but is also used as placeholder for statements that we make about more than a single formula. The latter is tantamount to an abbreviation if a finite number of formulae is supposed to be substituted but cannot be regarded as a mere abbreviation if an infinite number of formulae is considered.

After having translated these elementary mathematical thoughts into the formal language in Section 1.1, we may form formulae out of the symbols. This allows us to express more complicated mathematical statements in our formal language. We follow the rationale that every axiom, definition, and claim in the remainder of this text can in principle be expressed either in the formal language or in the nonformal language and translated in both directions without ambiguity. In practice, in some cases the formal language and in other cases the nonformal language are preferable with respect to legibility and brevity. Therefore we make use of both languages, even mix both deliberately. For example, in order to obtain a precise understanding of our axioms of set theory, we tend to use only the formal language in this context. In most other cases we partly use the formal language to form mathematical expressions, which are then surrounded by elements of the nonformal language. In the field of mathematical logic such formal languages and their interpretations in the nonformal language are probed. There also the nature of mathematical proofs, i.e. the derivation of theorems from assumptions, is analysed. Ways to formalize proofs are proposed so that in principle the derivation of a theorem could be written in a formal language. We do not formalize our proofs in this way but use the fomal language only to the extent described above.

1.1 Formal language

The only objects we consider are sets. Statements about sets are written in our formal language as formulae that consist of certain symbols. We distinguish between symbols that have a fixed meaning wherever they occur, variables that stand for sets, and variables for formulae, which may be substituted if a specific formula is meant. The symbols that have a fixed meaning are the symbol $\in$ and various logical symbols including $=$. As announced in the introduction to this Chapter we now define all these symbols by their meaning in the nonformal language and explain the meaning of variables that stand for sets and those that stand for formulae. We remark that in the remainder of the text further symbols with fixed meaning on a less elementary level are defined as abbreviations.
A variable that denotes a set is called a set variable. We may use as set variables

any small or capital letters of the Latin or Greek alphabet with or without subscripts, superscripts, or other add-ons.

Let x and y be set variables. We write the fact that x is an element of y in our formal language as $(x \in y)$. In this case we also say that x is a member of y. Furthermore we write the fact that x equals y, i.e. x and y denote the same set, as $(x = y)$ in our formal language. Similar translations from the nonformal language to the formal language are applied for membership and equality of any two variables different from x and y.

If x and y are set variables, each of the expressions $(x \in y)$ and $(x = y)$ is called an atomic formula. The same holds for all such expressions containing set variables different from x and y.

Generally, a formula may contain, apart from variables denoting sets, the symbols $\in$ and $=$, the logical connectives $\wedge$ (conjunction, and), $\vee$ (disjunction, nonexclusive or), $\neg$ (negation), $\Longrightarrow$ (implication), $\Longleftrightarrow$ (equivalence), and the quantifiers $\forall$ (for every) and $\exists$ (there exists). All these logical symbols are used in their conventional nonformal interpretation indicated after each symbol above. Additionally, the brackets (and) are used in order to express in which order a formula has to be read. Some of the symbols are clearly redundant to express our nonformal thoughts. For instance, if we use the symbols $\neg$ and $\vee$, then $\wedge$ is not required. Or, if we use the symbols $\neg$ and $\exists$, the symbol $\forall$ is redundant. However, it is often convenient to make use of all logical symbols. As the meaning of all these symbols is defined in our nonformal language, it is clear, for each expression that is written down using these symbols, whether such an expression is meaningful or not. For example, the expressions

(i) $\big(\neg(x \in y)\big) \Longrightarrow (\exists z\ z \in x)$

(ii) $\exists x\ (x \in x)$

(iii) $\big((x = y) \wedge (z \in x)\big) \quad \Longleftrightarrow \quad \big((x \in z) \vee (z = z)\big)$

(iv) $(x = y) \wedge \neg(x = y)$

where x, y, and z are set variables, are meaningful, though some may be logically false in a specific context, as e.g. (ii) in the theory presented in this text, or even false in any context like (iv).
In contrast, the expressions

(i) $\neg x$

(ii) $\forall\forall z$

(iii) $\exists x \, \big((x \in y) \vee \big)$

(iv) $\exists \, \big((x = y) \wedge (z \in x)\big)$

where x, y, and z are again set variables, are not meaningful. In (i), a set variable is used in a place where a formula is expected, in (ii) two quantifiers immediately follow each other, in (iii) the disjunction requires a formula on the right-hand side, and in (iv) a variable is expected on the right-hand side of $\exists$. If an expression is meaningful, then it is called a formula.
A variable by which we denote a formula is called a formula variable. If it is evident from the context that a letter is not a formula variable or a defined symbol of the theory, then it is understood that the letter denotes a set variable. For instance, we state the Existence Axiom in Section 1.2, $\exists x \, (x = x)$, and do not explicitly say that x is a set variable.
Given a formula φ, a set variable that occurs in φ is called free in φ if it does not occur directly after a quantifier. We use the convention that if we list set variables in brackets and separated by commas after a formula variable, then the formula variable denotes a formula that contains as free set variables only those listed in the brackets. For example, $\varphi(x, p)$ stands for a formula that has at most x and p as its free set variables.

We introduce one more symbol, $\notin$. By $x \notin y$ we mean $\neg(x \in y)$, and similarly for any set variables different from x and y.
Furthermore we introduce some variations of our formal notation, which is often very convenient. First, in a formula we may deliberately omit pairs of parentheses whenever the way how to reinsert the parentheses is obvious. Second, we sometimes use the following simplified notation after quantifiers. If x and X are set variables and φ is a formula variable, we write $\exists x \in X\ \varphi$ instead of $\exists x\ (x \in X \wedge \varphi)$. Similarly, we write $\forall x \in X\ \varphi$ instead of $\forall x\ (x \in X \implies \varphi)$. The same conventions apply, of course, to any other choice of set and formula variables. Third, given sets x, y, and X, the formula $x \in X \wedge y \in X$ is also written as $x, y \in X$, and similarly for more than two set variables.
In the nonformal language we adopt the convention that instead of saying that a statement holds "for every $x \in X$" we write $(x \in X)$ after the statement; for example we may write "$x \in Y\ (x \in X)$" instead of "$x \in Y$ for every $x \in X$". Moreover we use the acronym "iff" which means "if and only if" and thus corresponds to the symbol $\iff$ in the formal language.
Finally, we remark that the usage of formula variables in this text is restricted to a limited number of occasions. First, formula variables are used in the postulation of two Axiom schemas, the Separation schema, Axiom 1.4, and the Replacement schema, Axiom 1.47, and its immediate consequences Lemma and Definition 1.6, Definitions 1.7 and 1.24, and Lemmas 1.33 and 1.48. Whenever any of these is used later in the text, the formula variable is substituted by an actual formula. In particular, no further derivation is undertaken where formula variables are used without previous substitution of specific formulae. Second, formula variables are used in Definitions 6.2 and 6.17 of the notions "eventually" and "frequently". However, whenever these notions are used later, the formula variable of the definition is substituted by a specific formula.

1.2 Axioms of set theory

The axioms of set theory that we postulate in this Section and use throughout the text are widely accepted in the literature [Bernays,Ebbinghaus,Jech,Suppes]. They are called ZFC ("Zermelo Fraenkel with choice axiom"). There are other axioms that have similar implications for mathematical theories, e.g. NBG ("von Neumann Bernays Gödel"), see [Bernays]. Although we discuss certain aspects of ZFC in this work, the comparison with NBG or other axioms is beyond our scope.
First we postulate an axiom that says that the world of abstract mathematical objects, which are sets and only sets in our theory, contains at least some object.

Axiom 1.1 (Existence)

$$\exists x \; x = x$$

■

Logically, the formula $x = x$ is always true. Thus Axiom 1.1 postulates the existence of a set. The existence of sets does not follow from the other axioms presented subsequently. This is briefly discussed at the end of this Section.
Next we specify a condition under which two sets are equal.

Axiom 1.2 (Extensionality)

$$\forall x \, \forall y \, \big(\forall z \, (z \in x \Longleftrightarrow z \in y) \Longrightarrow x = y\big)$$

■

The interpretation of Axiom 1.2 is that two sets are equal if they have the same elements. The converse implication

$$\forall x\,\forall y\,\big(x = y \Longrightarrow \forall z\,(z \in x \Longleftrightarrow z \in y)\big)$$

is logically true in any theory because of the interpretation of the symbol $=$.

Definition 1.3
Given two sets X and Y, we say that Y **is a subset of** X, written $Y \subset X$ or $X \supset Y$, if the following statement holds:

$$\forall y\; y \in Y \Longrightarrow y \in X$$

We also write $Y \not\subset X$ for $\neg\,(Y \subset X)$. ■

Notice that Definition 1.3 introduces two new symbols $\subset$ and $\supset$ in the formal language, and also specifies a new notion in the nonformal language. Clearly, in the formal language the new symbol is in principle redundant, that is the same expressions can be written down without it. Thus the new symbols are merely abbreviations. Similarly as for $\in$, we also adopt the short notation $Y, Z \subset X$ for $Y \subset X \wedge Z \subset X$.
Next we postulate the axioms that allow us to specify a set in terms of a given property which is formalized by a formula.

Axiom 1.4 (Separation schema)
Let $\varphi(x, p)$ be a formula. We have:

$$\forall p\,\forall X\,\exists Y\,\forall x\,\big(x \in Y \Longleftrightarrow x \in X \wedge \varphi(x, p)\big)$$

■

In Axiom 1.4, for every formula that contains at most x and p as free variables, one axiom is postulated. Therefore not only a single mathematical expression is postulated here, but a method is given how to write down an axiom for every given formula $\varphi(x, p)$. This is called a schema. The general analysis of this concept is beyond our scope. However, it is clear that if we would like to specify sets in terms of certain properties one would either write down an axiom for each desired property and restrict oneself to a limited number of properties or define a generic method to specify the axioms. In ZFC the latter possibility is chosen. However, note that only in Lemma and Definition 1.6, Definitions 1.7 and 1.24, and Lemma 1.33 the schema is used with its formula variable. In all other instances when we refer in this text to the Axiom schema or one of the mentioned definitions or results, we substitute an explicit formula for the formula variable. Since in this text this happens only a finite number of times, we could postulate a finite number of axioms instead of the Separation schema, each with an explicit formula substituted. In this sense, Axiom 1.4 is only an abbreviated notation of a list of a finite number of axioms that do not contain formula variables. Notice however that the restriction to a finite number of axioms generally also constrains the implications that can be concluded *from* the statements proven in this text.

The Separation schema has an important consequence, viz. there exists a set that has no element.

Lemma and Definition 1.5

There is a unique set Y such that there exists no set x with $x \in Y$. Y is called the **empty set**, written $\emptyset$.

Proof. We may choose a set X by the Existence axiom. Let $\varphi(x)$ denote the formula $\neg(x = x)$. This formula is logically false in any theory for every x. There exists a set Y such that

$$\forall x \; x \in Y \Longleftrightarrow x \in X \wedge \neg(x = x)$$

by the Separation schema. Clearly, Y has no element. The uniqueness of Y follows by the Extensionality axiom. $\square$

We have postulated the existence of a set by the Existence axiom and concluded in Lemma and Definition 1.5 that the empty set exists. However, we have not proven so far that any other set exists. This is remedied by the Power set axiom to be introduced below in this Section, and, even without Power set axiom, by the Infinity axiom below.

We now introduce several notations that are all well-defined by the Axioms postulated so far, namely the curly bracket notation for sets, the intersection of two sets, the intersection of one set, and the difference of two sets.

Lemma and Definition 1.6

Let X and p be sets, and $\varphi(x,p)$ a formula. There is a unique set Y such that

$$\forall x\; x \in Y \iff x \in X \wedge \varphi(x,p)$$

We denote Y by $\{x \in X \,:\, \varphi(x,p)\}$.

Proof. The existence follows by the Separation schema. The uniqueness is a consequence of the Extensionality axiom. $\square$

In the particular case where the formula and the parameter are such that they define a set without restricting the members to a given set, we may use the following shorter notation.

Definition 1.7
Let p be a set and $\varphi(x,p)$ a formula. If there is a set Y such that

$$\forall x\ x \in Y \Longleftrightarrow \varphi(x,p),$$

then Y is denoted by $\{x : \varphi(x,p)\}$. ■

Definition 1.8
Let X and Y be sets. The set $\{z \in X : z \in Y\}$ is called **intersection of X and Y**, written $X \cap Y$. ■

Remark 1.9
Let X and Y be sets. We clearly have

$$\forall z\ z \in X \cap Y \Longleftrightarrow z \in X \wedge z \in Y$$

Thus in Definition 1.8 the sets X and Y may be interchanged without changing the result for their intersection. ■

Definition 1.10
Let X and Y be two sets. X and Y are called **disjoint** if $X \cap Y = \emptyset$. Given a set Z, the members of Z are called **disjoint** if $x \cap y = \emptyset$ for every $x, y \in Z$. ■

Definition 1.11
Let X be a set. If $X \neq \emptyset$, then the set $\{y : \forall x \in X\ y \in x\}$ is called **intersection of X**, written $\bigcap X$. ■

In Definition 1.11 the short notation introduced in Definition 1.7 can be used since we may choose $z \in X$ such that $\bigcap X = \{y \in z : \forall x \in X\ y \in x\}$. As we have not proven so far that any other than the empty set exists, "$X \neq \emptyset$" is stated as a condition in Definition 1.11, which may in principle never be satisfied. As already mentioned, the Power set axiom as well as the Infinity axiom each guarantee (without the other one) the existence of a large number of sets.

Remark 1.12

The intersection of a set as defined in Definition 1.11 is sometimes generalized in the following way in the literature (see e.g. [Jech]):

> Let p be a set and $\varphi(X,p)$ a formula. If there exists a set X such that $\varphi(X,p)$ is true, then the set $Y = \{x : \forall X\ \varphi(X,p) \Longrightarrow x \in X\}$ is well-defined. If, in addition, there is a set Z such that
>
> $$\forall z\ z \in Z \Longleftrightarrow \varphi(z,p),$$
>
> then we have $Y = \bigcap Z$.

The last claim shows that Definition 1.11 is a special case of the first claim. Note that this generalization involves a formula variable, which we prefer to avoid. In this text the generalization of the intersection of a set is only used once, viz. in the definition of the natural numbers (Definition 1.43). Their existence is a consequence of the Separation schema with a concrete formula. ■

The following result states that there exists no set that contains all sets.

Lemma 1.13

We have

$$\forall X\ \exists x\ x \notin X$$

Proof. Let X be a set. Then we have $\{x \in X : x \notin x\} \notin X$. □

We now postulate that for two given sets X and Y there is a set that contains all elements of X and Y.

Axiom 1.14 (Small union)

$$\forall X\ \forall Y\ \exists Z\ \forall z\ (z \in X \vee z \in Y \Longrightarrow z \in Z)$$

■

Definition 1.15
Given two sets X and Y, the set $\{x \in X : x \notin Y\}$ is called **difference of** X and Y, written $X \setminus Y$. If $Y \subset X$, the set $X \setminus Y$ is also called **complement of** Y whenever the set X is evident from the context. The complement of Y is also denoted by Y^c. ■

Lemma and Definition 1.16
Let X and Y be two sets. Furthermore, let Z be a set such that $X, Y \subset Z$. The set $\big(X^c \cap Y^c\big)^c$, where the complement is with respect to Z, is called **union of** X **and** Y, written $X \cup Y$.

Proof. The existence of Z follows by the Small union axiom. To see that the definition of $X \cup Y$ is independent of the choice of Z, let W be another set with $X, Y \subset W$. We clearly have

$$Z \setminus \big((Z \setminus X) \cap (Z \setminus Y)\big) = W \setminus \big((W \setminus X) \cap (W \setminus Y)\big)$$

□

Remark 1.17
Let X and Y be sets. We have

$$\forall z \; z \in X \cup Y \iff \varphi(z, X, Y)$$

where $\varphi(z, X, Y)$ denotes the formula $(z \in X) \vee (z \in Y)$. Since this formula contains z and two parameters as free variables, we cannot use the Separation schema in the above form (i.e. Axiom 1.4) to define the union of X and Y. ■

In the following two Lemmas we list several important equalities that hold for the unions, intersections, and differences of two or three sets, and for the complement of subsets of a given set.

Lemma 1.18
Given three sets A, B, and C, the following equalities hold:

(i) $A \cup B = B \cup A, \quad A \cap B = B \cap A$

(ii) $A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C$

(iii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

(iv) $A \setminus (A \setminus B) = A \cap B$

Proof. Exercise. □

Lemma 1.19 (De Morgan)
Given a set X and $A, B \subset X$, the following equalities hold:

$$(A \cap B)^c = A^c \cup B^c, \quad (A \cup B)^c = A^c \cap B^c$$

where the complement is with respect to X in each case.

Proof. Exercise. □

Axiom 1.20 (Great union)

$$\forall X \; \exists Y \; \forall y \; \big(\exists x \; (x \in X \wedge y \in x) \Longrightarrow y \in Y\big)$$

■

Lemma and Definition 1.21
Let X be a set. If $X \neq \emptyset$, then the set $\{y : \exists x \; (x \in X \wedge y \in x)\}$ is called **union of** X, written $\bigcup X$.

Proof. This set is well-defined by the Great union axiom, the Separation schema, and the Extensionality axiom. □

In some contexts, in particular when considering unions and intersections of sets, a set is called a system, or a system of sets. Similarly a subset of a set is sometimes called a subsystem. This somewhat arbitrary change in nomenclature is motivated by the fact that in such contexts there intuitively seems to be a hierarchy of, first, the system, second, the members of the system, and, third, the elements of the members of the system. Often this intuitive hierarchy is highlighted by three different types of letters used for the variables, namely script letters for the system (e.g. $\mathcal{A}$), capital Latin letters for the members of the system (e.g. A), and small Latin letters for their elements (e.g a). We emphasize, however, that all objects denoted by these variables are nothing but sets, that is also systems are sets, and that two sets are distinct if an only if they have different elements.
We now postulate that the system of all subsets of a given set is contained in a set.

Axiom 1.22 (Power set)

$$\forall X \; \exists Y \; \forall y \; (y \subset X \Longrightarrow y \in Y)$$

■

Lemma and Definition 1.23
Given a set X, the set $\{y : y \subset X\}$ is called **power set of** X, written $\mathcal{P}(X)$. We also write $\mathcal{P}^2(X)$ for $\mathcal{P}(\mathcal{P}(X))$.

Proof. This set is well-defined by the Power set axiom, the Separation schema, and the Extensionality axiom. □

The existence of the power set allows the following variation of Definition 1.6.

Definition 1.24
Let X and p be sets and $\varphi(x, p)$ a formula. The set $\{x \in \mathcal{P}(X) : \varphi(x, p)\}$ is also denoted by $\{x \subset X : \varphi(x, p)\}$. ■

As a consequence of the Power set axiom, for every set X there exists a set that contains X and only X as element. We introduce the following notation and nomenclature.

Definition 1.25
For every set X, the set $\{Y \subset X : Y = X\}$ is denoted by $\{X\}$. ■

Definition 1.26
Let X be a set. X is called a **singleton** if there is a set x such that $X = \{x\}$. ■

With the Axioms postulated so far and without the Power set axiom we do not know about the existence of any other set than the empty set. Including the Power set axiom we conclude that also $\{\emptyset\}$ and $\{\{\emptyset\}\}$ are sets. Clearly these three sets are distinct by the Extensionality axiom.
The following Lemma and Definition states that for every two sets X and Y there is a set whose members are precisely X and Y.

Lemma and Definition 1.27
Given two sets X and Y, there is a set Z such that

$$\forall z \; z \in Z \iff z = X \vee z = Y$$

We also denote Z by $\{X, Y\}$.

Proof. Notice that $\{X\}$, $\{Y\}$ are sets by the Power set axiom. Let $Z = \{X\} \cup \{Y\}$. □

This shows that e.g. also $\{\emptyset, \{\emptyset\}\}$ is a set.

Remark 1.28
Given a set X, we have $\{X, X\} = \{X\}$. ■

Remark 1.29
Let X and Y be two sets. We have

(i) $X \cap Y = \bigcap \{X, Y\}$

(ii) $X \cup Y = \bigcup \{X, Y\}$

Thus Definition 1.8 and Lemma and Definition 1.16 can be considered as special cases of Definition 1.11 and Lemma and Definition 1.21, respectively. ■

For convenience we also introduce a notation for a set with three members.

Lemma and Definition 1.30
Given sets X, Y, and Z, there is a set U such that

$$\forall u \; u \in U \iff (u = X) \vee (u = Y) \vee (u = Z)$$

We also denote U by $\{X, Y, Z\}$.

Proof. We may define $U = \{X, Y\} \cup \{Z\}$. □

It is obvious that the order in which we write the sets X and Y in Lemma and Definition 1.27, or the order in which we write X, Y, and Z in Lemma and Definition 1.30 does not play a role. In many contexts we need a system that identifies two (distinct or equal) sets and also specifies their order. This is achieved by the following concept.

Lemma and Definition 1.31
Let X be a set. X is called **ordered pair**, or short **pair**, if there are sets x and y such that $X = \{\{x, y\}, \{x\}\}$. In this case we have $\bigcup \bigcap X = x$ and $\bigcup \left(\bigcup X \setminus \{x\}\right) = y$. Moreover, in this case x and y are called **left** and **right coordinates** of X, respectively. Further, if X is an ordered pair, its unique left and right coordinates are denoted by X_l and X_r, respectively, and X is denoted by (X_l, X_r).

Proof. Exercise. □

Remark 1.32
Let x and y be two sets. Then $x \neq y$ implies $(x, y) \neq (y, x)$. ■

The concept of ordered pair allows the extension of the Separation schema, Axiom 1.4, to more than one parameter. The following is the result for two parameters.

Lemma 1.33 (Separation schema with two parameters)
Let $\varphi(x, p, q)$ be a formula. We have:

$$\forall p \,\forall q \,\forall X \,\exists Y \,\forall x \,\big(x \in Y \Longleftrightarrow x \in X \wedge \varphi(x, p, q)\big)$$

Proof. Let p, q, and X be sets. We define $r = (p, q)$ and $Y = \{x \in X : \psi(x, r)\}$ where $\psi(x, r)$ is the formula $\exists u \,\exists v \; r = (u, v) \wedge \varphi(x, u, v)$. Then Y is the required set. □

The Separation schema is used in the following two Lemmas. Before stating these Lemmas we would like to relax the rules for the usage of the set brackets $\{\ldots\}$ that are defined in Definition 1.6. Remember that such a modified notation is already defined in the case where the formula specifies a set (cf. Definition 1.7) and in the case of a subset of the power set (cf. Definition 1.24). We now agree that we may use a comma instead of $\wedge$ between two or more formulae on the right hand side of the colon; e.g. given two sets X and Y, we may write $\{x : x \in X,\ x \in Y\}$ instead of $\{x : x \in X \wedge x \in Y\}$. Moreover we agree that we may use all defined symbols on the left hand side of the colon, thereby eliminating one or more $\exists$ and one equality on the right hand side; e.g. given two sets X and Y, we may write $\{x \cap y : x \in X,\ x \in Y\}$ instead of $\{z : \exists x \in X \,\exists x \in Y \; z = x \cap y\}$, and $\{Y \setminus x : x \in X\}$ instead of $\{z : \exists x \in X \; z = Y \setminus x\}$. This is precisely the same kind of notation as in Definition 1.24.

Lemma 1.34
Let $\mathcal{A}$ and $\mathcal{B}$ be two systems of sets. If $\mathcal{A} \neq \emptyset$ and $\mathcal{B} \neq \emptyset$, the following equalities hold:

(i) $\big(\bigcup \mathcal{A}\big) \cap \big(\bigcup \mathcal{B}\big) = \bigcup \{A \cap B \,:\, A \in \mathcal{A},\, B \in \mathcal{B}\}$

(ii) $\big(\bigcap \mathcal{A}\big) \cup \big(\bigcap \mathcal{B}\big) = \bigcap \{A \cup B \,:\, A \in \mathcal{A},\, B \in \mathcal{B}\}$

In particular, the right hand sides of (i) and (ii) are well-defined.

Proof. Exercise. □

Lemma 1.35 (De Morgan)
Given a set X and a system $\mathcal{A} \subset \mathcal{P}(X)$ with $\mathcal{A} \neq \emptyset$, the following equalities hold:

(i) $\big(\bigcap \mathcal{A}\big)^c = \bigcup \{A^c \,:\, A \in \mathcal{A}\}$

(ii) $\big(\bigcup \mathcal{A}\big)^c = \bigcap \{A^c \,:\, A \in \mathcal{A}\}$

where the complements refer to the set X. In particular, the right hand sides of (i) and (ii) are well-defined.

Proof. Exercise. □

Remark 1.36
Notice that the equalities (iii) in Lemma 1.18 are special cases of those in Lemma 1.34, and that the equalities in Lemma 1.19 are special cases of those in Lemma 1.35. ■

We now introduce the set of all ordered pairs such that the left coordinate is in X and the right coordinate is in Y where X and Y are two sets. The following Definition uses again the Separation schema with two parameters.

Definition 1.37
Let X and Y be two sets. The set

$$\left\{z \in \mathcal{P}^2(X \cup Y) \,:\, \exists x \in X \; \exists y \in Y \; z = \{\{x, y\}, \{x\}\}\right\}$$

is called **Cartesian product of** X and Y, and denoted by $X \times Y$. ■

We now state some properties of the Cartesian product of two sets.

Lemma 1.38
Let U, V, X, and Y be sets. Then we have

(i) $\emptyset \times V = U \times \emptyset = \emptyset$

(ii) $(U \times V) \cap (X \times Y) = (U \cap X) \times (V \cap Y)$

(iii) $X \times (V \cup Y) = (X \times V) \cup (X \times Y)$

(iv) $X \times (Y \setminus V) = (X \times Y) \setminus (X \times V)$

Proof. Exercise. □

Lemma 1.39
Let U, V, X, and Y be sets with $X \subset U$ and $Y \subset V$. Then we have

$$(X \times Y)^c = (X^c \times Y^c) \cup (X^c \times Y) \cup (X \times Y^c)$$

where the first complement refers to $U \times V$, the complement of X refers to U, and the complement of Y refers to V.

Proof. Exercise. □

Similarly to ordered pairs we introduce the notion of ordered triple consisting of three sets in a specific order.

Definition 1.40
Let X be a set. X is called **ordered triple** if there are sets x, y, and z such that $X = \big((x,y),z\big)$. In this case X is also denoted by (x,y,z). ■

Notice that the Separation schema may be extended to three parameters by using ordered triples. Similarly, we may clearly define ordered quadruples etc. and define corresponding Separation schemas.

Axiom 1.41 (Infinity)

$$\exists X \ \Big(\emptyset \in X \wedge \forall x \ \big(x \in X \Longrightarrow x \cup \{x\} \in X\big)\Big)$$

■

Definition 1.42
A set X is called **inductive** if it has the following properties:

(i) $\emptyset \in X$

(ii) $\forall x \in X \quad x \cup \{x\} \in X$

■

Obviously, the Infinity axiom says that there exists an inductive set.

Definition 1.43
Let X be an inductive set. The members of the set

$$\{n \in X \,:\, \forall Y\, Y \text{ is inductive} \Longrightarrow n \in Y\}$$

are called **natural numbers**. The set of natural numbers is denoted by $\mathbb{N}$. ■

Notice that Definition 1.43 does not depend on the choice of the inductive set X. This is an example of the concept discussed in Remark 1.12, which can be understood as a generalized form of intersection.

Remark 1.44
The set $\mathbb{N}$ is inductive. ■

The following Axiom is part of ZFC, essentially in order to obtain the statements in the following Lemma.

Axiom 1.45 (Regularity)

$$\forall X\ X \neq \emptyset \Longrightarrow \exists x\ \ (x \in X) \wedge (X \cap x = \emptyset)$$

■

Lemma 1.46
The following statements hold:

(i) $\neg \exists X\ X \in X$

(ii) $\neg \exists X\ \exists Y\ \ (X \in Y) \wedge (Y \in X)$

(iii) $\neg \exists X\ \exists Y\ \exists Z\ \ (X \in Y) \wedge (Y \in Z) \wedge (Z \in X)$

Proof. To see (i), let X be a set. Notice that $\{X\}$ is non-empty and thus $\{X\} \cap X = \emptyset$ by the Regularity axiom.
To see (ii), we assume that X and Y are two sets such that $X \in Y$ and $Y \in X$. Since the set $Z = \{X, Y\}$ is non-empty there is $z \in Z$ such that $Z \cap z = \emptyset$ by the Regularity axiom. However, $z = X$ implies $Z \cap X = Y$, and $z = Y$ implies $Z \cap Y = X$, which is a contradiction.
Finally, to show (iii), we assume that there are sets X, Y, and Z such that $X \in Y$, $Y \in Z$, and $Z \in X$. We define $W = \{X, Y, Z\}$. Application of the Regularity axiom again leads to a contradiction. □

Axiom 1.47 (Replacement schema)
Let $\varphi(x, y, p)$ be a formula.

$$\begin{aligned}\forall p \Big(\big(\forall x \, \forall y \, \forall z \, (\varphi(x, y, p) \wedge \varphi(x, z, p) \Longrightarrow y = z)\big) \\ \Longrightarrow \forall X \, \exists Y \, \forall y \, \big(\exists x \, (x \in X \wedge \varphi(x, y, p)) \Longrightarrow y \in Y\big)\Big)\end{aligned}$$

■

Notice that Axiom 1.47 is not a single axiom but a schema of axioms. This concept is discussed above in the context of the Separation schema, Axiom 1.4. The Replacement schema is applied below to derive Theorem 3.49 where a concrete formula is substituted.
The premise in Axiom 1.47 says that for every x there is at most one y such that $\varphi(x, y, p)$ is satisfied. The conclusion states that if the sets x are taken out of a given set X, there exists a set Y that has those sets y among its members. Clearly one may also define a set Y that has precisely those sets y as its members (and no others) by the Separation schema. This is the result of the following Lemma.
We remark that, as in the context of the Separation schema (cf. Lemma 1.33),

the Replacement schema can be extended to two or more parameters. However, this is not required in the present work.

Lemma 1.48
Let X and p be sets, and $\varphi(x, y, p)$ a formula such that

$$\forall x\ \forall y\ \forall z\ \varphi(x, y, p) \wedge \varphi(x, z, p) \Longrightarrow y = z$$

holds. Then the set $\{y : \exists x \in X\ \varphi(x, y, p)\}$ is well-defined.

Proof. The existence follows by the Replacement schema and the Separation schema. The uniqueness is a consequence of the Extensionality axiom. $\square$

Axiom 1.49 (Choice)

$$\begin{aligned}\forall X\ \exists Z\ \forall Y \in \mathcal{P}(X)\backslash\{\emptyset\}\ \Big(&\big(\exists y \in Y\ \ (Y, y) \in Z\big)\\ &\wedge \big(\forall u, v \in X\ \ (Y, u), (Y, v) \in Z \Longrightarrow u = v\big)\Big)\end{aligned}$$

$\blacksquare$

Lemma and Definition 1.50
Let X be a set and $Y \subset \mathcal{P}(X)$ with $\emptyset \notin Y$. There exists a set $Z \subset Y \times X$ such that the following statements hold:

(i) $\forall z \in Z \quad z_r \in z_l$

(ii) $\forall y \in Y \quad \exists z \in Z \quad z_l = y$

(iii) $\forall w, z \in Z \quad (w_l = z_l \Longrightarrow w_r = z_r)$

Z is called a **choice function**.

Proof. Given the stated conditions, there exists, by Axiom 1.49, a set W such that $Z = W \cap (Y \times X)$ satisfies (i) to (iii). □

Given the other axioms above, the Choice axiom can be stated in several equivalent forms. Two versions are presented in Section 3.3, viz. the Well-ordering principle and Zorn's Lemma.

We finally introduce a notation that is convenient when dealing with topologies, topological bases and other concepts to be introduced below.

Definition 1.51

Let X be a set, $\mathcal{A} \subset \mathcal{P}(X)$, and $x \in X$. We define $\mathcal{A}(x) = \{A \in \mathcal{A} : x \in A\}$. ■

Notice that the existence of sets, or even of a single set, is not guaranteed if the Existence axiom is not postulated. This is because clearly none of the other axioms postulates the existence of a set—without any other set already existing—apart from the Infinity axiom 1.41 below. There however the definition of the empty set is used, which in turn is defined in Lemma and Definition 1.5 by usage of the Separation schema. The Separation schema always refers to an existing set. It is possible to modify the Infinity axiom such that it also postulates the existence of a set, see e.g. [Jech].

Chapter 2

Relations

2.1 Relations and orderings

In this Section we introduce the concept of relation, which is fundamental in the remainder of the text. Many important special cases are analysed, in particular orderings. Also functions, that are introduced in the next Section, are relations.

Definition 2.1
Given two sets X and Y, a subset $U \subset X \times Y$ is called a **relation on** $X \times Y$. The **inverse of** U is a relation on $Y \times X$ and defined as

$$U^{-1} = \big\{(y,x) \in Y \times X \,:\, (x,y) \in U\big\}$$

Given another set Z and a relation $V \subset Y \times Z$, the **product of** V **and** U is defined as

$$VU = \big\{(x,z) \in X \times Z \,:\, \exists y \in Y \ (x,y) \in U,\ (y,z) \in V\big\}$$

A relation R on $X \times X$ is also called **relation on** X. In this case the pair (X,R) is called **relational space**. Furthermore, the set $\Delta = \{(x,x) \,:\, x \in X\}$ is called **diagonal**. ■

Notice the order of U and V in the definition of the product, which may be counterintuitive.

Definition 2.2
Given two sets X and Y, a relation $R \subset X \times Y$, and a set $A \subset X$ we introduce the following notation:

(i) $R[A] = \big\{y \in X \,:\, \exists x \in A \; (x,y) \in R\big\}$

(ii) $R\{x\} = R[\{x\}]$, that is $R\{x\} = \{y \in X \,:\, (x,y) \in R\}$

(iii) $R\langle A\rangle = \big\{y \in X \,:\, \forall x \in A \; (x,y) \in R\big\}$

The set $R[X]$ is called the **range of** R, written $\operatorname{ran}(R)$ or $\operatorname{ran} R$. The set $R^{-1}[Y]$ is called the **domain of** R, written $\operatorname{dom}(R)$ or $\operatorname{dom} R$. We say that R **has full range** if $\operatorname{ran} R = Y$, and **full domain** if $\operatorname{dom} R = X$.
Given a relation S on X, the set $\big(S \cup S^{-1}\big)[X] = (\operatorname{dom} S) \cup (\operatorname{ran} S)$ is called the **field of** S, written $\operatorname{field}(S)$ or $\operatorname{field} S$. We say that S **has full field** if $\operatorname{field} S = X$. ■

Clearly, a relation S on X that has full domain or full range also has full field. Notice that Definition 2.1 of a relation R on a Cartesian product $X \times Y$ specifies the sets X and Y from which the product is formed although X may not be the domain and Y may not be the range of R. This is important in the context of functional relations to be defined in Section 2.2.

Definition 2.3
Let $\mathcal{R}$ be a system of relations on $X \times Y$. We define

(i) $\mathcal{R}[A] = \{R[A] : R \in \mathcal{R}\}$

(ii) $\mathcal{R}\{x\} = \mathcal{R}[\{x\}]$, that is $\mathcal{R}\{x\} = \{R\{x\} : R \in \mathcal{R}\}$

The sets

$$\bigcup(\mathcal{R}[X]) = \bigcup\{R[X] : R \in \mathcal{R}\}, \qquad \bigcup\left\{R^{-1}[Y] : R \in \mathcal{R}\right\}$$

are called **range** and **domain of** $\mathcal{R}$, respectively.
If $X = Y$, then the set

$$\bigcup\left\{R[X] : R, R^{-1} \in \mathcal{R}\right\}$$

is called **field of** $\mathcal{R}$. In this case $\mathcal{R}$ is said to **have full field** if its field is X.
■

Remark 2.4
Given two sets X and Y, $A \subset X$, and a relation $R \subset X \times Y$, the following statements hold:

(i) $R\{x\} = R\langle\{x\}\rangle$ for every $x \in X$

(ii) $R[\varnothing] = \varnothing$

(iii) $R\langle\varnothing\rangle = Y$

(iv) $R[A] = \bigcup\{R\{x\} : x \in A\}$

(v) $R\langle A\rangle = \bigcap\{R\{x\} : x \in A\}$

■

Definition 2.5
Let X and Y be sets, $R \subset X \times Y$ a relation, and $\mathcal{A} \subset \mathcal{P}(X)$. We define

$$R [\![\mathcal{A}]\!] = \{R[A] : A \in \mathcal{A}\}$$

■

We now list a few consequences of the above definitions.

Lemma 2.6
Given sets X, Y and Z, and relations $U, U' \subset X \times Y$, and $V, V' \subset Y \times Z$ where $U' \subset U$ and $V' \subset V$, and $W \subset Z \times S$ the following statements hold:

(i) $(VU)^{-1} = U^{-1}V^{-1}$

(ii) $(WV)U = W(VU)$

(iii) $U'^{-1} \subset U^{-1}$

(iv) $V'U \subset VU$

(v) $VU' \subset VU$

Proof. Exercise. □

By Lemma 2.6 (ii) we may drop the brackets in the case of multiple products of relations without generating ambiguities.

Lemma 2.7
Given sets X, Y, and Z, relations $U \subset X \times Y$ and $V \subset Y \times Z$, and a set $A \subset X$, we have $(VU)[A] = V[U[A]]$.

Proof. Exercise. □

Definition 2.8
Let (X, R) be a relational space. Then the relation $R \,|\, A = R \cap (A \times A)$ on A is called the **restriction of** R **to** A. ■

The following properties are important to characterize different types of relational spaces.

Definition 2.9
Let (X, R) be a relational space. Then R is called

(i) **reflexive** if $\Delta \subset R$

(ii) **antireflexive** if $\Delta \cap R = \emptyset$

(iii) **symmetric** if $R^{-1} = R$

(iv) **antisymmetric** if $R \cap R^{-1} \subset \Delta$

(v) **transitive** if $R^2 \subset R$

(vi) **connective** if $R \cup R^{-1} \cup \Delta = X \times X$

(vii) **directive** if $X \times X = R^{-1} R$

■

We remark that antisymmetry of a relation is defined in different ways in the literature, see for example [Gaal], p. 6, where the definition is $R \cap R^{-1} = \emptyset$, or [von Querenburg], p. 4, where the definition is $R \cap R^{-1} = \Delta$. The expressions "connective" and "directive" are not standard terms in the literature, see however [Ebbinghaus], p. 58.

Remark 2.10

Let (X, R) be a relational space and $A \subset X$. The following statements hold:

(i) R is connective iff for every $x, y \in X$ we have $(x, y) \in R$ or $(y, x) \in R$ or $x = y$.

(ii) R is directive iff for every $x, y \in X$ there is $z \in X$ such that $(x, z), (y, z) \in R$.

(iii) R^{-1} is reflexive, antireflexive, symmetric, antisymmetric, transitive, or connective if R has the respective property.

(iv) $R \,|\, A$ is reflexive, antireflexive, symmetric, antisymmetric, transitive, or connective if R has the respective property.

■

Lemma 2.11

Given a set X and relations U, V on X where V is symmetric, we have

$$VUV = \bigcup \{\, (V\{x\}) \times (V\{y\}) \;:\; (x, y) \in U \}$$

Proof. If $(u, v) \in VUV$, then there exists $(x, y) \in U$ such that $(u, x), (y, v) \in V$. Therefore we have $(u, v) \in (V\{x\}) \times (V\{y\})$. The converse is shown in a similar way. □

Definition 2.12

Let (X, R) be a related space and $A \subset X$. A is called a **chain** if $R \,|\, A$ is connective. For definiteness, we also define $\emptyset$ to be a chain. ■

Definition 2.13

Let X be a set and $\mathcal{A} \subset \mathcal{P}(X)$. $\mathcal{A}$ is called a **partition of** X if $\bigcup \mathcal{A} = X$ and $A \cap B = \emptyset$ for every $A, B \in \mathcal{A}$. ■

Definition 2.14

Let (X, R) be a relational space. R is called **equivalence relation** if it is reflexive, symmetric, and transitive. Given a point $x \in X$, the set $R\{x\}$ is called **equivalence class of** x, written $[x]$. ■

Remark 2.15

Let X be a set, R an equivalence relation on X, and $x, y \in X$. The following equivalences hold:

$$(x, y) \in R \quad \iff \quad x \in [y] \quad \iff \quad y \in [x] \quad \iff \quad [x] = [y]$$

■

Lemma 2.16

Given a set X and an equivalence relation R on X, the system of all equivalence classes is a partition of X, denoted by X/R.

Proof. We clearly have $\bigcup X/R = X$. Now we assume that $u, x, y \in X$ such that $u \in [x] \cap [y]$. It follows that $[x] = [u] = [y]$ by Remark 2.15. □

Definition 2.17
Let (X, R) be a relational space.

(i) R is called a **pre-ordering on** X if it is transitive. In this case we also write $\prec$ for R, and $x \prec y$ for $(x, y) \in R$. The pair $(X, \prec)$ is called **pre-ordered space**.

(ii) R is called an **ordering on** X if it is transitive and antisymmetric. The pair (X, R) is called **ordered space**.

(iii) R is called an **ordering in the sense of** "$<$" **on** X if it is antireflexive and transitive. In this case we also write $<$ for R, and $x < y$ or $y > x$ for $(x, y) \in R$. The pair $(X, <)$ is called **space ordered in the sense of** "$<$".

(iv) R is called an **ordering in the sense of** "$\leq$" **on** X if it is reflexive, antisymmetric, and transitive. In this case we also write $\leq$ for R, and $x \leq y$ or $y \geq x$ for $(x, y) \in R$. The pair $(X, \leq)$ is called **space ordered in the sense of** "$\leq$".

(v) R is called a **direction on** X if it is transitive, reflexive and directive. In this case we also write $\leq$ for R, and $x \leq y$ for $(x, y) \in R$. The pair $(X, \leq)$ is called **directed space**.

■

It follows by Definition 2.17 that there are no $x, y \in X$ such that both $x < y$ and $y < x$. Hence an ordering in the sense of "$<$" is antisymmetric. Therefore orderings in the sense of "$<$" and orderings in the sense of "$\leq$" are both orderings. We remark that subsequently the symbol $\prec$ is used only for pre-orderings, $<$ and $>$ only for orderings in the sense of "$<$", $\leq$ only for orderings in the sense of "$\leq$" and for directions, and $\geq$ only for orderings in the sense of "$\leq$". Regarding the symbol $\leq$ it is clarified in each case which kind of relation is considered. Of

course, any of such relations may have additional properties and are still denoted by the same symbol. For example, an ordering may be denoted by $\prec$ because it is a pre-ordering. Even an ordering in the sense of "$\leq$" or one in the sense of "$<$" may be denoted by $\prec$ in certain cases. Thus by using the symbols we implicitly imply that the relation satisfies certain properties but we do not exclude that it satisfies more.
We deliberately use the aggregated notation $x \prec y \prec z$ instead of "$x \prec y$ and $y \prec z$", and similarly for the other two symbols.

Remark 2.18
Let (X, R) be a relational space and $A \subset X$. R^{-1} and $R\,|\,A$ are pre-orderings, orderings, orderings in the sense of "$<$", or orderings in the sense of "$\leq$", if R has the respective property. ■

Lemma 2.19
Let (X, R) be an ordered space. Then $S = R \cup \Delta$ is an ordering in the sense of $\leq$, and $T = R \setminus \Delta$ is an ordering in the sense of $<$.

Proof. Exercise. □

Definition 2.20
Let $(X, \prec)$ be a pre-ordered space and $x \in X$. A point $y \in X$ is called **successor of** x if $x \prec y$, $x \neq y$, and if there is no $z \in X \setminus \{x, y\}$ such that $x \prec z \prec y$. A point $y \in X$ is called **predecessor of** x if $y \prec x$, $x \neq y$, and if there is no $z \in X \setminus \{x, y\}$ such that $y \prec z \prec x$. ■

It is clear that generally a successor or a predecessor of a point $x \in X$ need not exist and if it exists, need not be unique. Obviously, if y is a successor of x, then x is a predecessor of y.

Example 2.21
Let $X = \{a, b, c\}$ and $R = \{(a, b), (b, c), (a, c), (b, b)\}$. Then R is an ordering on X, $R \setminus \{(b, b)\}$ is an ordering in the sense of "$<$", and $R \cup \{(a, a), (c, c)\}$ is an ordering in the sense of "$\leq$". ■

Given a relational space, one can construct a pre-ordered space such that the set remains the same and the pre-ordering contains the original relation as subset. However this requires a recursive definition and is therefore postponed until Section 3.1.
Given a pre-ordered space, one can construct an ordered space by an antisymmetrization procedure as follows.

Lemma 2.22
Let (X, R) be a pre-ordered space and Q a relation on X defined by

$$(x, y) \in Q \quad \Longleftrightarrow \quad (x, y), (y, x) \in R \ \wedge \ x = y$$

Then Q is an equivalence relation. Let $S \subset (X/Q) \times (X/Q)$ be the relation defined by

$$(s, t) \in S \quad \Longleftrightarrow \quad \exists x \in s,\, y \in t \ \ (x, y) \in R$$

Then $(X/Q, S)$ is an ordered space. If R is reflexive, then S is reflexive.

Proof. Exercise. □

Lemma 2.22 is used in Theorem 3.56. The analysis of the equivalence relation Q also enhances our understanding how a general pre-ordered space, i.e. a set with a relation satisfying transitivity, looks like. It shows that there may be disjoint groups each consisting of several elements of X and "isolated elements" in the following sense. For every pair of distinct elements x and y within the same group we have $(x,y),(y,x) \in R$ and thus also $(x,x),(y,y) \in R$ by transitivity. For an "isolated element" x we may have $(x,x) \in R$ or $(x,x) \notin R$. The relation S on X/Q always leads to $(s,s) \in S$ if s corresponds to a group of elements, and it may lead to $(s,s) \in S$ or to $(s,s) \notin S$ for "isolated elements" depending on which statement holds for the original elements of X. Therefore the ordering S need not be in the sense of "$\leq$" nor in the sense of "$<$". However, we have already examined another method how to construct such orderings from arbitrary orderings in Lemma 2.19.

Lemma and Definition 2.23

Let X be a set and $\mathcal{A} \subset \mathcal{P}(X)$. Let the relation $R \subset \mathcal{A} \times \mathcal{A}$ be defined by $(A,B) \in R$ if $A \subset B$. We also write $(\mathcal{A}, \subset)$ for $(\mathcal{A}, R)$ and $(\mathcal{A}, \supset)$ for $\big(\mathcal{A}, R^{-1}\big)$. Each of the pairs $(\mathcal{A}, \subset)$ and $(\mathcal{A}, \supset)$ is an ordered space in the sense of "$\leq$", and if $\mathcal{A} = \mathcal{P}(X)$, a directed space.

Proof. Exercise. $\square$

Definition 2.24

Given a set X, a connective ordering $\prec$ on X is called **total ordering**. In this case $(X, \prec)$ is called **totally ordered space**. $\blacksquare$

Example 2.25

Let X be a set. If X has more than one member, then $(\mathcal{P}(X), \subset)$ is not a totally ordered space. $\blacksquare$

Definition 2.26
Let $(X, \prec)$ be a pre-ordered space. For every $x, y \in X$ with $x \prec y$ the set $]x, y[\ = \{z \in X \,:\, x \prec z \prec y\}$ is called **proper interval**. Moreover, for every $x \in X$, the set $]-\infty, x[\ = \{z \in X \,:\, z \prec x\}$ is called the **lower segment of** x, and the set $]x, \infty[\ = \{z \in X \,:\, x \prec z\}$ is called the **upper segment of** x. A lower or upper segment is also called an **improper interval**. A proper or improper interval is also called an **interval**. ■

Clearly, if $x \prec y$, then $]x, y[\ =\ ]-\infty, y[\ \cap\ ]x, \infty[$. We remark that ∞ and $-\infty$ are merely used as symbols here. In particular, they do not generally refer to any of the number systems to be introduced below in this Chapter, neither does their usage imply that there is an infinite number of elements—for a Definition of "infinite" see Section 3.4 below—in an improper interval.

Definition 2.27
Let X be a set and $\mathcal{R} = \{R_i \,:\, i \in I\}$ a system of pre-orderings on X. Intervals with respect to a pre-ordering $R \in \mathcal{R}$ are denoted by subscript R, i.e. $]-\infty, x[_R$ and $]x, \infty[_R$ where $x \in X$, and $]x, y[_R$ where $x, y \in X$, $(x, y) \in R$. Alternatively intervals with respect to R_i for some $i \in I$ are denoted by index i, i.e. $]-\infty, x[_i$, etc. ■

Remark 2.28
Let $(X, \prec)$ be a pre-ordered space and $x \in X$.

(i) If $\prec$ has full range, then $]-\infty, x[\ = \bigcup \{\ ]y, x[\ :\ y \in X,\ y \prec x\}$

(ii) If $\prec$ has full domain, then $]x, \infty[\ = \bigcup \{\ ]x, y[\ :\ y \in X,\ x \prec y\}$

■

Definition 2.29

Let $(X, \prec)$ be a pre-ordered space. A subset $Y \subset X$ is called **$\prec$-dense in** X or **order dense in** X if for every $x, y \in X$ with $x \prec y$ there exists $z \in Y$ such that $x \prec z \prec y$. X is called **$\prec$-dense** or **order dense** if it is order dense in itself. ■

Definition 2.30

Let X be a set and $\mathcal{R}$ a system of pre-orderings on X. A subset $Y \subset X$ is called **$\mathcal{R}$-dense in** X if for every $R \in \mathcal{R}$ and $x, y \in X$ with $(x, y) \in R$ there exists $z \in Y$ such that $(x, z), (z, y) \in R$. X is called **$\mathcal{R}$-dense** if it is $\mathcal{R}$-dense in itself. ■

Remark 2.31

Let $(X, \prec)$ be a pre-ordered space and $Y \subset X$ order dense. For every $x, y \in X$ the following equalities hold:

$$\begin{aligned}]-\infty, y[\;&=\; \bigcup \{\,]-\infty, z[\;:\; z \in Y,\; z \prec y \} \\ ]x, \infty[\;&=\; \bigcup \{\,]z, \infty[\;:\; z \in Y,\; x \prec z \} \\ ]x, y[\;&=\; \bigcup \{\,]u, v[\;:\; u, v \in Y,\; x \prec u \prec v \prec y \} \end{aligned}$$

■

Definition 2.32

Let (X, R) be a relational space. A member $x \in X$ is called a **weak minimum of** X if $(y, x) \in R$ implies $(x, y) \in R$. Moreover a member $x \in X$ is called a **weak maximum of** X if $(x, y) \in R$ implies $(y, x) \in R$. Further let $A \subset X$. Then $x \in A$ is called a **weak minimum** (**weak maximum**) **of** A if it is a weak minimum (weak maximum) of A with respect to the restriction $R \,|\, A$. ■

Definition 2.33
Let (X, R) be a relational space. A member $x \in X$ is called a **minimum** or **least element of** X if $(x, y) \in R$ for every $y \in X \setminus \{x\}$. Moreover a member $x \in X$ is called a **maximum** or **greatest element of** X if $(y, x) \in R$ for every $y \in X \setminus \{x\}$. Further let $A \subset X$. Then $x \in A$ is called a **minimum** (**maximum**) **of** A if it is a minimum (maximum) of A with respect to the restriction $R \,|\, A$. If A has a unique minimum (maximum), then it is denoted by $\min A$ ($\max A$). ■

Notice that the singleton $\{x\}$, where x is a set, trivially has x as its minimum and maximum. Although Definitions 2.32 and 2.33 are valid for any relation R on X, they are mainly relevant in the case where R is a pre-ordering.
The following result shows that the notions defined in Definition 2.33 are invariant under a change from the original relation to the relations defined in Lemma 2.19.

Lemma 2.34
Let (X, R) be a relational space, $T \in \{R \cup \Delta,\ R \setminus \Delta\}$, and $x \in X$. x is a minimum (maximum) of X with respect to T iff it is a minimum (maximum) of X with respect to R.

Proof. Exercise. □

Remark 2.35
Let (X, R) be a relational space. If $x \in X$ is a minimum (maximum) of X, then x is also a weak minimum (weak maximum) of X. ■

Remark 2.36
Let $(X, <)$ be an ordered space and x a weak minimum of X. Then there is no $y \in X$ with $y < x$. ■

Remark 2.37

Let $(X, \prec)$ be an ordered space. If X has a minimum (maximum), then this minimum (maximum) is unique. ■

Remark 2.38

Let $(X, \prec)$ be a totally ordered space. If X has a weak minimum (weak maximum), then this weak minimum (weak maximum) is the minimum (maximum) of X. ■

Definition 2.39

Let (X, R) be a relational space. We say that R **has the minimum property** if every $A \subset X$ with $A \neq \emptyset$ has a minimum. ■

Remark 2.40

Let (X, R) be a relational space. If R has the minimum property, then R is connective. ■

Definition 2.41

Given a set X, an ordering R on X that has the minimum property is called **well-ordering**. In this case we say that R **well-orders** X, and (X, R) is called **well-ordered space**. ■

Notice that according to Definition 2.41 a well-ordering may be an ordering in the sense of "$<$" or "$\leq$" or neither. In the literature "well-ordering" is often used only in the sense of "$<$" (see e.g. [Kelley] or [Ebbinghaus]) or only in the sense of "$\leq$" (see e.g. [von Querenburg]).

Lemma 2.42

Every well-ordered space is totally ordered.

Proof. This follows from Remark 2.40. □

Lemma 2.43
Let (X, R) be a well-ordered space. Then $S = R \cup \Delta$ is a well-ordering in the sense of "$\leq$", and $T = R \setminus \Delta$ is a well-ordering in the sense of "$<$".

Proof. S and T are clearly well-orderings. The claim follows by Lemma 2.19. □

Lemma 2.44
Let (X, R) be a relational space. If R is antisymmetric and has the minimum property, then it is a well-ordering.

Proof. R is connective by Lemma 2.40. Now assume that R is not transitive. Let $x, y, z \in X$ such that $(x, y), (y, z) \in R$ and $(x, z) \notin R$. If all three or any two of the elements of $\{x, y, z\}$ are equal, then this is a contradiction. If x, y, and z are distinct, then we have $(z, x) \in R$ since R is connective, and $(y, x), (z, y) \notin R$ since R is antisymmetric. Thus $\{x, y, z\}$ has no minimum, which is a contradiction. □

Definition 2.45
Let $(X, \prec)$ be a pre-ordered space and $A \subset X$. A member $x \in X$ is called an **upper bound of** A if $y \prec x$ for every $y \in A \setminus \{x\}$. A member $x \in X$ is called a **lower bound of** A if $x \prec y$ for every $y \in A \setminus \{x\}$. A member $x \in X$ is called a **supremum of** A or a **least upper bound of** A if it is a minimum of the set of all upper bounds of A. A member $x \in X$ is called an **infimum of** A or **a greatest lower bound of** A if it is a maximum of the set of all lower bounds of A. If A has a unique supremum (infimum), then it is denoted by $\sup A$ ($\inf A$).
■

Again, the notions defined in Definition 2.45 are invariant under a change from the original relation to the relations defined in Lemma 2.19.

Lemma 2.46

Let (X, R) be a relational space, $T \in \{R \cup \Delta,\ R \backslash \Delta\}$, $A \subset X$, and $x \in X$. x is an upper bound, lower bound, supremum, or infimum of A with respect to T iff it has the respective property with respect to R.

Proof. Exercise. □

Lemma and Definition 2.47

Let $(X, \prec)$ be an ordered space. Every $A \subset X$ has at most one supremum and at most one infimum. Given a set Y, a subset $B \subset Y$, and a function $f : Y \longrightarrow X$, the supremum of the set $f[B] = \{f(y) : y \in B\}$ is also denoted by $\sup_{y \in B} f(y)$ and its infimum by $\inf_{y \in B} f(y)$.

Proof. Let C be the set of all upper bounds of A. If C has a minimum, then this minimum is unique by Remarks 2.18 and 2.37. Therefore A has at most one supremum. The proof regarding the minimum is similar. □

The following is a property that, for example, the real numbers have as demonstrated in Lemma 4.39 below.

Definition 2.48

Let $(X, \prec)$ be a pre-ordered space. We say that $\prec$ **has the least upper bound property** if every set $A \subset X$ with $A \neq \emptyset$ which has an upper bound has a supremum. ■

The least upper bound property is equivalent to the intuitively reversed property as stated in the following Theorem. In the proof we follow [Kelley], p. 14.

Theorem 2.49
Let $(X, \prec)$ be a pre-ordered space. $\prec$ has the least upper bound property iff every set $A \subset X$ with $A \neq \emptyset$ which has a lower bound has an infimum.

Proof. First assume that $\prec$ has the least upper bound property. Let $A \subset X$ such that $A \neq \emptyset$ and A has a lower bound. Further let B be the set of all lower bounds of A. Let $x \in A$. It follows that, for every $y \in B$, we have $y = x$ or $y \prec x$. Hence x is an upper bound of B. Therefore all members of A are upper bounds of B. By assumption B has a supremum, say y. Since y is the minimum of all upper bounds of B, we have $y \prec x$ for every $x \in A \backslash \{y\}$. Thus y is a lower bound of A. In order to see that y is the greatest lower bound of A, let z be a lower bound of A, i.e. $z \in B$. Since y is an upper bound of B we have $z \prec y$ or $z = y$.
The converse can be proven similarly. □

Example 2.50
Let X be a set. Then $(\mathcal{P}(X), \subset)$ is an ordered space. Further let $\mathcal{A} \subset \mathcal{P}(X)$. Then X is an upper bound of $\mathcal{A}$ and $\emptyset$ is a lower bound of $\mathcal{A}$. If $\mathcal{A} \neq \emptyset$, then $\bigcup \mathcal{A}$ is the supremum of $\mathcal{A}$, and $\bigcap \mathcal{A}$ is the infimum of $\mathcal{A}$. Thus the relation $\subset$ on $\mathcal{P}(X)$ has the least upper bound property. ■

Lemma and Definition 2.51
Given a set X, we define

$$\begin{aligned}\mathcal{Q}(X) &= \big\{\, (x,A) \in X \times \mathcal{P}(X) \,:\, x \in A \,\big\} \\ &= \bigcup \big\{\, \{x\} \times \big(\mathcal{P}(X)(x)\big) \,:\, x \in X \,\big\} \subset X \times \mathcal{P}(X)\end{aligned}$$

A relation $R \subset \mathcal{Q}(X)$ is called a **structure relation on** X. The relation $\leq$ on R defined by

$$(y,B) \leq (x,A) \quad \Longleftrightarrow \quad x = y \;\wedge\; A \subset B$$

is an ordering in the sense of "$\leq$". If X has more than one member, this ordering is not connective.

Proof. Exercise. □

2.2 Functions

In this Section we introduce the important concept of function. We analyse various fundamental properties of functions, in particular the interplay of functions with unions and intersections of sets as well as with pre-orderings.

Definition 2.52

Let X and Y be two sets. A **functional relation** f is a relation $f \subset X \times Y$ such that for every $x \in X$ there exists at most one $y \in Y$ with $(x, y) \in f$. Let D be the domain of f. Then f is called a **function from** D **to** Y. We use the standard notation $f : D \longrightarrow Y$. A function is also called **map** in the sequel. For every $x \in D$ we denote by $f(x)$ or f_x the member $y \in Y$ such that $(x, y) \in f$. $f(x)$ is called the **value of** f **at** x. For every $A \subset X$ we call $f[A] = \{f(x) : x \in A\}$ the **image of** A **under** f. For every $B \subset Y$ the set $f^{-1}[B] = \{x \in X : f(x) \in B\}$ is called the **inverse of** B **under** f. For a system $\mathcal{A} \subset \mathcal{P}(X)$ the system $f[\![\mathcal{A}]\!] = \{f[A] : A \in \mathcal{A}\}$ is called **image of** $\mathcal{A}$ **under** f. For a system $\mathcal{B} \subset \mathcal{P}(Y)$ the system $f^{-1}[\![\mathcal{B}]\!] = \left\{f^{-1}[B] : B \in \mathcal{B}\right\}$ is called **inverse of** $\mathcal{B}$ **under** f. ■

Definition 2.53

Let X and Y be two sets. The set of all functions from X to Y is denoted by Y^X. ■

Notice that slightly different definitions are used if X is a natural number (see Definition 3.9) or if Y is a relation (see Definition 3.14). Generally there is no risk of confusion.

Definition 2.54

Let X, Y, and Z be sets, $f : X \times Y \longrightarrow Z$ a function, and $x \in X$, $y \in Y$. Then we also write $f(x, y)$ instead of $f((x, y))$ for the value of f at (x, y). ■

Lemma and Definition 2.55
Given two sets X and Y, and a function $f : X \longrightarrow Y$, f is called **surjective** if $f[X] = Y$, i.e. the range of f is Y. f is called **injective** if $f^{-1}\{y\}$ contains at most one member for each $y \in Y$. f is called **bijective** if f is both surjective and injective.
If f is bijective, then the inverse relation f^{-1} is a functional relation with domain Y, i.e. $f^{-1} : Y \longrightarrow X$. f^{-1} is called **inverse function of** f, or short, **inverse of** f. We have $f^{-1}(f(x)) = x$ for every $x \in X$. f^{-1} is bijective.
If $X = Y$ and $f = \Delta$, then f is called the **identity map on** X, and also denoted by id_X, or, when the set is evident from the context, by id.

Proof. Exercise. □

Remark 2.56
Let X, Y be two sets, $f : X \longrightarrow Y$ a bijection, $\mathcal{A} \subset \mathcal{P}(X)$, and $\mathcal{B} = f[\![\mathcal{A}]\!]$. Then the map $F : \mathcal{A} \longrightarrow \mathcal{B}$, $F(A) = f[A]$ is a bijection too. ■

Definition 2.57
Given a set X and a map $f : X \longrightarrow X$, a member $x \in X$ is called **fixed point of** f if $f(x) = x$. ■

Definition 2.58
Given two sets X, Y, a function $f : X \longrightarrow Y$, and a set $A \subset X$, the functional relation $\{(x, y) \in f : x \in A\}$ is called **restriction of** f **to** A. It is denoted by $f \mid A$. ■

Lemma and Definition 2.59

Given sets X, Y and Z, and functions $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$, the product gf of g and f as defined in Definition 2.1 is also denoted by $g \circ f$. It is a function from X to Z, i.e. $g \circ f : X \longrightarrow Z$. It is also called the **composition of f and g**. We have $g(f(x)) = (gf)(x)$ for every $x \in X$. We also write $gf(x)$ for $(gf)(x)$.

Proof. Exercise. □

Definition 2.60

Given a set X and a map $f : X \longrightarrow X$, f is called a **projection** or **projective**, if $f \circ f = f$. ■

Using the notion of a function, a system of sets and the union and intersection of a system as defined in Lemma and Definition 1.21 and Definition 1.11, respectively, can be written in a different form as follows.

Definition 2.61

A set I is called an **index set** if $I \neq \emptyset$. Given a non-empty system $\mathcal{A}$, an index set I, and a function $A : I \longrightarrow \mathcal{A}$, we define the following notations:

$$\bigcup_{i \in I} A_i = \bigcup \mathcal{B}, \qquad \bigcap_{i \in I} A_i = \bigcap \mathcal{B}$$

where $\mathcal{B} = A\,[I]$. If A is surjective, then it follows that

$$\bigcup_{i \in I} A_i = \bigcup \mathcal{A}, \qquad \bigcap_{i \in I} A_i = \bigcap \mathcal{A}$$

■

We mainly use the notion "index set" for a set that is the non-empty domain of a function to a system of sets as in Definition 2.61, but not for arbitrary non-empty

sets; of course, formally also the system $\mathcal{A}$ is an index set. With the notation of Definition 2.61 we clearly have $\mathcal{A} = \{A_i : i \in I\}$ if A is surjective. It is often more convenient to use the index notations than an abstract letter for the system of sets. Notice that there is a slight difference between the two notations because we may have $A_i = A_j$ for $i, j \in I$ with $i \neq j$. This happens if the map A is not injective. However, in most cases this turns out to be irrelevant. When using index notation, we often do neither explicitly introduce a letter for the range system (e.g. $\mathcal{A}$) nor a letter for the function (e.g. A). Instead we only introduce an index set I and the sets A_i $(i \in I)$ that specify the values of the function and that are precisely the members of the range system. In particular, a definition of the index set I and the sets A_i $(i \in I)$ does not tacitly imply that the letter A without subscript stands for the corresponding function unless this is explicitly said; we may even use the letter A for other purposes, for example we may define $A = \bigcup_{i \in I} A_i$.
The following identities are the analogues of Lemmas 1.34 and 1.35.

Lemma 2.62
Let I and J be index sets and A_i $(i \in I)$, B_j $(j \in J)$ sets. Then the following equalities hold:

(i) $\big(\bigcup_{i\in I} A_i\big) \cap \big(\bigcup_{j\in J} B_j\big) = \bigcup \{A_i \cap B_j \,:\, i \in I,\, j \in J\}$

(ii) $\big(\bigcap_{i\in I} A_i\big) \cup \big(\bigcap_{j\in J} B_j\big) = \bigcap \{A_i \cup B_j \,:\, i \in I,\, j \in J\}$

(iii) $\big(\bigcap_{i\in I} A_i\big)^c = \bigcup_{i\in I} A_i^c$

(iv) $\big(\bigcup_{i\in I} A_i\big)^c = \bigcap_{i\in I} A_i^c$

Proof. Exercise. □

The following two Lemmas show how the image and the inverse under f behave together with intersections and unions.

Lemma 2.63
Given a function $f : X \longrightarrow Y$, an index set I, and sets $A_i \subset X$ $(i \in I)$, the following statements hold:

(i) $f\left[\bigcup_{i\in I} A_i\right] = \bigcup_{i\in I} f\left[A_i\right]$

(ii) $f\left[\bigcap_{i\in I} A_i\right] \subset \bigcap_{i\in I} f\left[A_i\right]$

Proof. Exercise. □

Lemma 2.64
Given a function $f : X \longrightarrow Y$, an index set I, a set $A \subset Y$, and sets $A_i \subset Y$ $(i \in I)$, the following relations hold:

(i) $f^{-1}\left[\bigcup_{i\in I} A_i\right] = \bigcup_{i\in I} f^{-1}\left[A_i\right]$

(ii) $f^{-1}\left[\bigcap_{i\in I} A_i\right] = \bigcap_{i\in I} f^{-1}\left[A_i\right]$

(iii) $f^{-1}\left[A^c\right] = \left(f^{-1}\left[A\right]\right)^c$

where the complement refers to Y.

Proof. Exercise. □

The following Definition generalizes the Cartesian product of two sets as defined in Definition 1.37.

Definition 2.65

Let I be an index set, $\mathcal{A}$ a non-empty system, $A : I \longrightarrow \mathcal{A}$ a map, and $B = \bigcup_{i\in I} A_i$. We define the **Cartesian product of** A as follows:

$$\bigtimes_{i\in I} A_i = \left\{ f \in B^I \ : \ \forall i \in I \ \ f(i) \in A_i \right\}$$

For each $i \in I$, the map $p_i : \bigtimes_{i\in I} A_i \longrightarrow A_i$, $p_i(f) = f(i)$, is called the **projection on** A_i. ■

Notice that in our definition of the Cartesian product we use index notation, which allows identical factors. Of course, using index notation for the projections we formally have to think of a surjective map $p : I \longrightarrow \{p_i \ : \ i \in I\}$.

It is a consequence of the Choice axiom that the Cartesian product is not always empty. The following Remark is a repetition of Lemma and Definition 1.50, now using functional notation.

Remark 2.66

Let X be a set and $\mathcal{A} \subset \mathcal{P}(X)$ with $\emptyset \notin \mathcal{A}$. Then there exists a function $f : \mathcal{A} \longrightarrow X$ such that $f(A) \in A$ for every $A \in \mathcal{A}$. f is a choice function. ■

Corollary 2.67

With definitions as in Definition 2.65, $\emptyset \notin \mathcal{A}$ implies that $\bigtimes_{i\in I} A_i \neq \emptyset$.

Proof. This is a consequence of Remark 2.66. □

Remark 2.68

With definitions as in Definition 2.65, the following statements hold:

(i) If $A_i = A$ $(i \in I)$ for some set A, then $\bigtimes_{i\in I} A_i = A^I$.

(ii) If $A_i = \emptyset$ for some $i \in I$, then $\bigtimes_{i\in I} A_i = \emptyset$.

■

If the index set in Definition 2.65 is a singleton, the Cartesian product can obviously be identified with the single factor set in the following manner.

Remark 2.69

Let X and a be two sets and $I = \{a\}$. We define the map $f : X^I \longrightarrow X$, $f(h) = h(a)$. Then f is bijective. ■

The following Remark says that Definitions 1.37 and 2.65 are in agreement with each other.

Remark 2.70

Let X, Y, a, and b be sets with $a \neq b$. We further define the sets $I = \{a, b\}$, $X_a = X$, $X_b = Y$, and the function $f : \bigtimes_{i\in I} X_i \longrightarrow X \times Y$, $f(h) = (h(a), h(b))$. Then f is bijective. In particular, this gives us a bijection from $X \times X$ to X^I. ■

The following result says that an iterated Cartesian product can be identified with a simple Cartesian product.

Remark 2.71

Let $\mathcal{J}$ be a non-empty system of disjoint index sets, $J : I \longrightarrow \mathcal{J}$ a bijection where I is an index set, $K = \bigcup \mathcal{J}$, $\mathcal{A}$ a non-empty system, and $F : K \longrightarrow \mathcal{A}$ a map. Then the Cartesian product $\bigtimes_{j\in K} F_j$ is well-defined. Further let $G : I \longrightarrow \mathcal{P}\,(K \times \mathcal{A})$ be the map such that, for every $i \in I$, $G(i)$ is a functional relation with domain J_i, i.e. $G_i : J_i \longrightarrow \mathcal{A}$, and $G_i(j) = F(j)$. Thus, for every $i \in I$, the Cartesian product $\bigtimes_{j\in J(i)} G_i(j)$ is well-defined. Now let $A = \bigcup \mathcal{A}$ and $H : I \longrightarrow \mathcal{P}^2\,(K \times A)$, $H(i) = \bigtimes_{j\in J(i)} G_i(j)$. We define

$$f : \bigtimes_{j\in K} F_j \longrightarrow \bigtimes_{i\in I} H_i \,,$$

$$\Big(\big(f(h)\big)(i)\Big)(j) = h(j) \quad \text{for every } i \in I \text{ and } j \in J_i$$

Then f is a bijection. ∎

Remarks 2.69, 2.70, and 2.71 can be combined in different ways. The following is a useful example.

Remark 2.72

Let I be an index set, X_i $(i \in I)$ sets, $j \in I$, and $J = I \setminus \{j\}$. If $J \neq \emptyset$, then there is a bijection from $\bigtimes_{i\in I} X_i$ to $\Big(\bigtimes_{i\in J} X_i\Big) \times X_j$ by Remarks 2.69, 2.70, and 2.71. ∎

Definition 2.73

Let X be a set and I an index set. Further let Y_i $(i \in I)$ be sets and $f_i : X \longrightarrow Y_i$ maps. We say that $\{f_i \,:\, i \in I\}$ **distinguishes points** if for every $x, y \in X$ with $x \neq y$ there is $i \in I$ such that $f_i(x) \neq f_i(y)$. ∎

Remark 2.74

With definitions as in Definition 2.65, the set of functions $\{p_i : i \in I\}$ distinguishes points. ■

Definition 2.75

Let X be a set, $(Y, \prec)$ a pre-ordered space, and $f : X \longrightarrow Y$ a function. f is called **bounded** if there exist $x, y \in Y$ such that $f[X] \subset\ ]x, y[\ \cup \{x, y\}$. Otherwise f is called **unbounded**. f is called **bounded from below** if there is $x \in Y$ such that $f[X] \subset\ ]x, \infty[\ \cup \{x\}$. f is called **bounded from above** if there is $y \in Y$ such that $f[X] \subset\ ]-\infty, y[\ \cup \{y\}$. ■

Lemma 2.76

Let $(X, \prec)$ be an ordered space where $\prec$ has the least upper bound property, Y a non-empty set, and $f : Y \longrightarrow X$, $g : Y \longrightarrow X$ two functions such that $f(y) \prec g(y)$ for every $y \in Y$. The following two statements hold:

(i) If f is bounded from below, then $\inf f[Y] \prec \inf g[Y]$ or $\inf f[Y] = \inf g[Y]$.

(ii) If g is bounded from above, then $\sup f[Y] \prec \sup g[Y]$ or $\sup f[Y] = \sup g[Y]$.

Proof. In order to prove (i), let L_f be the set of all lower bounds of $f[Y]$ and L_g the set of all lower bounds of $g[Y]$. Under the stated conditions we have $L_f \subset L_g$. Since $f[Y]$ has a lower bound, it has an infimum by Theorem 2.49. Moreover, the infimum of $f[Y]$ is unique since $\prec$ is an ordering. Similarly, also $\inf g[Y]$ exists. The claim now follows by the fact that $L_f \subset L_g$.
The proof of (ii) is similar. □

Definition 2.77
Let $(X, \prec)$ and $(Z, \prec)$ be ordered spaces, $A \subset X$, $B \subset Z$, and $f : A \longrightarrow B$ a function. f is called **monotonically increasing** or **increasing** or **non-decreasing** if, for every $x, y \in A$, $x \prec y$ implies $f(x) \prec f(y)$ or $f(x) = f(y)$. f is called **monotonically decreasing** or **decreasing** or **non-increasing** if, for every $x, y \in A$, $x \prec y$ implies $f(y) \prec f(x)$ or $f(x) = f(y)$. f is called **monotonic** if it is either increasing or decreasing.
Further, f is called **strictly increasing** if, for every $x, y \in A$ with $x \neq y$, $x \prec y$ implies $f(x) \prec f(y)$ and $f(x) \neq f(y)$. f is called **strictly decreasing** if, for every $x, y \in A$ with $x \neq y$, $x \prec y$ implies $f(y) \prec f(x)$ and $f(x) \neq f(y)$. f is called **strictly monotonic** if it is either strictly increasing or strictly decreasing.
■

The following result shows that the notions defined in Definition 2.77 are invariant under the change from the original orderings to the orderings in the sense of "$<$" and "$\leq$" as defined in Lemma 2.19, both in the domain and in the range space.

Lemma 2.78
Let (X, R) and (Y, S) be ordered spaces, and $f : X \longrightarrow Y$ a function. Further let $T \in \{R \cup \Delta,\, R \setminus \Delta\}$ and $U \in \{S \cup \Delta,\, S \setminus \Delta\}$. f is increasing, decreasing, strictly increasing, or strictly decreasing with respect to the orderings T on X and U on Y iff it has the respective property with respect to the orderings R and S.

Proof. Exercise. □

Definition 2.79
Given a set X, a function $f : X \times X \longrightarrow X$ is called **binary function on** X. The symbol f is also denoted by $\bullet$, and for every $x, y \in X$ we also write $x \bullet y$ for $f(x, y)$. If the equality $x \bullet (y \bullet z) = (x \bullet y) \bullet z$ holds for every $x, y, z \in X$, the function is called **associative**. If the equality $x \bullet y = y \bullet x$ holds for every $x, y \in X$, the function is called **commutative**.
The triple (X, f, e)—or $(X, \bullet, e)$—with $e \in X$ is called a **group** if the following statements hold:

(i) The function $\bullet$ is associative.

(ii) For every $x \in X$, we have $x \bullet e = x$. e is called **neutral element**.

(iii) For every $x \in X$, there is $y \in X$ such that $x \bullet y = e$. y is called **inverse of** x.

If $\bullet$ is commutative, the group is called **Abelian**. ■

In the remainder of the text various binary functions are introduced and different symbols are defined, for example $+$ is used instead of $\bullet$ for the addition of natural numbers.

Theorem 2.80
Given a group $(X, \bullet, e)$, the following statements hold:

(i) For every $x \in X$, we have $e \bullet x = x$.

(ii) There is no $d \in X \setminus \{e\}$ such that $x \bullet d = x$ for every $x \in X$.

(iii) For every $x \in X$, there is a unique $y \in X$ such that $x \bullet y = y \bullet x = e$.

Proof. We first show that $x \bullet y = e$ implies $y \bullet x = e$ for every $x, y \in X$. Assume $x \bullet y = e$. There is $z \in X$ such that $y \bullet z = e$. It follows that $y \bullet x = y \bullet (x \bullet e) =$

$y \bullet (x \bullet (y \bullet z)) = y \bullet ((x \bullet y) \bullet z) = y \bullet (e \bullet z) = (y \bullet e) \bullet z = y \bullet z = e$.
To prove (i), let $x \in X$. There is $y \in X$ such that $x \bullet y = e$. Thus $e \bullet x = (x \bullet y) \bullet x = x \bullet (y \bullet x) = x$.
To prove (ii), assume that such a member d exists. Then $d = e \bullet d$ by (i) and $e \bullet d = e$ by assumption. Thus $d = e$.
To show the uniqueness in (iii), let $x, y, z \in X$ such that $x \bullet y = x \bullet z = e$. It follows that $y \bullet x = z \bullet x = e$. Therefore $y = y \bullet e = y \bullet (x \bullet z) = (y \bullet x) \bullet z = e \bullet z = z$.
□

2.3 Relations and maps

In this Section we analyse how relations behave under maps. This is used subsequently for various purposes.

Lemma and Definition 2.81
Let (X, R) and (Y, S) be two relational spaces, and $f : X \longrightarrow Y$ a map. We use the same symbol for the function $f : X \times X \longrightarrow Y \times Y$, $f(x, z) = (f(x), f(z))$, as the two functions can be distinguished by their arguments. We have $f[R] = \{ (f(x), f(z)) : (x, z) \in R \}$, which is a relation on Y, and $f^{-1}[S] = \{ (x, z) \in X \times X : (f(x), f(z)) \in S \}$, which is a relation on X. The following statements hold:

(i) If S is transitive, then $f^{-1}[S]$ is transitive.

(ii) If S is reflexive, then $f^{-1}[S]$ is reflexive.

(iii) If S is antisymmetric and f is injective, then $f^{-1}[S]$ is antisymmetric.

(iv) If S is antireflexive, then $f^{-1}[S]$ is antireflexive.

Proof. Exercise. □

Example 2.82
Let (X_i, R_i) $(i \in I)$ be pre-ordered spaces, where I is an index set, and $X = \bigtimes_{i \in I} X_i$. Then $\mathcal{R} = \left\{p_i^{-1}[R_i] : i \in I\right\}$ is a system of pre-orderings on X. ■

Remark 2.83
Let X be a set, $\mathcal{R}$ a system of relations on X, and $S = \bigcap \mathcal{R}$. Then the following statements hold:

(i) If every $R \in \mathcal{R}$ is transitive, then S is transitive.

(ii) If every $R \in \mathcal{R}$ is reflexive, then S is reflexive.

(iii) If there is $R \in \mathcal{R}$ such that R is antisymmetric, then S is antisymmetric.

(iv) If there is $R \in \mathcal{R}$ such that R is antireflexive, then S is antireflexive.

In other words, if every $R \in \mathcal{R}$ is a pre-ordering, then S is a pre-ordering. Moreover, if the members of $\mathcal{R}$ are pre-orderings and at least one member is an ordering, then S is an ordering. Finally notice that the above also states conditions under which S is an ordering in the sense of "$<$" or "$\leq$". ■

Example 2.84
Let (X_i, R_i) $(i \in I)$ be pre-ordered spaces, where I is an index set, and $X = \bigtimes_{i \in I} X_i$. Then $S = \bigcap \left\{p_i^{-1}[R_i] : i \in I\right\}$ is a pre-ordering on X. ■

Lemma and Definition 2.85
Let (X_i, R_i) $(i \in I)$ be directed spaces, where I is an index set, and $X = \bigtimes_{i \in I} X_i$. Then $S = \bigcap \left\{p_i^{-1}[R_i] : i \in I\right\}$ is a direction on X. (X, S) is called **product directed space**.

Proof. Exercise. □

The following notion is used in Section 5.5 where we consider interval topologies.

Definition 2.86

Let X be a set, $\mathcal{R} = \{\prec_i \,:\; i \in I\}$ a system of pre-orderings on X where I is an index set, and $S = \bigcap \mathcal{R}$. The pre-ordering S is also denoted by $\prec$. $\mathcal{R}$ is called **upwards independent** if for every $i \in I$, every $s \in X$, and every $x \in X$ with $x \prec_i s$, there exists $y \in X$ such that $x \prec y$ and $y \prec_i s$.
$\mathcal{R}$ is called **downwards independent** if for every $i \in I$, every $r \in X$, and every $x \in X$ with $r \prec_i x$, there exists $y \in X$ such that $y \prec x$ and $r \prec_i y$.
$\mathcal{R}$ is called **independent** if it is both upwards and downwards independent. ■

Lemma 2.87

Let X be a set, $\mathcal{R} = \{\prec_i \,:\; i \in I\}$ a system of pre-orderings on X where I is an index set, and $S = \bigcap \mathcal{R}$. The pre-ordering S is also denoted by $\prec$. Intervals with respect to the pre-ordering $\prec_i$ are denoted by subscript i, those with respect to the pre-ordering $\prec$ are denoted without subscript.

(i) If $\mathcal{R}$ is upwards independent, then we have for every $i \in I$ and $s \in X$

$$]{-\infty}, s[_i \;=\; \bigcup \big\{\,]{-\infty}, x[\;:\; x \in X,\; x \prec_i s \big\}$$

(ii) If $\mathcal{R}$ is downwards independent, then we have for every $i \in I$ and $r \in X$

$$]r, \infty[_i \;=\; \bigcup \big\{\,]x, \infty[\;:\; x \in X,\; r \prec_i x \big\}$$

(iii) If $\mathcal{R}$ is independent, then we have for every $i \in I$ and $r, s \in X$

$$]r, s[_i \;=\; \bigcup \big\{\,]x, y[\;:\; x, y \in X,\; x \prec y,\; r \prec_i x,\; y \prec_i s \big\}$$

Proof. To see (i), assume the stated condition and let $i \in I$ and $s \in X$. We have

$$\begin{aligned}]{-\infty}, s[_i &= \{z \in X : z \prec_i s\} \\ &= \{z \in X : \exists x \in X \quad z \prec x,\ x \prec_i s\} \\ &= \bigcup \Big\{ \{z \in X : z \prec x\} : x \in X,\ x \prec_i s \Big\} \end{aligned}$$

The proof of (ii) is similar.
To show (iii), assume the stated condition and let $i \in I$ and $r, s \in X$. We have

$$\begin{aligned}]r, s[_i &= \ ]r, \infty[_i \cap \ ]{-\infty}, s[_i \\ &= \bigcup \{ \ ]x, \infty[\ : \ x \in X,\ r \prec_i x\} \cap \bigcup \{ \ ]{-\infty}, y[\ : \ y \in X,\ y \prec_i s\} \\ &= \bigcup \{ \ ]x, y[\ : \ x, y \in X,\ x \prec y,\ r \prec_i x,\ y \prec_i s\} \end{aligned}$$

where the second equation follows by (i) and (ii), and the third equation by Lemma 1.34 (i). □

In the following Definition we introduce a notation that is convenient for the analysis of set functions in Section 5.1.

Definition 2.88
Given a relational space (X, R) and a function $f : X \longrightarrow X$, the relation $f^{-1}[R]$ is also denoted by R_f. If R is a pre-ordering, we also write $x \prec_f y$ for $(x, y) \in R_f$. ■

The notation defined in the above Definition is meaningful since R_f is a pre-ordering if R is a pre-ordering.

Definition 2.89
Given a relational space (X, R), a function $f : X \longrightarrow X$ is called R-**increasing**, if $(x, y) \in R$ implies $(x, f(y)), (f(x), f(y)) \in R$ for every $x, y \in X$. ■

Remark 2.90
Let (X, R) be a relational space, and $f : X \longrightarrow X$ and $g : X \longrightarrow X$ two R-increasing maps. Then $g \circ f$ is R-increasing. ■

Lemma 2.91
Given a set X, a reflexive pre-ordering $\prec$ on X, and an $\prec$-increasing projective map $f : X \longrightarrow X$, we have

$$x \prec_f y \quad \Longleftrightarrow \quad x \prec f(y)$$

Proof. Fix $x, y \in X$. We have $x \prec x$, and therefore $x \prec f(x)$. Assume $x \prec_f y$. It follows that $f(x) \prec f(y)$, and thus $x \prec f(y)$, since f is transitive. Now assume instead that $x \prec f(y)$. Since f is $\prec$-increasing and projective, we obtain $f(x) \prec f(y)$. □

Chapter 3

Numbers I

3.1 Natural numbers, induction, recursion

In Definition 1.43 we have defined the set $\mathbb{N}$ of natural numbers. In this Section we derive two important Theorems: the Induction principle for natural numbers and the Recursion theorem for natural numbers. Based on these Theorems we define and analyse the addition, multiplication, and exponentiation on the natural numbers. The natural numbers are the starting point for the construction of the other number systems below in this Chapter.
We first introduce the conventional symbols for four specific sets, that are natural numbers.

Definition 3.1
We define the sets $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, and $3 = \{0, 1, 2\}$. Furthermore, we define the function $\sigma : \mathbb{N} \longrightarrow \mathbb{N} \setminus \{0\}$, $\sigma(m) = m \cup \{m\}$. ■

We clearly have $\sigma(0) = 1$, $\sigma(1) = 2$, and $\sigma(2) = 3$. Notice that σ is well-defined since $\mathbb{N}$ is inductive and $m \cup \{m\}$ is non-empty for every $m \in \mathbb{N}$.

Theorem 3.2 (Induction principle for natural numbers)
Let $A \subset \mathbb{N}$. If $0 \in A$ and if $\sigma(n) \in A$ for every $n \in A$, then $A = \mathbb{N}$.

Proof. Assume A satisfies the conditions. Then A is inductive. It follows that $\mathbb{N} \subset A$ by Definition 1.43. □

Theorem 3.3
The natural numbers have the following properties:

(i) $\forall m \in \mathbb{N} \quad m \subset \mathbb{N}$

(ii) $\forall n \in \mathbb{N} \quad \forall m \in n \quad m \in \mathbb{N} \wedge m \subset n$

(iii) $\forall n \in \mathbb{N} \quad \forall m \in n \quad \sigma(m) \in \sigma(n)$

(iv) $\forall m, n \in \mathbb{N} \quad m = n \vee m \in n \vee n \in m$

(v) $\forall m, n, p \in \mathbb{N} \quad m \in n \wedge n \in p \implies m \in p$

(vi) $\neg \exists m \in \mathbb{N} \quad m \notin m$

(vii) $\forall m \in \mathbb{N} \quad \neg \exists n \in \mathbb{N} \quad m \in n \in m \cup \{m\}$

Proof. To see (i), let $A = \{m \in \mathbb{N} : m \subset \mathbb{N}\}$. Clearly, $0 \in A$. Now assume that $m \in A$ for some $m \in \mathbb{N}$. Then we have $m \subset \mathbb{N}$. Therefore $\sigma(m) \subset \mathbb{N}$, and thus $\sigma(m) \in A$. It follows that $A = \mathbb{N}$ by the Induction principle.
To show (ii), let $A = \{n \in \mathbb{N} : \forall m \in n \ (m \in \mathbb{N} \wedge m \subset n)\}$. Clearly, $0 \in A$. Now assume that $n \in A$ for some $n \in \mathbb{N}$. Let $m \in \sigma(n)$. We have either $m \in n$ or $m = n$. In the first case, we obtain $m \in \mathbb{N}$ and $m \subset n \subset \sigma(n)$ by assumption. In the second case, we obviously have $m \in \mathbb{N}$ and $m \subset \sigma(n)$. We obtain $A = \mathbb{N}$ by the Induction principle.
To show (iii) we again apply the Induction principle. The claim is trivially true for $n = 0$ and every $m \in n$. Assume the claim is true for some $n \in \mathbb{N}$ and every $m \in n$. Let $m \in \sigma(n)$. If $m \in n$, then $\sigma(m) \in \sigma(n)$ by assumption. If $m = n$, then $\sigma(m) = \sigma(n)$. It follows that $\sigma(m) \in \sigma(\sigma(n))$.
To prove (iv), we use the Induction principle with respect to m. First let $m = 0$. If $n = 0$, then we have $m = n$. If $n \neq 0$, then $0 \in n$, which is easily shown by the Induction principle. Thus the claim is true for $m = 0$. Assume the claim holds for some $m \in \mathbb{N}$ and every $n \in \mathbb{N}$. Fix $n \in \mathbb{N}$. If $n \in m$ or $n = m$, then $n \in \sigma(m)$.

If $m \in n$, then either $\sigma(m) = n$ or $\sigma(m) \in n$ by (iii).
To see (v), notice that $n \subset p$ under the stated conditions by (ii).
(vi) is a consequence of Lemma 1.46 (i).
To see (vii), notice that if such n exists, then we have either $n = m$ or $n \in m$ both of which is excluded by Lemma 1.46. □

Lemma and Definition 3.4
We define a total ordering in the sense of "$<$" on the natural numbers by:

$$m < n \quad \Longleftrightarrow \quad m \in n$$

For every $m \in \mathbb{N}$, $\sigma(m)$ is the unique successor of m, and we have $m = \{n \in \mathbb{N} : n < m\}$.
We further define $\leq$ to be the total ordering in the sense of "$\leq$" on the natural numbers by the method of Lemma 2.19.

Proof. This follows by Theorem 3.3. □

Notice that, in particular, 1 is the successor of 0 etc.

Definition 3.5
We adopt the convention that all notions related to orderings on $\mathbb{N}$ refer to the ordering $<$ as defined in Lemma and Definition 3.4 unless otherwise specified. ■

Note that in many contexts it is irrelevant whether the ordering $<$ or the ordering $\leq$ on $\mathbb{N}$ is considered since most properties related to orderings are invariant, cf. Lemmas 2.34, 2.46, and 2.78.
The following version of the Induction principle allows us to prove statements inductively for all natural numbers that are larger than a fixed number.

Corollary 3.6
Let $A \subset \mathbb{N}$ and $m \in \mathbb{N}$. If $\sigma(m) \in A$ and if $\sigma(n) \in A$ for every $n \in A$ with $n > m$, then we have $\{n \in \mathbb{N} : n > m\} \subset A$.

Proof. We show that, under the stated conditions, $n > m$ implies $n \in A$ for every $n \in \mathbb{N}$ by the Induction principle. This implication is trivially true for $n = 0$. Now assume that it is true for some $n \in \mathbb{N}$. We distinguish the cases $n < m$, $n = m$, and $n > m$ by Theorem 3.3 (iv). If $n < m$, then either $\sigma(n) < m$ or $\sigma(n) = m$ by Theorem 3.3 (iii), and thus the implication again holds trivially for $\sigma(n)$. If $n = m$, then $\sigma(n) \in A$ as this is amongst the conditions. If $n > m$, then $n \in A$ by assumption, and thus $\sigma(n) \in A$ since this is amongst the conditions. $\square$

We also refer to Corollary 3.6 as the Induction principle.

Theorem 3.7
$\sigma : \mathbb{N} \longrightarrow \mathbb{N} \setminus \{0\}$ is a bijection.

Proof. To see that σ is surjective, notice that $1 \in \sigma[\mathbb{N}]$ and $\sigma(m) \in \sigma[\mathbb{N}]$ whenever $m \in \sigma[\mathbb{N}]$. It follows that $\sigma[\mathbb{N}] = \mathbb{N} \setminus \{0\}$ by the Induction principle in the form of Corollary 3.6.
To show that σ is injective, let $m, n \in \mathbb{N}$ such that $\sigma(m) = \sigma(n)$. Hence we have $m \cup \{m\} = n \cup \{n\}$. This implies

$$m = n \vee \big(m \in n \wedge n \in m\big)$$

By Theorem 3.3 it follows that $m = n$. $\square$

Corollary 3.8
Every $m \in \mathbb{N} \setminus \{0\}$ has a unique predecessor with respect to the ordering $<$.

Proof. This follows by Lemma and Definition 3.4 and Theorem 3.7. $\square$

The next definition is a modification of Definition 2.53 for the case where the superscript is a natural number larger than 0.

Definition 3.9
Let X be a set and $m \in \mathbb{N}$, $m > 0$. The system of functions X^I where $I = \sigma(m)\backslash\{0\}$ is also denoted by X^m. ∎

Notice that this deviates from Definition 2.53 where the superscript would be the domain and thus would contain 0 but not m. By Remark 2.70 we may write members of X^2 as ordered pairs as follows.

Definition 3.10
Given a set X, we also write $(f(1), f(2))$ for $f \in X^2$. ∎

Definition 3.11
Let $m, n \in \mathbb{N}$ with $m < n$. Further let $I = \sigma(n)\backslash m$ and A_i $(i \in I)$ be sets. We define

$$\bigcup\nolimits_{i=m}^{n} A_i = \bigcup\nolimits_{i\in I} A_i\,, \qquad \bigcap\nolimits_{i=m}^{n} A_i = \bigcap\nolimits_{i\in I} A_i\,, \qquad \bigtimes\nolimits_{i=m}^{n} A_i = \bigtimes\nolimits_{i\in I} A_i$$

∎

Definition 3.12
Let $m \in \mathbb{N}$, $I = \mathbb{N}\backslash m$, and A_i $(i \in I)$ be sets. We define

$$\bigcup\nolimits_{i=m}^{\infty} A_i = \bigcup\nolimits_{i\in I} A_i\,, \qquad \bigcap\nolimits_{i=m}^{\infty} A_i = \bigcap\nolimits_{i\in I} A_i\,, \qquad \bigtimes\nolimits_{i=m}^{\infty} A_i = \bigtimes\nolimits_{i\in I} A_i$$

∎

The following Theorem states that one may define a function from $\mathbb{N}$ to a set X

recursively.

Theorem 3.13 (Recursion for natural numbers)
Given a set X, a point $x \in X$ and a function $f : X \longrightarrow X$, there exists a unique function $g : \mathbb{N} \longrightarrow X$ with the following properties:

(i) $g(0) = x$

(ii) $g(\sigma(n)) = f(g(n)) \quad$ for every $n \in \mathbb{N}$

Proof. For every $p \in \mathbb{N}$ there exists a map $G : \sigma(p) \longrightarrow X$ with the following properties:

(i) $G(0) = x$

(ii) $G(\sigma(n)) = f(G(n)) \quad$ for every $n \in p$

[This is clear for $p = 0$. Assume there exists such a function G for $p \in \mathbb{N}$. We define $H : \sigma(\sigma(p)) \longrightarrow X$, $H|\sigma(p) = G$, $H(\sigma(p)) = f(G(p))$. The assertion follows by the Induction principle.]

We call a function G with these properties a "debut of size $\sigma(p)$". Let G and H be two debuts of sizes $\sigma(p)$ and $\sigma(q)$, respectively, where $p, q \in \mathbb{N}$. We may assume that $p < q$ or $p = q$. We have $G = H|\sigma(p)$.

[Clearly, $G(0) = x = H(0)$. Now let $n \in \mathbb{N}$, $n < p$ and assume that $G(n) = H(n)$. Then also $G(\sigma(n)) = H(\sigma(n))$ holds. The assertion follows by the Induction principle.]

Now, for every $n \in \mathbb{N}$, let $g(n) = G(n)$ where G is the debut of size $\sigma(n)$. Clearly, g satisfies (i) and (ii) of the claim.

[(i) is satisfied since $g(0) = x$. (ii) is satisfied for $n = 0$ since $g(1) = G(1) = f(G(0)) = f(x) = f(g(0))$ where G is the debut of size 2. Now assume that (ii) is true for some $n \in \mathbb{N}$, that is $g(\sigma(n)) = f(g(n))$. Then we have $g\big(\sigma(\sigma(n))\big) = G\big(\sigma(\sigma(n))\big) = f\big(G(\sigma(n))\big) = f\big(H(\sigma(n))\big) = f\big(g(\sigma(n))\big)$ where G is the debut of size $\sigma\big(\sigma(\sigma(n))\big)$ and H the debut of size $\sigma(\sigma(n))$. Thus (ii) is true for every $n \in \mathbb{N}$ by the Induction principle.]

To see that g is unique, assume that also $h : \mathbb{N} \longrightarrow X$ satisfies (i) and (ii) of the claim. Obviously, $g(0) = x = h(0)$. Moreover, $g(n) = h(n)$ for some $n \in \mathbb{N}$ implies $g(\sigma(n)) = h(\sigma(n))$. It follows by the Induction principle that $g = h$. □

Theorem 3.13 can be used to define powers (i.e. multiple products) of a relation on a given set X as follows.

Lemma and Definition 3.14
Given a relational space (X, R), we define, for every $m \in \mathbb{N}$, $0 < m$, a relation R^m on X by

(i) $R^1 = R$

(ii) $R^{\sigma(m)} = R^m R$

Proof. The existence and uniqueness of R^m for every $m \in \mathbb{N}$, $0 < m$ follows by Theorem 3.13. □

Notice that this definition deviates from Definition 2.53. Since Lemma and Definition 3.14 is valid only for relations, there is generally no risk of confusion.
As a further immediate consequence of Theorem 3.13 we obtain the following result that allows us to define binary functions on $\mathbb{N}$ recursively.

Corollary 3.15
Given two maps $e : \mathbb{N} \longrightarrow \mathbb{N}$ and $f : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$, there exists a unique function $g : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ with the following properties:

(i) $g(m, 0) = e(m) \quad$ for every $m \in \mathbb{N}$

(ii) $g(m, \sigma(n)) = f(m, g(m, n)) \quad$ for every $m, n \in \mathbb{N}$

Proof. For each $m \in \mathbb{N}$ there is a unique function $g_m : \mathbb{N} \longrightarrow \mathbb{N}$ by Theorem 3.13 with the following properties:

(i) $g_m(0) = e(m)$

(ii) $g_m(\sigma(n)) = f(m, g_m(n)) \quad$ for every $n \in \mathbb{N}$

We may then define g by $g(m, n) = g_m(n)$. □

In the remainder of this Section we introduce three binary functions on $\mathbb{N}$, viz. addition, multiplication, and exponentiation.

Lemma and Definition 3.16
There is a unique binary function $+$ on $\mathbb{N}$ such that for every $m, n \in \mathbb{N}$ we have

(i) $m + 0 = m$

(ii) $m + \sigma(n) = \sigma(m + n)$

The function $+$ is called **addition on** $\mathbb{N}$. For every $m, n \in \mathbb{N}$, the expression $m + n$ is called the **sum of** m **and** n.
$+$ is associative and commutative. Furthermore, for every $m, n, p \in \mathbb{N}$, the following implication holds:

$$m < n \quad \Longrightarrow \quad m + p < n + p$$

Proof. The existence and uniqueness of the function $+$ follows by Corollary 3.15.

$+$ is commutative since it follows by the Induction principle that for every $p \in \mathbb{N}$ the following equations hold for $m, n \in \mathbb{N}$, $m + n = p$:

(i) $m + n = n + m$

(ii) $\sigma(n) + m = n + \sigma(m)$

[First notice that $m+n = 0$ implies $m = n = 0$. Thus (i) holds for $m, n \in \mathbb{N}$, $m + n = 0$. Moreover, we have $0 + m = m$ for every $m \in \mathbb{N}$. In fact, this equation clearly holds for $m = 0$ and, assuming it holds for some $m \in \mathbb{N}$, it also holds for $\sigma(m)$ since $0 + \sigma(m) = \sigma(0 + m) = \sigma(m)$. As a special case, we obtain (ii) for $m, n \in \mathbb{N}$, $m + n = 0$, viz. $1 + 0 = 0 + 1$. Now assume that (i) and (ii) hold for some $p \in \mathbb{N}$ and for every $m, n \in \mathbb{N}$ with $m + n = p$. Let $m, n \in \mathbb{N}$ such that $m + n = \sigma(p)$. If $n = 0$, then (i) holds as shown above. If $0 < n$, then let q be the predecessor of n. We then have $m + n = m + \sigma(q) = \sigma(m + q) = \sigma(q + m) = q + \sigma(m) = \sigma(q) + m = n + m$. Hence (i) also holds for $0 < n$. If $m = 0$, then (ii) clearly holds. If $0 < m$, then let q be the predecessor of m. We then obtain $\sigma(n)+m = \sigma(n)+\sigma(q) = \sigma(\sigma(n) + q) = \sigma(n + \sigma(q)) = \sigma(n + m) = n + \sigma(m)$, which is equation (ii).]

To see that the addition is associative, let $m, p \in \mathbb{N}$ and

$$A = \big\{n \in \mathbb{N} \,:\, (m + n) + p = m + (n + p)\big\}$$

Clearly, $0 \in A$ by commutativity. Now assume that $n \in A$ for some $n \in \mathbb{N}$. We have $(m + \sigma(n)) + p = \sigma(m + n) + p = p + \sigma(m + n) = \sigma(p + (m + n)) = \sigma((m+n)+p) = \sigma(m+(n+p)) = m+\sigma(n+p) = m+\sigma(p+n) = m+(p+\sigma(n))$. Also the last claim can be shown by the Induction principle. It clearly holds for $m, n \in \mathbb{N}$ and $p = 0$. Assume it holds for some $p \in \mathbb{N}$ and every $m, n \in \mathbb{N}$. Let $m, n \in \mathbb{N}$ with $m < n$. Then $m + \sigma(p) = \sigma(m + p) < \sigma(n + p) = n + \sigma(p)$. □

The associativity of the addition allows us to write multiple sums without brackets, i.e. $m + n + p$ instead of $(m + n) + p$ or $m + (n + p)$, and similarly for sums with more than three terms.

Lemma 3.17
For every $m, n \in \mathbb{N}$ with $m < n$ there is $p \in \mathbb{N}$ such that $m + p = n$.

Proof. Let $m \in \mathbb{N}$. Then the claim follows by the Induction principle in the form of Corollary 3.6 as follows. The claim clearly holds for $n = \sigma(m)$. Assuming it holds for some $n \in \mathbb{N}$ with $n > m$, we may choose $p \in \mathbb{N}$ such that $m + p = n$. It follows that $m + \sigma(p) = \sigma(n)$. $\square$

Lemma 3.18
Let (X, R) be a relational space, and $m, n \in \mathbb{N} \setminus \{0\}$. Then the following statements hold:

(i) $R^{m+n} = R^m R^n = R^n R^m$

(ii) $\Delta \subset R \implies R^m \subset R^{m+n}$

Proof. The first equation in (i) clearly holds for $m \in \mathbb{N} \setminus \{0\}$ and $n = 1$. Assume it holds for every $m \in \mathbb{N} \setminus \{0\}$ and some $n \in \mathbb{N} \setminus \{0\}$. Then we obtain, for every $m \in \mathbb{N} \setminus \{0\}$, $R^{m+\sigma(n)} = R^{m+n} R = (R^m R^n) R = R^m (R^n R) = R^m R^{\sigma(n)}$. Thus the first equation is true for every $m, n \in \mathbb{N} \setminus \{0\}$ by the Induction principle, Corollary 3.6. The second equation is a consequence of the first one and the commutativity of addition.
Also (ii) can be shown by the Induction principle. The claim clearly holds for $n = 1$ and every $m \in \mathbb{N}$. Assume it holds for some $n \in \mathbb{N}$ and every $m \in \mathbb{N}$. We have, for every $m \in \mathbb{N}$, $R^m \subset R^{m+n} \subset R^{(m+n)+1} = R^{m+\sigma(n)}$. $\square$

Using arbitrary powers of relations one can construct a pre-ordered space from an arbitrary relational space such that the set is the same and the pre-ordering contains the original relation. The following result also says that the constructed pre-ordering is minimal.

Lemma 3.19
Let (X, R) be a relational space. Then $S = \bigcup\{R^n : n \in \mathbb{N},\, n > 0\}$ is a pre-ordering on X. If T is a pre-ordering on X with $R \subset T$, then $S \subset T$.

Proof. Let $(x, y), (y, z) \in S$. Then there exist $m, n \in \mathbb{N}\backslash\{0\}$ such that $(x, y) \in R^m$ and $(y, z) \in R^n$. Hence $(x, z) \in R^{m+n} \subset S$.
Now let T be a pre-ordering on X with $R \subset T$. Then $R^m \subset T$ for every $m \in \mathbb{N} \setminus \{0\}$ by the Induction principle. It follows that $S \subset T$. □

Lemma and Definition 3.20
There is a unique binary function $\cdot$ on $\mathbb{N}$ such that for every $m, n \in \mathbb{N}$ we have

(i) $m \cdot 0 = 0$

(ii) $m \cdot \sigma(n) = (m \cdot n) + m$

The function $\cdot$ is called **multiplication on** $\mathbb{N}$. For every $m, n \in \mathbb{N}$, the expression $m \cdot n$ is called the **product of** m **and** n, also written $m\,n$. The multiplication on $\mathbb{N}$ is commutative and associative. Furthermore, for every $m, n, p \in \mathbb{N}$, $0 < p$, we have the distributive law

$$(m + n) \cdot p \;=\; m \cdot p + n \cdot p$$

and the following implication holds:

$$m < n \quad\Longrightarrow\quad m \cdot p < n \cdot p$$

We define that, in the absence of brackets, products are evaluated before sums. Thus we may write $m \cdot n + p$ instead of $(m \cdot n) + p$, and $m + n \cdot p$ instead of $m + (n \cdot p)$ without ambiguity.

Proof. The existence and uniqueness of the function follows by Corollary 3.15. In order to prove that the multiplication is commutative and that the distributive law holds, first notice that, for every $m, n \in \mathbb{N}$, we have $\sigma(m)\,n = m\,n + n$.

> [For every $m \in \mathbb{N}$, we have $\sigma(m) \cdot 0 = 0 = m \cdot 0 + 0$. Assuming the claim is true for some $n \in \mathbb{N}$ and every $m \in \mathbb{N}$, we have $\sigma(m)\,\sigma(n) = \sigma(m)\,n + \sigma(m) = m\,n + n + m + 1 = m\,n + m + \sigma(n) = m\,\sigma(n) + \sigma(n)$.]

It follows that $m\,n = n\,m$ for every $m, n \in \mathbb{N}$.

[We show that, for every $p \in \mathbb{N}$, the equation $m\,n = n\,m$ holds for every $m, n \in \mathbb{N}$ with $m + n = p$. Clearly, this is true for $p = 0$ because this implies $m = n = 0$. Assuming it is true for $p \in \mathbb{N}$, let $m, n \in \mathbb{N}$ with $m + n = \sigma(p)$. We may assume $0 < m$. Let q be the predecessor of m, that is $\sigma(q) = m$. Then we have $q + n = p$. It follows that $m\,n = \sigma(q)\,n = q\,n + n = n\,q + n = n\,\sigma(q) = n\,m$.]

It also follows that $(m + n)\,p = m\,p + n\,p$ for every $m, n, p \in \mathbb{N}$.

[For every $m, p \in \mathbb{N}$ this equation is clearly satisfied for $n = 0$. Now assume this equation holds for some $n \in \mathbb{N}$ and for every $m, p \in \mathbb{N}$. We then have $(m + \sigma(n))\;p = \sigma(m + n)\,p = (m + n)\,p + p = m\,p + n\,p + p = m\,p + \sigma(n)\,p$.]

We next show that the multiplication is associative. We have $(m \cdot 0) \cdot p = 0 = m \cdot (0 \cdot p)$ for every $m, p \in \mathbb{N}$. Assume that $(m\,n)\,p = m\,(n\,p)$ for some $n \in \mathbb{N}$ and every $m, p \in \mathbb{N}$. Then $(m\,\sigma(n))\,p = (m\,n + m)\,p = (m\,n)\,p + m\,p = m\,(n\,p) + m\,p = m\,(n\,p + p) = m\,(\sigma(n)\,p)$.
To see the last claim, notice that it clearly holds for $p = 1$ and every $m, n \in \mathbb{N}$. Assume it holds for some $p \in \mathbb{N}$, $0 < p$, and every $m, n \in \mathbb{N}$. Let $m, n \in \mathbb{N}$ and $m < n$. Then $m\,\sigma(p) = m\,p + m < n\,p + n = n\,\sigma(p)$. □

As in the case of addition, we may write multiple products without brackets because of the associativity of the multiplication, i.e. we may write $m\,n\,p$ instead of $(m\,n)\,p$ or $m\,(n\,p)$. Clearly, also the distributive law $p\,(m + n) = p\,m + p\,n$ holds, since addition and multiplication are commutative.

Definition 3.21

Let $m \in \mathbb{N}$. m is called **even** if there is $n \in \mathbb{N}$ such that $m = 2n$, otherwise it is called **odd**. ■

The following Lemma and Corollary are immediate consequences of this Definition.

Lemma 3.22
Let $m \in \mathbb{N}$. If m is odd, then there is $n \in \mathbb{N}$ such that $m = 2n + 1$.

Proof. We prove that for every $m \in \mathbb{N}$ there is $n \in \mathbb{N}$ such that $m = 2n$ or $m = 2n + 1$. This is clearly true for $m = 0$. Assuming it is true for some $m \in \mathbb{N}$, we have either $m + 1 = 2n + 1$ or $m + 1 = 2n + 1 + 1 = 2(n + 1)$. $\square$

Corollary 3.23
Let $m, n, p, q \in \mathbb{N}$ where m and n are even, and p and q are odd. Then $m + n$ and $p + q$ are even, and $m + p$ is odd.

Proof. We may choose $m_0, n_0, p_0, q_0 \in \mathbb{N}$ such that $m = 2m_0$, $n = 2n_0$, $p = 2p_0 + 1$, and $q = 2q_0 + 1$. It follows that $m + n = 2(m_0 + n_0)$, $p + q = 2(p_0 + q_0 + 1)$, and $m + p = 2(m_0 + p_0) + 1$.

$\square$

The following proposition is applied in Section 3.4.

Proposition 3.24
For every $m \in \mathbb{N}$ there is $n \in \mathbb{N}$ such that $2n = m\,\sigma(m)$.

Proof. The claim clearly holds for $m = 0$. Assume that it holds for some $m \in \mathbb{N}$. Let $n \in \mathbb{N}$ such that $2n = m\,\sigma(m)$. Then we have $\sigma(m)\,\sigma(\sigma(m)) = \sigma(m)\,\sigma(m) + \sigma(m) = \sigma(m)\,m + \sigma(m) + \sigma(m) = 2\,(n + \sigma(m))$. $\square$

Lemma and Definition 3.25
There is a unique binary function h on $\mathbb{N}$ such that for every $m, n \in \mathbb{N}$ we have

(i) $h(m, 0) = 1$

(ii) $h(m, \sigma(n)) = h(m, n) \cdot m$

This function is called **exponentiation on** $\mathbb{N}$. We also write m^n for $h(m, n)$. We define that, in the absence of brackets, exponentiation is evaluated before sums or products. Thus, for $m, n, p \in \mathbb{N}$, we may write $m^n + p$ instead of $(m^n) + p$, and $m^n \cdot p$ instead of $(m^n) \cdot p$. We have, for every $m, n, p \in \mathbb{N}$

$$m^{n+p} = m^n\, m^p, \qquad (m^n)^p = m^{n\,p}, \qquad (m\,n)^p = m^p\, n^p$$

and the implication

$$(m < n) \wedge (1 < p) \quad \Longrightarrow \quad (m^p < n^p) \wedge (p^m < p^n)$$

Proof. The existence and uniqueness of the function follows from Corollary 3.15. Further notice that the three equations clearly hold for $p = 0$ and every $m, n \in \mathbb{N}$. Now assume they hold for some $p \in \mathbb{N}$ and every $m, n \in \mathbb{N}$. We then have $m^{n+\sigma(p)} = m^{n+p}\, m = m^n\, m^p\, m = m^n\, m^{\sigma(p)}$, which proves the first equation. To show the second equation, notice that $\left(m^n\right)^{\sigma(p)} = \left(m^n\right)^p m^n = m^{np}\, m^n = m^{np+n} = m^{n\,\sigma(p)}$. Finally, we have $(m\,n)^{\sigma(p)} = (m\,n)^p\; m\,n = m^p\, n^p\, m\,n = m^{\sigma(p)}\, n^{\sigma(p)}$, which proves the third equation.
We now show that $(m < n) \wedge (1 < p)$ implies $m^p < n^p$ for every $m, n, p \in \mathbb{N}$. Let $m, n \in \mathbb{N}$. The implication clearly holds for $p = 2$. Further, assuming it holds for some $p > 1$, it also holds for $\sigma(p)$.
Finally we show that $m < n$ implies $p^m < p^n$ for every $m, n, p \in \mathbb{N}$ where $p > 1$. Let $m, p \in \mathbb{N}$ with $p > 1$. The claim is clearly true for $n = \sigma(m)$. Assuming it holds for some $n \in \mathbb{N}$ with $n \geq \sigma(m)$, we have $p^{\sigma(n)} = p^n\, p > p^m\, p > p^m$. $\square$

Notice that exponentiation is neither associative nor commutative.

3.2 Ordinal numbers

In this Section we introduce the notion of ordinal numbers. Ordinals are a natural extension of the natural numbers. In this text they are used to analyse the Choice axiom in Section 3.3 and as a basis of the concept of cardinality in Section 3.4, which corresponds, in a sense, to the "number of elements" of a set. In order to define the ordinal numbers we begin with the following definition following [Kelley].

Definition 3.26
A set X is called **full** if $A \in X$ implies $A \subset X$. ■

In the literature full sets are also called transitive (cf. [Jech]). However we do not adopt this notion here because we use it as a property that relations may have.
We now define a relation that corresponds to the property of a set to be member of another, but is however restricted to a specified set. Only when restricted to a given set it is a relation according to our Definition 2.1.

Definition 3.27
Given a set X, the relation R on X defined by

$$(x, y) \in R \quad \Longleftrightarrow \quad x \in y$$

is called **element relation on** X. We also write $x \in_X y$ for $(x, y) \in R$. ■

Remark 3.28

Given a set X, the element relation $\in_X$ is irreflexive. If $Y \subset X$, the restriction of $\in_X$ to Y is $\in_Y$. ■

Definition 3.29

A non-empty set X is called **ordinal number** or **ordinal** if X is full and $\in_X$ has the minimum property. We also define that $\emptyset$ is an ordinal. ■

We mainly use Greek letters for ordinals—as we do for positive reals. However, we always say explicitly when a variable is meant to be an ordinal. This property is never automatically implied by the mere usage of a Greek letter.

Lemma 3.30

Every natural number is an ordinal.

Proof. Let $m \in \mathbb{N}$. If $m = 0$, then m is an ordinal by definition. If $0 < m$, it follows by Theorem 3.3 (ii) that m is full. We show by the Induction principle that m has the minimum property for every $m \in \mathbb{N}$, $0 < m$. $\in_m$ clearly has the minimum property if $m = 1$. Now assume that $\in_m$, where $m \in \mathbb{N}$ with $m > 0$, has the minimum property, and let $A \subset \sigma(m)$ with $A \neq \emptyset$. If $A \cap m \neq \emptyset$, the minimum of A is the same as the minimum of $A \setminus \{m\}$. If $A \cap m = \emptyset$, then m is the only member and hence the minimum of A. □

Remark 3.31

Let α be a non-empty ordinal and $\beta \in \alpha$. Then $\beta \subset \alpha$ and, if $\beta \neq \emptyset$, then $\in_\beta$ has the minimum property. ■

Lemma 3.32

If α is a non-empty ordinal, then $\in_\alpha$ is a well-ordering in the sense of "$<$".

Proof. Notice that $\in_\alpha$ is connective by Remark 2.40 since it has the minimum property.
It remains to show that $\in_\alpha$ is transitive. Let $\beta, \gamma, \delta \in \alpha$ with $\beta \in \gamma$ and $\gamma \in \delta$. Since $\in_\alpha$ is connective, we have $\beta = \delta$, $\delta \in \beta$, or $\beta \in \delta$. The first two cases are excluded by Lemma 1.46. □

Lemma 3.33
The following statements hold for ordinals:

(i) If α is an ordinal and $\beta \in \alpha$, then β is an ordinal.

(ii) If α and β are ordinals with $\alpha \subset \beta$ and $\alpha \neq \beta$, then $\alpha \in \beta$.

(iii) If α and β are ordinals, then $\alpha \subset \beta$ or $\beta \subset \alpha$.

(iv) If α is an ordinal, then $\alpha \cup \{\alpha\}$ is an ordinal.

(v) If α is an ordinal, then there is no ordinal β such that $\alpha \in \beta \in \alpha \cup \{\alpha\}$.

Proof. To see (i), assume the stated conditions. If $\beta = \emptyset$, then β is an ordinal. Now assume that $\beta \neq \emptyset$. Then $\in_\beta$ has the minimum property by Remark 3.31. To see that β is full, let $\delta \in \gamma \in \beta$. Since α is full, we have $\gamma, \delta \in \alpha$. It follows that $\delta \in \beta$ because $\in_\alpha$ is transitive by Lemma 3.32. Thus $\gamma \subset \beta$.
To prove (ii) let, under the stated conditions, γ be the minimum of $\beta \setminus \alpha$. We clearly have $\gamma \subset \alpha$. This shows the claim for $\alpha = \emptyset$. Now assume that $\alpha \neq \emptyset$. Let $\delta \in \alpha$. Since $\in_\beta$ is connective, we have either $\gamma \in \delta$ or $\delta \in \gamma$. The case $\gamma \in \delta$ is excluded, because this implies $\gamma \in \alpha$, since α full. It follows that $\alpha \subset \gamma$, and thus $\alpha = \gamma \in \beta$.
To see (iii), notice that for ordinals α and β, $\alpha \cap \beta$ clearly is an ordinal, say γ. By (ii) it follows that $\gamma \in \alpha$ or $\gamma = \alpha$. Similarly, it follows that $\gamma \in \beta$ or $\gamma = \beta$. It is not possible that $\gamma \in \alpha \cap \beta = \gamma$. Hence we have either $\gamma = \alpha$ or $\gamma = \beta$.
(iv) follows by Definition 3.29.

To see (v), assume there are ordinals α and β such that $\alpha \in \beta \in \alpha \cup \{\alpha\}$. It follows that $\alpha \subset \beta \subset \alpha \cup \{\alpha\}$, and thus $\beta = \alpha$ or $\beta = \alpha \cup \{\alpha\}$, which is a contradiction. $\square$

Note that (ii) implies that every ordinal that is not the empty set contains the empty set.

Lemma 3.34
Let A be a non-empty set of ordinals. Then $\in_A$ is a total ordering in the sense of "$<$".

Proof. Let $\alpha, \beta, \gamma \in A$. If $\alpha \in \beta \in \gamma$, then $\beta \subset \gamma$, and therefore $\alpha \in \gamma$. Thus $\in_A$ is transitive.
Moreover $\in_A$ is connective by Lemma 3.33 (ii) and (iii). $\square$

Lemma 3.35
Let A be a non-empty set of ordinals. A has a minimum with respect to the ordering $\in_A$. $\in_A$ is a well-ordering.

Proof. We may choose $\alpha \in A$ such that $\alpha \cap A = \emptyset$ by the Regularity axiom. For every $\beta \in A \setminus \{\alpha\}$, we have either $\alpha \in \beta$ or $\beta \in \alpha$ since $\in_A$ is connective by Lemma 3.34. The latter is a contradiction. Therefore α is a minimum of A.
The second claim is a consequence of the first one. $\square$

Lemma 3.36
Let A be a non-empty set of ordinals. Then $\bigcap A$ and $\bigcup A$ are ordinals too. In particular, $\mathbb{N}$ is an ordinal.

Proof. Exercise. $\square$

Remark 3.37

Notice that the ordering $<$ on $\mathbb{N}$ as defined in Lemma and Definition 3.4 is identical to the well-ordering $\in_{\mathbb{N}}$ on $\mathbb{N}$ when $\mathbb{N}$ is considered as ordinal. ■

Remark 3.38

Let α be an ordinal such that $\alpha \notin \mathbb{N}$ and $\alpha \neq \mathbb{N}$. Then we have $m \in \alpha$ for every $m \in \mathbb{N}$ and $\mathbb{N} \in \alpha$. ■

Lemma 3.39

There is no set that contains every ordinal number.

Proof. Assume A is such a set. We define $\alpha = \bigcup A$ and $\beta = \alpha \cup \{\alpha\}$. α and β are ordinals by Lemmas 3.36 and 3.33 (iv), and thus $\beta \in A$. Moreover we have $A \subset \beta$. It follows that $\beta \in \beta$, which is a contradiction. □

We have seen in Lemma 3.32 that for every ordinal α the relation $\in_\alpha$ is a well-ordering in the sense of "$<$". We now show that well-orderings as defined on ordinals are essentially the only well-orderings in the sense of "$<$" that exist. To this end we first need the notion of isomorphism between two pre-ordered spaces. We then establish some important results between well-orderings that do not explicitly refer to ordinals. Thereby we essentially follow [Jech].

Definition 3.40

Let (X, R) and (Y, S) be ordered spaces, and $f : X \longrightarrow Y$ a map. f is called **order preserving**, if $(x, y) \in R$ implies $\big(f(x), f(y)\big) \in S$. If f is bijective and f as well as f^{-1} are order preserving, then f is called an **order isomorphism**, or short an **isomorphism**. If such an isomorphism exists, the ordered spaces (X, R) and (Y, S) are called **order isomorphic**, or short **isomorphic**. ■

Remark 3.41
Let (X, R) be a totally ordered space and (Y, S) be an ordered space, and $f : X \longrightarrow Y$ a map. If f is bijective and order preserving, then f is an order isomorphism. ■

We recall that the symbol $<$ always denotes an ordering in the sense of "$<$".

Lemma 3.42
Given a well-ordered space $(X, <)$ and an order preserving map $f : X \longrightarrow X$, we have $x < f(x)$ or $x = f(x)$ for every $x \in X$.

Proof. Let y be the minimum of $A = \{x \in X \,:\, f(x) < x\}$. Further let $z = f(y)$. Since f is order preserving, $f(y) < y$ implies $f(z) < z$. Hence $z \in A$ and $z = f(y) < y$, which is a contradiction. □

Corollary 3.43
Given a well-ordered space $(X, <)$, the identity is the only isomorphism from X to itself.

Proof. This follows by Lemma 3.42. □

Corollary 3.44
Given two isomorphic well-ordered spaces $(X, <)$ and $(Y, <)$, there is a unique isomorphism $f : X \longrightarrow Y$.

Proof. This is a direct consequence of Corollary 3.43. □

Corollary 3.45
Let $(X,<)$ be a well-ordered space. Then for every subset $A \subset X$, the restriction of the ordering $<$ to A is a well-ordering in the sense of "$<$" on A, and denoted by $<$ too. There exists no point $x \in X$ such that $(\,]-\infty, x[\,,<)$ and $(X,<)$ are isomorphic.

Proof. Assume there is such a point $x \in X$ and $f : X \longrightarrow\ ]-\infty, x[$ is an isomophism. Then $f(x) < x$, which is a contradiction by Lemma 3.42. □

Proposition 3.46
Let $(X,<)$ and $(Y,<)$ be two well-ordered spaces, $f : X \longrightarrow Y$ an isomorphism, and $x \in X$. Then $f(\,]-\infty, x[\,) = \,]-\infty, f(x)[$, and

$$f \,|\,]-\infty, x[\ : \ ]-\infty, x[\ \longrightarrow \ ]-\infty, f(x)[$$

is an isomorphism.

Proof. Exercise. □

Theorem 3.47
Given two well-ordered spaces $(X,<)$ and $(Y,<)$, exactly one of the following statements is true:

(i) $(X,<)$ and $(Y,<)$ are isomorphic.

(ii) There is $y \in Y$ such that $(X,<)$ and $(\,]-\infty, y[\,,<)$ are isomorphic.

(iii) There is $x \in X$ such that $(\,]-\infty, x[\,,<)$ and $(Y,<)$ are isomorphic.

Proof. Following [Jech] we define the relation f on $X \times Y$ by

$$(x,y) \in f \quad \Longleftrightarrow \quad (\,]-\infty, x[\,,<) \text{ and } (\,]-\infty, y[\,,<) \text{ are isomorphic}$$

Let $D \subset X$ be the domain of f and $R \subset Y$ the range of f. Then f is a bijection from D to R by Corollary 3.45. Let $x_1, x_2 \in D$ with $x_1 < x_2$ and let

$$g : \left]-\infty, x_2\right[\longrightarrow \left]-\infty, f(x_2)\right[$$

be an isomorphism. Then the restriction

$$g \,|\, \left]-\infty, x_1\right[\;:\; \left]-\infty, x_1\right[\longrightarrow \left]-\infty, g(x_1)\right[$$

is an isomorphism by Proposition 3.46. Hence $f(x_1) = g(x_1) < f(x_2)$. Therefore f is order preserving. Thus f is an isomorphism from D to R by Remark 3.41. Clearly at most one of the three statements (i) to (iii) is true by Corollary 3.45 and Proposition 3.46.

Now note that either $D = X$ or $D = \left]-\infty, x\right[$ for some $x \in X$. Similarly, we either have $R = Y$ or $R = \left]-\infty, y\right[$ for some $y \in Y$.

> [If $D \neq X$, then $X \setminus D$ has a minimum, say x. For every $z \in X$ with $z < x$, we have $z \in D$. On the other hand, for every $z \in X$ with $z \geq x$ we have $z \notin D$ by Proposition 3.46. Therefore $D = \left]-\infty, x\right[$. The proof for R is similar.]

Now assume that $D = \left]-\infty, x\right[$ for some $x \in X$, and $R = \left]-\infty, y\right[$ for some $y \in Y$. Since D and R are isomorphic, we have $(x, y) \in f$ by definition of f, which is a contradiction. □

Corollary 3.48

Let α and β be two ordinals such that $(\alpha, <)$ and $(\beta, <)$ are isomorphic. Then $\alpha = \beta$.

Proof. This follows from Lemma 3.33 (iii) and Theorem 3.47.

[Assume that, under the stated conditions, we have $\alpha \subset \beta$ and $\alpha \neq \beta$. Then we have $\alpha \in \beta$ and $\alpha = \;]-\infty, \alpha[$ where the interval refers to the ordering $\in_\beta$. Thus the identity is an isomorphism from α to $]-\infty, \alpha[$. It follows that β and $]-\infty, \alpha[$ are isomorphic, which is a contradiction by Corollary 3.45.] $\square$

We now establish two important results: first it is shown that for any given set, there is a larger ordinal. Second, for any given well-ordered space that is ordered in the sense of "$<$", there is a specific ordinal that is isomorphic. We proceed similiarly to [Ebbinghaus].

Theorem 3.49
Given a set X, there is an ordinal α such that there exists no injection $f : \alpha \longrightarrow X$.

Proof. Assume that X is a set such that for every ordinal α there exists an injection $f : \alpha \longrightarrow X$. We define

$$\mathcal{R} = \big\{ R \subset X \times X \ : \ R \text{ is a well-ordering on (field } R) \big\}$$

Note that, for every ordinal α, there is $A \subset X$ and a well-ordering $R \in \mathcal{R}$ in the sense of "$<$" on A such that $(\alpha, <)$ and (A, R) are isomorphic.

[By assumption there is $A \subset X$ and a bijection $f : \alpha \longrightarrow A$. Then

$$R = \big\{ \big(f(\beta), f(\gamma)\big) \ : \ \beta, \gamma \in \alpha \ \wedge \ \beta < \gamma \big\}$$

is a well-ordering in the sense of "$<$" on A. Moreover $(\alpha, <)$ and (A, R) are isomorphic.]

Now consider the following statement, which we write down partially in the informal language, but we are aware that it could be also written entirely in our formal language with R, X, and z as free variables:

> If $R \in \mathcal{R}$ and if there is an ordinal α such that $(\alpha, <)$ and $(\text{field}\, R, R)$ are isomorphic, then $z = \alpha$, else $z = \emptyset$.

Clearly, for every set R, this statement is true for exactly one set z by Corollary 3.48. Thus, by the Replacement schema 1.47, the following set exists:

$$W = \big\{\alpha : \alpha \text{ is ordinal}, \\ \exists R \in \mathcal{R} \quad (\alpha, <) \text{ and } (\text{field}\, R, R) \text{ are isomorphic}\big\}$$

Now, W contains every ordinal as member by the above result, which is a contradiction to Lemma 3.39. □

Theorem 3.50
For every well-ordered space $(X, <)$ there exists a unique isomorphic ordinal α.

Proof. Let α be an ordinal such that there is no injection $f : \alpha \longrightarrow X$ by Theorem 3.49. By Theorem 3.47 there is $\beta < \alpha$ such that $(\beta, <)$ and $(X, <)$ are isomorphic.
The uniqueness follows by Corollary 3.48. □

The following Theorem is a generalization of the Induction principle, Theorem 3.2 (which applies to the set of natural numbers) to any given ordinal number.

Theorem 3.51 (Induction principle for ordinal numbers)
Let α be an ordinal and $A \subset \alpha$. If, for every ordinal β with $\beta < \alpha$, $\beta \subset A$ implies $\beta \in A$, then $A = \alpha$.

Proof. Assume that, under the stated conditions, $\alpha \setminus A \neq \emptyset$. Let γ be the minimum of $\alpha \setminus A$. It follows that $\gamma \subset A$, and therefore $\gamma \in A$, which is a contradiction. □

We remark that again we take a prudent approach in Theorem 3.51 by formulating an Induction principle for an arbitrary ordinal number but not for the collection of all ordinals. This approach allows us to state the Theorem in terms of sets—while the collection of all ordinals is not a set. One may derive a similar theorem for the class of all ordinals as done in [Jech], or in terms of a formula holding for every ordinal as done in [Ebbinghaus].
The following result is a generalization of the Recursion theorem for natural numbers, Theorem 3.13, to any given ordinal number.

Theorem 3.52 (Local recursion)
Let α be an ordinal, X a set, $D = \bigcup \{X^\beta : \beta < \alpha\}$, and $F : D \longrightarrow X$ a map. There is a unique function $f : \alpha \longrightarrow X$ such that $f(\beta) = F(f \,|\, \beta)$ for every $\beta < \alpha$.

Proof. Notice that if such a function exists, then it is unique.

> [Assume f and g are two such functions and $f \neq g$. Let $A = \{\beta : \beta < \alpha,\ f(\beta) \neq g(\beta)\}$ and $\gamma = \min A$. Then we have $f \,|\, \gamma = g \,|\, \gamma$, and therefore $f(\gamma) = F(f \,|\, \gamma) = F(g \,|\, \gamma) = g(\gamma)$, which is a contradiction.]

Now assume there exists no such function. Let $M \subset \alpha \cup \{\alpha\}$ be the set of those ordinals $\gamma < \alpha \cup \{\alpha\}$ for which there is no function $f : \alpha \longrightarrow X$ such that $f(\beta) = F(f \,|\, \beta)$ for every $\beta < \gamma$. We have $\alpha \in M$ by assumption. Let δ be the minimum of M. Then for every $\varepsilon < \delta$ there exists a function $f : \alpha \longrightarrow X$ such that $f(\xi) = F(f \,|\, \xi)$ for every $\xi < \varepsilon$. For every $\varepsilon < \delta$ we may denote the system of such functions by F_ε. Moreover, if $\varepsilon_1, \varepsilon_2 < \delta$ with $\varepsilon_1 < \varepsilon_2$ or $\varepsilon_1 = \varepsilon_2$, and $f_1 \in F_{\varepsilon(1)}$, $f_2 \in F_{\varepsilon(2)}$, then we have $f_1 \,|\, \varepsilon_1 = f_2 \,|\, \varepsilon_1$ by the result above.
We now distinguish two cases. If δ has no predecessor, then we may choose a

point $x \in X$ and define the map $g : \alpha \longrightarrow X$ by

$$g(\xi) = \begin{cases} f(\xi) & \text{if} \quad \xi < \delta, \quad \text{where } f \in F_\varepsilon,\ \xi < \varepsilon < \delta \\ x & \text{if} \quad \xi = \delta \ \text{ or } \ \delta < \xi \end{cases}$$

Now, if δ has a predecessor, say ε, we may choose $x \in X$, $f \in F_\varepsilon$, and define $g : \alpha \longrightarrow X$ by

$$g(\xi) = \begin{cases} f(\xi) & \text{if} \quad \xi < \varepsilon \\ F(f \mid \varepsilon) & \text{if} \quad \xi = \varepsilon \\ x & \text{if} \quad \varepsilon < \xi \end{cases}$$

In both cases we find that $g(\xi) = F(g \mid \xi)$ for every $\xi < \delta$, which is a contradiction.

□

3.3 Choice

Several important consequences of the Choice axiom, Axiom 1.49, that are required subsequently are proven in this Section. Since our aim is not an exhaustive discussion of ZFC, but to explain its consequences for the foundations of analysis, we do not prove the equivalence of the different forms of the Choice axiom neither provide a comprehensive list of known equivalent forms. Instead, we always assume the Choice axiom and derive some of its relevant implications. Remember that the Choice axiom implies the existence of a choice function as stated in Lemma and Definition 1.50 and Remark 2.66.

Theorem 3.53 (Well-ordering principle)
For every set X there exists a well-ordering on X.

Proof. We may choose an ordinal β such that there is no injection from β to X by Theorem 3.49. Moreover, let $g : \mathcal{P}(X)\backslash\{\emptyset\} \longrightarrow X$ a choice function and y a point such that $y \notin X$. Further let

$$Y = X \cup \{y\}\,, \quad D = \bigcup\{Y^\gamma : \gamma < \beta\}$$

We define the map $F : D \longrightarrow Y$ by

$$F(h) = \begin{cases} g\big(X\backslash \operatorname{ran} h\big) & \text{if} \quad X \not\subset \operatorname{ran} h \\ y & \text{if} \quad X \subset \operatorname{ran} h \end{cases}$$

By Theorem 3.52 there exists a unique function $f : \beta \longrightarrow Y$ such that $f(\gamma) = F(f\,|\,\gamma)$ for every $\gamma < \beta$.

$f^{-1}[X]$ is an ordinal, since it is full and the relation $<$ on this set has the minimum property by Lemma 3.35.

> [Let $\delta < \gamma \in f^{-1}[X]$. Then $f(\gamma) \in X$, and thus $X \not\subset f[\gamma]$. It follows that $X \not\subset f[\delta]$. Therefore $f(\delta) \in X$.]

We define $\alpha = f^{-1}[X]$.

For every $x \in X$ the set $f^{-1}\{x\}$ is a singleton or empty.

> [Assume $\gamma, \delta \in f^{-1}\{x\}$ with $\gamma < \delta$. Then we have $f(\delta) \in X$, and thus $f(\delta) = g\big(X\backslash f[\delta]\big) \neq x$, which is a contradiction.]

Moreover $X \subset f[\beta]$.

> [Assume there is $x \in X$ such that $x \notin f[\beta]$. Let $\gamma < \beta$. Then $X \not\subset f[\gamma]$, and therefore $f(\gamma) = g\big(X\backslash f[\gamma]\big) \in X$. Thus f is an injection that maps β on X, which is a contradiction.]

It follows that the map $t : \alpha \longrightarrow X$, $t = f\,|\,\alpha$ is bijective.

We now define a relation $<$ on X by

$$x < y \iff t^{-1}\{x\} < t^{-1}\{y\}$$

This relation is clearly a well-ordering on X. $\square$

Notice that Theorem 3.53 and Lemma 2.43 imply that for every set X there exists a well-ordering in the sense of "$<$" and a well-ordering in the sense of "$\leq$". Theorem 3.53 is used to prove the pseudo-metrization theorem.

Corollary 3.54
For every set X there exists an ordinal α and a bijection $t : \alpha \longrightarrow X$.

Proof. Such a bijection is explicitily constructed in the proof of Theorem 3.53. $\square$

The next result is another important implication of the Choice axiom. Our proof does not make use of the Well-ordering principle.

Theorem 3.55 (Zorn's Lemma)
Let (X, R) be an ordered space. If every chain has an upper bound, then X has a weak maximum.

Proof. We may choose an ordinal β such that there is no injection from β to X by Theorem 3.49. Let $g : \mathcal{P}(X)\backslash\{\emptyset\} \longrightarrow X$ be a choice function, and $\mathcal{A} \subset \mathcal{P}(X)$ the system of all chains. Moreover let

$$H : \mathcal{A} \longrightarrow \mathcal{P}(X), \quad H(A) = \big\{z \in X\backslash A \ : \ A \cup \{z\} \in \mathcal{A}\big\}$$

Further let

$$x \in X, \quad y \notin X, \quad Y = X \cup \{y\}, \quad D = \bigcup \{Y^\gamma : \gamma < \beta\},$$

We define the map $F : D \longrightarrow Y$ by

$$F(h) = \begin{cases} x & \text{if} \quad h = \emptyset \\ g\big(H(\operatorname{ran} h)\big) & \text{if} \quad h \neq \emptyset, \ \operatorname{ran} h \in \mathcal{A}, \ H(\operatorname{ran} h) \neq \emptyset \\ y & \text{if} \quad h \neq \emptyset, \ \operatorname{ran} h \in \mathcal{A}, \ H(\operatorname{ran} h) = \emptyset \\ y & \text{if} \quad h \neq \emptyset, \ \operatorname{ran} h \notin \mathcal{A} \end{cases}$$

By the Local recursion theorem there is a unique function $f : \beta \longrightarrow Y$ such that $f(\gamma) = F(f \mid \gamma)$ for every $\gamma < \beta$. We define $\alpha = f^{-1}[X]$. α is an ordinal since it is full and the relation $<$ is a well-ordering on α by Lemma 3.35.

[Let $\delta < \gamma < \alpha$. Then $f(\gamma) \in X$, and thus $f[\gamma] \in \mathcal{A}$ and $H(f[\gamma]) \neq \emptyset$. It follows that $f[\delta] \in \mathcal{A}$ and $H(f[\delta]) \neq \emptyset$. Therefore $f(\delta) \in X$.]

For every $z \in X$ the set $f^{-1}\{z\}$ is a singleton or empty.

[Assume $\gamma, \delta \in f^{-1}\{z\}$ with $\gamma < \delta$. Then we have $f(\delta) \in X$, and thus $f(\delta) = g\big(H(f[\delta])\big) \neq z$, which is a contradiction.]

Therefore we have $\alpha < \beta$ by the choice of β. We define $B = f[\alpha]$. Then B is a chain.

[Let $\gamma, \delta \in \alpha$ with $\gamma < \delta$. Then $f(\delta) = g\big(H(f[\delta])\big)$, and therefore $f[\delta] \cup \{f(\delta)\}$ is a chain. We have $\gamma \in \delta$, and thus $f(\gamma) \in f[\delta]$.]

Furthermore, $H(B) = \emptyset$.

[Assume that $H(B) \neq \emptyset$. It follows that $f(\alpha) = g(H(B)) \in X$, and hence $\alpha \in \alpha$, which is a contradiction.]

Let $b \in X$ be an upper bound of B, which exists by assumption. Then b is a weak maximum of X.

[Assume that $(b, c) \in R$. Then $B \cup \{b, c\}$ is a chain, and thus $b, c \in B$. It follows that $(c, b) \in R$.]

$\square$

Finally we generalize Zorn's Lemma to pre-ordered spaces, which is applied in the proof of the existence of an ultrafilter base finer than a given filter base in Theorem 5.72.

Theorem 3.56
Let (X, R) be a pre-ordered space. If every chain has an upper bound, then X has a weak maximum.

Proof. Assume the stated condition. Let (Y, S) be the ordered space constructed from (X, R) as in Lemma 2.22. Every chain in Y has an upper bound.

[Let $A \subset Y$ be a chain. Then $B = \bigcup A$ is chain too. Therefore B has an upper bound, say x. Then $[x]$ is an upper bound of A.]

Let b be the weak maximum of Y by Theorem 3.55. Then every point $x \in b$ is a weak maximum of X (exercise). $\square$

3.4 Cardinality

In this Section we define several notions in order to describe what could be considered as the size of a set.

Definition 3.57
Two sets X and Y are said to be **of the same cardinality**, also written $X \sim Y$, if there exists a bijection $f : X \longrightarrow Y$. Let X be a set. X is called **finite** if there is $m \in \mathbb{N}$ such that $m \sim X$, else it is called **infinite**. If $X \sim \mathbb{N}$ or X is finite, then X is called **countable**. If X is not countable, it is called **uncountable**. ■

Remark 3.58
For every set X there is an ordinal α such that $X \sim \alpha$ by Corollary 3.54. ■

Lemma 3.59
Let X be a set. If X is infinite, then for every $m \in \mathbb{N}$ there is a subset $Y \subset X$ such that $Y \sim m$. If X is uncountable, then there is a subset $Y \subset X$ such that $Y \sim \mathbb{N}$.

Proof. This follows by Remarks 3.38 and 3.58. □

Lemma 3.60
Let $(X, \prec)$ be a connective pre-ordered space and $A \subset X$ where A is finite and $A \neq \emptyset$. Then A has a minimum and a maximum. If $\prec$ is a total ordering, then the minimum and the maximum of A are unique.

Proof. If $A \sim 1$, then A is a singleton. Thus it has a unique minimum and a unique maximum.
To see the first claim, assume that every $A \subset X$ with $A \sim m$ for some $m \in \mathbb{N}$, $m > 0$, has a minimum and a maximum. Let $B \subset X$ with $B \sim \sigma(m)$. There is a bijection $f : \sigma(m) \longrightarrow B$. By assumption $B \setminus \{f(m)\}$ has a minimum, say x, and a maximum, say y. Then the set $\{x, f(m)\}$ has a minimum, which is a minimum of B, and the set $\{y, f(m)\}$ has a maximum, which is a maximum of B.
The second claim follows by Remark 2.37. □

Corollary 3.61
Let X be a set. If $X \sim \mathbb{N}$, then X is infinite.

Proof. It is enough to show that $\mathbb{N}$ is infinite. Let $m \in \mathbb{N}$, $m > 0$, and $f : m \longrightarrow \mathbb{N}$ be an injection. Then the set $f[m]$ is finite and not empty, and therefore has a unique maximum by Lemma 3.60. Thus f is not surjective. $\square$

Proposition 3.62
Let X be a finite non-empty set, $x \in X$, and $m \in \mathbb{N}$. Then $X \sim \sigma(m)$ implies $X \setminus \{x\} \sim m$.

Proof. Exercise. $\square$

Lemma 3.63
Let X and Y be finite sets and $Z \subset X$. Then Z, $X \cup Y$, and $X \times Y$ are finite.

Proof. We first show that Z is finite. This is clearly true if $X \sim 0$. Now let $m \in \mathbb{N}$. Assume that the claim is true for every X with $X \sim m$. Let U be a set with $U \sim \sigma(m)$. Further let $u \in U$, $V = U \setminus \{u\}$, and $Z \subset U$. Then we have $V \sim m$ by Proposition 3.62, and hence $Z \cap V$ is finite by assumption. If $Z \subset V$, then Z is finite. If $Z \not\subset V$, we may choose $n \in \mathbb{N}$ and a bijection $g : n \longrightarrow Z \cap V$. We define a bijection $h : \sigma(n) \longrightarrow Z$ by $h \,|\, n = g$ and $h(n) = u$, and therefore Z is finite.
Next we prove that $X \cup Y$ is finite. Since in the case $X = Y = \emptyset$ the claim is obvious, we show that $X \sim m$ and $Y \sim n$, where $m, n \in \mathbb{N}$, $0 < m$, implies that there is an injection $h : X \cup Y \longrightarrow m + n$. Let $m \in \mathbb{N}$, $0 < m$. Then this implication is clearly true for $n = 0$. Assuming that it is true for some $n \in \mathbb{N}$, let X and Y be sets with $X \sim m$, $Y \sim \sigma(n)$. Further let $y \in Y$, and $V = Y \setminus \{y\}$. Then we have $V \sim n$, and there is an injection $f : X \cup V \longrightarrow m+n$ by assumption.

We define the injection $h : X \cup Y \longrightarrow m+\sigma(n)$ by $h \,|\, (X \cup V) = f$ and, if $y \notin X$, then $h(y) = m + n$.
To see that $X \times Y$ is finite, first notice that this is clear if $X = \emptyset$ or $Y = \emptyset$. Now assume $X \neq \emptyset$ and $Y \neq \emptyset$. We show that $X \sim m$ and $Y \sim n$ $(m, n \in \mathbb{N} \setminus \{0\})$ implies that $X \times Y \sim m\,n$. Let $m \in \mathbb{N}$, $0 < m$. Then this implication is clearly true for $n = 1$. Assuming that it is true for some $n \in \mathbb{N}$, $0 < n$, let X and Y be sets with $X \sim m$, $Y \sim \sigma(n)$. Further let $y \in Y$, $V = Y \setminus \{y\}$, and $h : m \longrightarrow X$ a bijection. Then there is a bijection $f : X \times V \longrightarrow m\,n$ by assumption. We define a bijection $s : X \times Y \longrightarrow m\,\sigma(n)$ by $s \,|\, (X \times V) = f$ and $s\big((h(k), y)\big) = m\,n + \sigma(k)$ for every $k \in \mathbb{N}$, $k < m$. □

The following Lemma extends the results of Lemma 3.63 to arbitrary finite unions and products.

Lemma 3.64
Let I be a finite index set, and for every $i \in I$, let X_i be a finite non-empty set. Then $\bigcup_{i \in I} X_i$ and $\bigtimes_{i \in I} X_i$ are finite.

Proof. We first show that $\bigcup_{i \in I} X_i$ is finite. This is clear if $I \sim 1$. Now let $m \in \mathbb{N}$, $0 < m$, and assume that the claim holds for every index set I with $I \sim m$. Let J be a set with $J \sim \sigma(m)$. Further let $k \in J$, $K = J \setminus \{k\}$, and X_j a finite non-empty set for every $j \in J$. Then $K \sim m$ by Proposition 3.62, and $\bigcup_{j \in K} X_j$ is finite by assumption. It follows that $\bigcup_{j \in J} X_j = \bigcup_{j \in K} X_j \cup X_k$ is finite by Lemma 3.63.
We now prove that $\bigtimes_{i \in I} X_i$ is finite. The claim clearly holds if $I \sim 1$. Now let $m \in \mathbb{N}$, $0 < m$, and assume that the claim holds for every index set I with $I \sim m$. Let J be a set with $J \sim \sigma(m)$. Further let $k \in J$, $K = J \setminus \{k\}$, and X_j a finite non-empty set for every $j \in J$. Then $\bigtimes_{j \in K} X_j$ is finite by assumption. It follows that $\bigtimes_{j \in K} X_j \times X_k$ is finite by Lemma 3.63. Thus $\bigtimes_{j \in J} X_j$ is finite

by Remark 2.72. □

Remark 3.65

Let X and Y be finite non-empty sets. Then X^Y is finite. ■

Proposition 3.66

Let X be a countable set. There is a set Y such that $X \subset Y$ and $Y \sim \mathbb{N}$.

Proof. If X is infinite, the claim clearly holds. If X is finite, then there is $m \in \mathbb{N}$ and a bijection $f : m \longrightarrow X$. If the set $X \cap \mathbb{N}$ is non-empty, then let n be its unique maximum by Lemma 3.60. In this case we may choose a number $p \in \mathbb{N}$ such that $n < p$ and $m < p$. If $X \cap \mathbb{N}$ is empty, then let $p = \sigma(m)$. We define $Y = X \cup \{q \in \mathbb{N} : q \geq p\}$ and the map $g : \mathbb{N} \longrightarrow Y$ by

$$g(r) = \begin{cases} f(r) & \text{if } r < m \\ p & \text{if } r = m \\ p + s & \text{if } r > m \end{cases}$$

where, in the last case, s is the number such that $m + s = r$ by Lemma 3.17. Then g is bijective. □

Lemma 3.67

A set X is countable iff there is an injection $f : X \longrightarrow \mathbb{N}$.

Proof. Assume there is an injection $f : X \longrightarrow \mathbb{N}$. Let $A = \operatorname{ran} f$. If A is finite, then X is finite. If A is infinite, we may recursively define the function $g : \mathbb{N} \longrightarrow A$ by Lemma 3.13 and Lemma 3.36 as follows: Let $g(0)$ be the minimum of A, and, for every $m \in \mathbb{N}$, let $g(\sigma(m))$ be the minimum of $\{n \in A : n > g(m)\}$. Then g is strictly increasing.

[We show by the Induction principle that g is strictly increasing, i.e. $m < n$ implies $g(m) < g(n)$ for every $m, n \in \mathbb{N}$. Let $m \in \mathbb{N}$. Then the claim clearly holds for $n = \sigma(m)$. Assume it holds for some $n \in \mathbb{N}$ with $m < n$. Then $g(\sigma(n)) > g(n)$, and therefore the claim also holds for $\sigma(n)$.]

Since g is strictly increasing, it is injective. To see that g is surjective, assume there exists $m \in A \setminus \operatorname{ran} g$. Then we have $\operatorname{ran} g \subset m$.

[Let $B = \{n \in \mathbb{N} : g(n) > m\}$, and assume that $B \neq \emptyset$. Let p be the minimum of B by Lemma 3.36. We have $p > 0$ by the definition of $g(0)$. Let q be the predecessor of p. Then $g(q) < m$. It follows that $g(\sigma(q)) \leq m$ since $m \in A$, which is a contradiction.]

Hence $\operatorname{ran} g$ is finite, which is a contradiction. Thus g is bijective, and therefore X is countable.
The converse follows by Theorem 3.3 (i). □

Proposition 3.68
We have $\mathbb{N}^2 \sim \mathbb{N}$.

Proof. We define the functions

$$s : \mathbb{N} \longrightarrow \mathbb{N}\ , \qquad s(m) = m(m+1)/2\ ;$$

$$h : \mathbb{N}^2 \longrightarrow \mathbb{N}\ , \quad h(p,q) = s(p+q) + q$$

Note that s is well-defined by Proposition 3.24.
We first show that h is injective. For given $r \in \mathbb{N}$ there is at most one pair $(m,q) \in \mathbb{N}^2$ such that $q \leq m$ and $s(m) + q = r$.

[For $i \in \{1, 2\}$, assume that $m_i, q_i \in \mathbb{N}$ with $q_i \leq m_i$ and $s(m_i) + q_i = r$. Further we may assume that $m_1 < m_2$. Since s is increasing, we have

$$s(m_2) + q_1 \geq s(m_1 + 1) + q_1 = s(m_1) + m_1 + 1 + q_1 = r + m_1 + 1 > r + q_1$$

It follows that $s(m_2) > r$, which is a contradiction.]

Thus there is at most one pair $(p, q) \in \mathbb{N}^2$ such that $s(p + q) + q = r$.
Thus $\mathbb{N}^2$ is countable by Lemma 3.67. This set is clearly not finite, so it is of the same cardinality than $\mathbb{N}$. □

Lemma 3.69
Let X and Y be countable sets and $Z \subset X$. Then Z and $X \times Y$ are countable.

Proof. Z is countable by Lemma 3.67.
In order to show that $X \times Y$ is countable, we may choose two injections $f : X \longrightarrow \mathbb{N}$ and $g : Y \longrightarrow \mathbb{N}$ by Lemma 3.67, and an injection $h : \mathbb{N}^2 \longrightarrow \mathbb{N}$ by Proposition 3.68. We define the function

$$t : X \times Y \longrightarrow \mathbb{N}, \quad t(x, y) = h\big(f(x), g(y)\big)$$

t is clearly an injection. The claim follows by Lemma 3.67. □

Lemma 3.70
Let I be a countable index set, and for each $i \in I$ let X_i be a countable set. Then $X = \bigcup_{i \in I} X_i$ is countable.

Proof. We may assume that $I \sim \mathbb{N}$ and $X_i \sim \mathbb{N}$ for every $i \in I$.

[If I is countable, then there is a set J and a bijection $f : J \longrightarrow \mathbb{N}$ such that $I \subset J$ by Proposition 3.66. For every $j \in J$, we may choose a set Y_j and a bijection $f_j : Y_j \longrightarrow \mathbb{N}$ such that $X_j \subset Y_j$ $(j \in I)$ by the same Proposition. Then $X \subset Y$ where $Y = \bigcup_{j \in J} Y_j$. If there is an injection $t : Y \longrightarrow \mathbb{N}$, then $t \,|\, X$ is an injection.]

Let $g : \mathbb{N} \longrightarrow I$ and $g_i : \mathbb{N} \longrightarrow X_i$ $(i \in I)$ be bijections. We may choose a bijection $h : \mathbb{N} \longrightarrow \mathbb{N}^2$ by Lemma 3.69. For $i \in \{1, 2\}$, let $p_i : \mathbb{N}^2 \longrightarrow \mathbb{N}$ be the projections on the coordinates. We define

$$G : \mathbb{N} \longrightarrow X, \qquad G(m) = g_{g\, p_1 h(m)}\big(p_2\, h(m)\big)$$

G is clearly surjective. Let $H : X \longrightarrow \mathbb{N}$ where, for every $x \in X$, $H(x)$ is the minimum of $G^{-1}\{x\}$, which exists since $<$ is a well-ordering on $\mathbb{N}$ by Remark 3.37. Then H is injective. The claim follows by Lemma 3.67. $\square$

Lemma 3.71
Let X be a countable non-empty set and $\mathcal{A} = \{A \subset X \,:\, A \text{ is finite}\}$. Then $\mathcal{A}$ is countable.

Proof. We have $\mathcal{A} = \bigcup_{m \in \mathbb{N}} \mathcal{A}_m$ where $\mathcal{A}_m = \{A \subset X \,:\, A \sim m\}$ $(m \in \mathbb{N})$.
First we show by the Induction principle that, for every $m \in \mathbb{N}$, $\mathcal{A}_m$ is countable. Clearly, $\mathcal{A}_0$ is finite. Now assume that $\mathcal{A}_m$ is countable for some $m \in \mathbb{N}$. We have

$$\begin{aligned} \mathcal{A}_{\sigma(m)} &= \big\{A \subset X \,:\, \exists x \in A \quad A \backslash \{x\} \in \mathcal{A}_m\big\} \\ &\subset \bigcup_{x \in X} \big\{A \cup \{x\} \,:\, A \in \mathcal{A}_m\big\} \end{aligned}$$

which is countable by Lemma 3.70.
Now it follows that $\mathcal{A}$ is countable by the same Lemma. $\square$

Chapter 4

Numbers II

4.1 Positive dyadic rational numbers

In this Section we define the positive dyadic rational numbers, and in the next Section the positive real numbers. It is then possible to construct the full system of real numbers from its positive counterpart. The set of (positive and negative) integers and the set of (positive and negative) dyadic rationals can finally be identified with subsets of the reals.

Lemma and Definition 4.1
We define an equivalence relation Q on $\mathbb{N}^2$ by

$$\big((m,u),(n,v)\big) \in Q \quad \Longleftrightarrow \quad m\,2^v = n\,2^u$$

and $\mathbb{D}_+ = \mathbb{N}^2/Q$. For every $m, u \in \mathbb{N}$, the equivalence class of (m,u) is denoted by $\lfloor m,u \rfloor$. The members of $\mathbb{D}_+$ are called **positive dyadic rational numbers**. Furthermore, we define the relation $<$ on $\mathbb{D}_+$ as follows:

$$\lfloor m,u \rfloor < \lfloor n,v \rfloor \quad \Longleftrightarrow \quad m\,2^v < n\,2^u$$

This is a total ordering in the sense of "$<$" on $\mathbb{D}_+$. Moreover, we define $\leq$ to be the total ordering in the sense of "$\leq$" on $\mathbb{D}_+$ obtained from the ordering $<$ by the method of Lemma 2.19.

Proof. Clearly, Q is an equivalence relation. To see that the relation $<$ on $\mathbb{D}_+$ is well defined, let $\lfloor m,u \rfloor < \lfloor n,v \rfloor$, and $(p,w) \in \lfloor m,u \rfloor$, $(q,r) \in \lfloor n,v \rfloor$. Then we have $p\,2^u = m\,2^w$ and $q\,2^v = n\,2^r$. It follows that

$$p\,2^{u+v+r} = m\,2^{w+v+r} < n\,2^{u+w+r} = q\,2^{v+u+w}$$

and thus $p\,2^r < q\,2^w$. To see that it is transitive let $\lfloor m,u \rfloor, \lfloor n,v \rfloor, \lfloor p,w \rfloor \in \mathbb{D}_+$ with $\lfloor m,u \rfloor < \lfloor n,v \rfloor < \lfloor p,w \rfloor$. Then we have $m\,2^v < n\,2^u$ and $n\,2^w < p\,2^v$. It follows that $m\,2^{v+w} < n\,2^{u+w} < p\,2^{v+u}$, and therefore $m\,2^w < p\,2^u$. Thus we obtain $\lfloor m, u \rfloor < \lfloor p, w \rfloor$. Moreover, it is obviously antireflexive and connective. □

Definition 4.2
We adopt the convention that all notions related to orderings on $\mathbb{D}_+$, in particular intervals, refer to the ordering in the sense of "$<$" as defined in Lemma and Definition 4.1 unless otherwise specified. ■

Notice that this convention agrees with the one in the context of natural numbers, cf. Definition 3.5. Note again that in many cases it is irrelevant whether the ordering $<$ or the ordering $\leq$ on $\mathbb{D}_+$ is considered as most order properties are invariant, cf. Lemmas 2.34, 2.46, and 2.78. However, the choice of the ordering *is* relevant for intervals.

Corollary 4.3
The set $\mathbb{D}_+$ is countable.

Proof. This follows from Lemma 3.69. □

Lemma 4.4
$\mathbb{D}_+$ is $<$-dense.

Proof. Let $\lfloor m, u \rfloor, \lfloor n, v \rfloor \in \mathbb{D}_+$ with $\lfloor m, u \rfloor < \lfloor n, v \rfloor$. We define $m_0 = m\, 2^{v+1}$, and $n_0 = n\, 2^{u+1}$. Then we have $\lfloor m, u \rfloor = \lfloor m_0, u + v + 1 \rfloor$ and $\lfloor n, v \rfloor = \lfloor n_0, u + v + 1 \rfloor$. Since m_0 and n_0 are even, there is $k \in \mathbb{N}$ such that $m_0 < k < n_0$. Thus $\lfloor m, u \rfloor < \lfloor k, u + v + 1 \rfloor < \lfloor n, v \rfloor$. □

Lemma and Definition 4.5
Let $+$ be the binary function on $\mathbb{D}_+$ defined by

$$\lfloor m, u \rfloor + \lfloor n, v \rfloor = \lfloor m\, 2^v + n\, 2^u,\ u + v \rfloor$$

This function is called **addition on** $\mathbb{D}_+$. The expression $\lfloor m, u \rfloor + \lfloor n, v \rfloor$ is called the **sum of** $\lfloor m, u \rfloor$ **and** $\lfloor n, v \rfloor$. $+$ is commutative and associative. Moreover, for every $d, e, f \in \mathbb{D}_+$, we have

$$d < e \quad \Longrightarrow \quad d + f < e + f$$

Proof. To see that $+$ is well-defined, let $\lfloor m, u \rfloor = \lfloor p, w \rfloor \in \mathbb{D}_+$ and $\lfloor n, v \rfloor = \lfloor q, r \rfloor \in \mathbb{D}_+$. Then we have

$$\begin{aligned} \lfloor m, u \rfloor + \lfloor n, v \rfloor &= \lfloor m\, 2^v + n\, 2^u,\ u + v \rfloor \\ &= \lfloor m\, 2^{v+w+r} + n\, 2^{u+w+r},\ u + v + w + r \rfloor \\ &= \lfloor p\, 2^{v+u+r} + q\, 2^{u+w+v},\ u + v + w + r \rfloor \\ &= \lfloor p\, 2^r + q\, 2^w,\ w + r \rfloor \ = \ \lfloor p, w \rfloor + \lfloor q, r \rfloor \end{aligned}$$

$+$ is clearly commutative.
Now let $\lfloor m, u \rfloor, \lfloor n, v \rfloor, \lfloor p, w \rfloor \in \mathbb{D}_+$. To see that $+$ is associative, notice that

$$\begin{aligned} \big(\lfloor m, u \rfloor + \lfloor n, v \rfloor\big) + \lfloor p, w \rfloor &= \lfloor m\, 2^v + n\, 2^u,\ u + v \rfloor + \lfloor p, w \rfloor \\ &= \lfloor m\, 2^{v+w} + n\, 2^{u+w} + p\, 2^{u+v},\ u + v + w \rfloor \\ &= \lfloor m, u \rfloor + \big(\lfloor n, v \rfloor + \lfloor p, w \rfloor\big) \end{aligned}$$

To see the last assertion, notice that

$$\lfloor m, u \rfloor + \lfloor p, w \rfloor \ = \ \lfloor m\, 2^w + p\, 2^u,\ u + w \rfloor \ = \ \lfloor m\, 2^{v+w} + p\, 2^{u+v},\ u + v + w \rfloor$$

and

$$\lfloor n, v \rfloor + \lfloor p, w \rfloor \ = \ \lfloor n\, 2^w + p\, 2^v,\ v + w \rfloor \ = \ \lfloor n\, 2^{u+w} + p\, 2^{u+v},\ u + v + w \rfloor$$

□

As in the case of natural numbers, associativity of $+$ for positive dyadic rationals allows one to write multiple sums without brackets, i.e. for every $d, e, f \in \mathbb{D}_+$ we may write $d + e + f$ instead of $(d + e) + f$ or $d + (e + f)$, and similarly for sums of more than three terms.

Proposition 4.6
Let $d, e \in \mathbb{D}_+$. If $d < e$, then there is $f \in \mathbb{D}_+$ such that $d + f = e$.

Proof. Let $\lfloor m, u \rfloor, \lfloor n, v \rfloor \in \mathbb{D}_+$ with $\lfloor m, u \rfloor < \lfloor n, v \rfloor$. There is $p \in \mathbb{N}$ such that $m\,2^v + p = n\,2^u$ by Lemma 3.17. It follows that

$$\begin{aligned} \lfloor m, u \rfloor + \lfloor p,\ u + v \rfloor &= \lfloor m\,2^v,\ u + v \rfloor + \lfloor p,\ u + v \rfloor \\ &= \lfloor m\,2^v + p,\ u + v \rfloor \\ &= \lfloor n\,2^u,\ u + v \rfloor \ = \ \lfloor n, v \rfloor \end{aligned}$$

□

Lemma and Definition 4.7
Let $\cdot$ be the binary function on $\mathbb{D}_+$ defined by $\lfloor m,u \rfloor \cdot \lfloor n,v \rfloor = \lfloor m\,n,\ u+v \rfloor$. This function is called **multiplication on** $\mathbb{D}_+$. The expression $\lfloor m,u \rfloor \cdot \lfloor n,v \rfloor$ is called the **product of** $\lfloor m,u \rfloor$ **and** $\lfloor n,v \rfloor$. For every $a,b \in \mathbb{D}_+$ we also write $a\,b$ for $a \cdot b$. $\cdot$ is commutative and associative. For every $d,e,f \in \mathbb{D}_+$ the following distributive law holds:

$$(d+e)\cdot f \;=\; (d\cdot f)+(e\cdot f)$$

Moreover, we have

$$(d<e) \wedge (0<f) \quad\Longrightarrow\quad d\cdot f < e\cdot f$$

We define that in the absence of brackets products are evaluated before sums.

Proof. To see that $\cdot$ is well-defined, let $\lfloor m,u \rfloor = \lfloor p,w \rfloor \in \mathbb{D}_+$ and $\lfloor n,v \rfloor = \lfloor q,r \rfloor \in \mathbb{D}_+$. Then we have

$$\begin{aligned}
\lfloor m,u \rfloor \cdot \lfloor n,v \rfloor &= \lfloor m\,n,\ u+v \rfloor \\
&= \lfloor m\,2^w\,n\,2^r,\ u+v+w+r \rfloor \\
&= \lfloor p\,2^u\,q\,2^v,\ u+v+w+r \rfloor \\
&= \lfloor p\,q,\ w+r \rfloor \;=\; \lfloor p,w \rfloor \cdot \lfloor q,r \rfloor
\end{aligned}$$

$\cdot$ is clearly commutative.
Let $\lfloor m,u \rfloor, \lfloor n,v \rfloor, \lfloor p,w \rfloor \in \mathbb{D}_+$. To see that $\cdot$ is associative, notice that

$$\begin{aligned}
\big(\lfloor m,u \rfloor \cdot \lfloor n,v \rfloor\big) \cdot \lfloor p,w \rfloor &= \lfloor m\,n\,p,\ u+v+w \rfloor \\
&= \lfloor m,u \rfloor \cdot \big(\lfloor n,v \rfloor \cdot \lfloor p,w \rfloor\big)
\end{aligned}$$

The distributive law is seen by the following calculation:

$$\begin{aligned}\big(\lfloor m,u\rfloor + \lfloor n,v\rfloor\big)\cdot\lfloor p,w\rfloor &= \lfloor m\,2^v + n\,2^u,\ u+v\rfloor\cdot\lfloor p,w\rfloor \\ &= \lfloor m\,p\,2^v + n\,p\,2^u,\ u+v+w\rfloor \\ &= \lfloor m\,p\,2^{v+w} + n\,p\,2^{u+w},\ u+v+2w\rfloor \\ &= \lfloor m\,p,\ u+w\rfloor + \lfloor n\,p,\ v+w\rfloor \\ &= \lfloor m,u\rfloor\cdot\lfloor p,w\rfloor + \lfloor n,v\rfloor\cdot\lfloor p,w\rfloor\end{aligned}$$

Now assuming that $\lfloor m,u\rfloor < \lfloor n,v\rfloor$, we have $m\,2^v < n\,2^u$ by definition. Thus $m\,p\,2^{v+w} < n\,p\,2^{u+w}$, and therefore $\lfloor m\,p,\ u+w\rfloor < \lfloor n\,p,\ v+w\rfloor$. This shows the asserted implication. $\square$

Again, associativity allows one to write multiple products without brackets.

Lemma 4.8
Let $g : \mathbb{N} \longrightarrow \mathbb{D}_+$, $g(m) = \lfloor m,0\rfloor$. Then g is injective. For every $m,n \in \mathbb{N}$ we have

(i) $m < n \iff g(m) < g(n)$

(ii) $g(m+n) = g(m) + g(n)$

(iii) $g(m\cdot n) = g(m)\cdot g(n)$

Proof. Exercise. $\square$

The injection from $\mathbb{N}$ to $\mathbb{D}_+$ in Lemma 4.8 preserves the ordering in the sense of $<$ (and that in the sense of $\leq$ too) as well as the binary functions addition and multiplication. This justifies the usage of the same symbols $<$, $\leq$, $+$, and $\cdot$. Furthermore this allows the deliberate usage of mixed notations such as $\lfloor m,u\rfloor + n$ for $\lfloor m,u\rfloor + g(n)$, $n < \lfloor m,u\rfloor$ for $g(n) < \lfloor m,u\rfloor$, etc. where $m,n,u \in \mathbb{N}$. In each case such mixed notation is understood as shorthand notation for the full expression including the required injections. Moreover we may write 0 for $g(0)$ and 1 for $g(1)$. Similarly, if $A \subset \mathbb{N}$ and $B \subset \mathbb{D}_+$, we may write $A \cap B$ instead of $g\,[A] \cap B$ without ambiguity. Occasionally, given $d \in \mathbb{D}_+$, we may even write $d \in \mathbb{N}$ instead of $d \in g\,[\mathbb{N}]$, or, given $A \subset \mathbb{D}_+$, we may write $A \subset \mathbb{N}$ instead of $A \subset g\,[\mathbb{N}]$.

Proposition 4.9
Let $d, e, f \in \mathbb{D}_+$ with $d < f$ and $e \neq 0$. There is $g \in \mathbb{D}_+$ such that $d < e\,g < f$.

Proof. Let $\lfloor m, u\rfloor, \lfloor n, v\rfloor, \lfloor p, w\rfloor \in \mathbb{D}_+$ where $\lfloor n, v\rfloor < \lfloor p, w\rfloor$ and $m > 0$. We may choose $r \in \mathbb{N}$ such that $m < 2^r$. Then we have $\lfloor m, r\rfloor < 1$. Further we may choose $s \in \mathbb{N}$ such that $v + w < u + s$. We have $\lfloor m, u\rfloor \cdot \lfloor q,\ r + s\rfloor = \lfloor m, r\rfloor \cdot \lfloor q,\ u + s\rfloor$ for every $q \in \mathbb{N}$. We define $n_0 = n\,2^w$ and $p_0 = p\,2^v$. There exists $q \in \mathbb{N}$ such that $\lfloor n, v\rfloor = \lfloor n_0,\ v + w\rfloor < \lfloor m, r\rfloor \cdot \lfloor q,\ u + s\rfloor < \lfloor p_0,\ v + w\rfloor = \lfloor p, w\rfloor$.

[Let $q_0 \in \mathbb{N}$ be the maximum natural number such that $\lfloor m, r\rfloor \cdot \lfloor q_0,\ u+s\rfloor \leq \lfloor n_0,\ v + w\rfloor$. Then

$$\begin{aligned}
\lfloor m, r\rfloor \cdot \lfloor q_0 + 1,\ u + s\rfloor &= \lfloor m, r\rfloor \cdot \big(\lfloor q_0,\ u + s\rfloor + \lfloor 1,\ u + s\rfloor\big) \\
&< \lfloor n_0,\ v + w\rfloor + \lfloor 1,\ v + w\rfloor \\
&\leq \lfloor p_0,\ v + w\rfloor
\end{aligned}$$

]

□

Proposition 4.10
Let $d, e, f, g \in \mathbb{D}_+$. If $d < e$ and $f < g$, then we have $d\,g + e\,f < d\,f + e\,g$.

Proof. If the stated condition holds, then there are $a, b \in \mathbb{D}_+$ such that $d + a = e$

and $f + b = g$ by Proposition 4.6. Then we have

$$\begin{aligned} d\,g + e\,f &= d\,(f + b) + (a + d)\,f \\ &< (d + a)\,(f + b) + d\,f \\ &= e\,g + d\,f \end{aligned}$$

□

Proposition 4.11
Let $a, b, c \in \mathbb{D}_+$ with $a < b + c$. Then there are $f, g \in \mathbb{D}_+$ such that

$$f < b\,, \quad g < c\,, \quad a < f + g$$

Proof. We may choose $d \in \mathbb{D}_+$, $d > 0$, such that $a + 2d < b + c$ by Proposition 4.6. If $b < d$, then we may choose $f \in \mathbb{D}_+$ with $f < b$ by Lemma 4.4. If $b \geq d$, then there is h with $h + d = b$, and we may choose $f \in \mathbb{D}_+$ such that $h < f < b$ by Lemma 4.4. In both cases we have $b < f + d$. In a similar way we may choose a number $g \in \mathbb{D}_+$ such that $c < g + d$. We then have $a + 2d < f + g + 2d$, and hence $a < f + g$. □

Proposition 4.12
Let $a, b, c \in \mathbb{D}_+$ with $a < b\,c$. Then there are $f, g \in \mathbb{D}_+$ such that

$$f < b\,, \quad g < c\,, \quad a < f\,g$$

Proof. We may choose $d, e \in \mathbb{D}_+$ such that $a < d < e < b\,c$ by Lemma 4.4. There is $m \in \mathbb{N}$ such that $d + \lfloor 1, m \rfloor < e$ by Proposition 4.6. Further there is $k \in \mathbb{N}$ such that $(b + c)\lfloor 1, k \rfloor < \lfloor 1, 2m \rfloor$ by Proposition 4.9. We define n to be the maximum of $\{m, k\}$. We may choose $f, g \in \mathbb{D}_+$ such that

$$f < b\,, \quad b < f + \lfloor 1, n \rfloor\,, \quad g < c\,, \quad c < g + \lfloor 1, n \rfloor$$

It follows that $(f+g)\lfloor 1,n\rfloor < (b+c)\lfloor 1,k\rfloor < \lfloor 1,2m\rfloor$. Thus we have

$$\begin{aligned} b\,c \;&<\; f\,g + (f+g)\lfloor 1,n\rfloor + \lfloor 1,2n\rfloor \\ &<\; f\,g + \lfloor 1,2m\rfloor + \lfloor 1,2m\rfloor \;=\; f\,g + \lfloor 1,m\rfloor \end{aligned}$$

□

4.2 Positive real numbers

In this Section we introduce the positive real numbers, its orderings, as well as addition, multiplication, and exponentiation on positive real numbers. We also show how to identify the positive dyadic rationals as a subset of the positive reals.

Lemma and Definition 4.13
We define $\mathbb{D}_0 = \{\,]-\infty, d[\; : \, d \in \mathbb{D}_+\}$ where the lower segments refer to the relation $<$ on $\mathbb{D}_+$, and $\mathbb{R}_+ = \left\{\bigcup\mathcal{A} \, : \, \mathcal{A} \subset \mathbb{D}_0, \, \mathcal{A} \neq \emptyset\right\} \setminus \{\mathbb{D}_+\}$. The members of $\mathbb{R}_+$ are called **positive real numbers**. We further define a total ordering in the sense of "$\leq$" on $\mathbb{R}_+$ by

$$\alpha \leq \beta \quad\Longleftrightarrow\quad \alpha \subset \beta$$

Moreover, we define $<$ to be the total ordering in the sense of "$<$" on $\mathbb{R}_+$ obtained from the ordering $\leq$ by the method of Lemma 2.19.

Proof. It follows from Lemma 2.23 that $\leq$ is an ordering in the sense of "$\leq$". To see that $\leq$ is connective let $\alpha, \beta \in \mathbb{R}_+$ and assume that $\alpha \leq \beta$ does not hold. Then there is $d \in \mathbb{D}_+$ such that $d \in \alpha$, $d \notin \beta$. It follows that $e < d$ for every $e \in \beta$, and thus $\beta \subset \alpha$. □

Remark 4.14
Notice that $\lfloor 0,0 \rfloor \in \mathbb{D}_+$, and $\emptyset = \ ]{-\infty}, \lfloor 0,0 \rfloor[\ \in \mathbb{D}_0 \subset \mathbb{R}_+$. ■

Remember that the usage of the symbol $-\infty$ in an interval denotes a lower segment but does generally not imply that the interval has no lower bound. On the contrary, the members of $\mathbb{D}_0$ all have a lower bound, viz. $\emptyset$.

Lemma 4.15
$\mathbb{D}_0$ is $<$-dense in $\mathbb{R}_+$. $\mathbb{R}_+$ is $<$-dense.

Proof. To see the first claim, let $\alpha, \beta \in \mathbb{R}_+$ with $\alpha < \beta$. Then there exists $d \in \beta \setminus \alpha$. Furthermore, there exists $e \in \mathbb{D}_+$ such that $d \in \ ]{-\infty}, e[\ \subset \beta$. Thus we have $\alpha \leq \ ]{-\infty}, d[\ < \ ]{-\infty}, e[\ \leq \beta$. The claim follows by Lemma 4.4.
The second claim is a consequence of the first one. □

Lemma 4.16
The ordered space $(\mathbb{R}_+, <)$ has the least upper bound property. Specifically, if $A \subset \mathbb{R}_+$, $A \neq \emptyset$, and A has an upper bound, then $\sup A = \bigcup A$.

Proof. Assume the stated conditions and let $\alpha = \bigcup A$. We may choose $\beta \in \mathbb{R}_+$ such that $\gamma < \beta$ for every $\gamma \in A \setminus \{\beta\}$. It follows that $\alpha \subset \beta$, and thus $\alpha \in \mathbb{R}_+$. Moreover, we have $\gamma \leq \alpha$ for every $\gamma \in A$. It is also clear that α is the least upper bound. □

Lemma and Definition 4.17

We define two binary functions $+$ (called **addition**) and $\cdot$ (called **multiplication**) on $\mathbb{R}_+$ by

$$\begin{aligned}\alpha+\beta &= \bigcup\{\,]-\infty, d+e[\ : d\in D,\, e\in E\}\\ \alpha\cdot\beta &= \bigcup\{\,]-\infty, d\cdot e[\ : d\in D,\, e\in E\}\end{aligned}$$

where $D, E \subset \mathbb{D}_+$ such that $D, E \neq \emptyset$, $\alpha = \bigcup\{\,]-\infty, d[\ : d\in D\}$ and $\beta = \bigcup\{\,]-\infty, e[\ : e\in E\}$. $\alpha+\beta$ and $\alpha\cdot\beta$ are called the **sum** and the **product of** α **and** β, respectively. We also write $\alpha\,\beta$ for $\alpha\cdot\beta$. Both functions are commutative and associative, and the distributive law

$$(\alpha+\beta)\cdot\gamma = (\alpha\cdot\gamma)+(\beta\cdot\gamma)$$

holds for $\alpha,\beta,\gamma\in\mathbb{R}_+$.

We define that in the absence of brackets products are evaluated before sums.

Proof. We first show that $+$ and $\cdot$ are well-defined.

We may choose upper bounds $d_0, e_0 \in \mathbb{D}_+$ of D and E, respectively. Then $d+e \le d_0+e_0$ and $d\,e \le d_0\,e_0$ for every $d\in D$, $e\in E$. Thus $(\alpha+\beta), (\alpha\cdot\beta) \neq \mathbb{D}_+$. In order to see that the definitions of $+$ and $\cdot$ do not depend on the choice of the index sets D and E, let $\alpha,\beta\in\mathbb{R}_+$ and, for $i\in\{1,2\}$, let $D_i, E_i \subset \mathbb{D}_+$ such that $D_i, E_i \neq \emptyset$, $\alpha = \bigcup\{\,]-\infty, d[\ : d\in D_i\}$ and $\beta = \bigcup\{\,]-\infty, e[\ : e\in E_i\}$. Further let $d_1\in D_1$ and $e_1\in E_1$, and assume that not both $d_1 = 0$ and $d_2 = 0$.

To show the claim for $+$, let $f\in\mathbb{D}_+$ with $f < d_1+e_1$. First we consider the case $d_1 > 0$, $e_1 > 0$. We may choose $a, b\in\mathbb{D}_+$ such that $a < d_1$, $b < e_1$, and $f < a+b$ by Proposition 4.11. Since $a\in\alpha$, there is $d_2\in D_2$ such that $a < d_2$. Similarly, since $b\in\beta$, there is $e_2\in E_2$ such that $b < e_2$. It follows that $f < d_2+e_2$. Second consider the case $d_1 = 0$, $e_1 > 0$. We may choose any $d_2\in D_2$, and $b\in\mathbb{D}_+$ such that $f < b < e_1$ by Proposition 4.4, and $e_2\in E_2$ such that $b < e_2$. It follows

that $f < d_2 + e_2$. The case $d_1 > 0$, $e_1 = 0$ is handled similarly.
To show the claim for $\cdot$, assume that $d_1 > 0$ and $e_1 > 0$, and let $f \in \mathbb{D}_+$ with $f < d_1\, e_1$. We may choose $a, b \in \mathbb{D}_+$ such that $a < d_1$, $b < e_1$, and $f < a\, b$ by Proposition 4.12. There is $d_2 \in D_2$ such that $a < d_2$. Further there is $e_2 \in E_2$ such that $b < e_2$. It follows that $f < d_2\, e_2$.
The commutativity and associativity of $+$ and $\cdot$ on $\mathbb{R}_+$ is a consequence of the respective properties of $+$ and $\cdot$ on $\mathbb{D}_+$.
In order to prove the distributive law we define the following sets:

$$\begin{aligned} \alpha &= \bigcup \big\{ \,]-\infty, d[\; : \, d \in D \big\}\,, & R &= \{ d + e \, : \, d \in D,\, e \in E \}\,, \\ \beta &= \bigcup \big\{ \,]-\infty, e[\; : \, e \in E \big\}\,, & S &= \{ d\, f \, : \, d \in D,\, f \in F \}\,, \\ \gamma &= \bigcup \big\{ \,]-\infty, f[\; : \, f \in F \big\}\,, & T &= \{ e\, f \, : \, e \in E,\, f \in F \} \end{aligned}$$

The distributive law then follows from the following calculation:

$$\begin{aligned} (\alpha + \beta)\, \gamma &= \bigcup \big\{ \,]-\infty, d + e[\; : \, d \in D,\, e \in E \big\} \cdot \bigcup \big\{ \,]-\infty, f[\; : \, f \in F \big\} \\ &= \bigcup \big\{ \,]-\infty, r[\; : \, r \in R \big\} \cdot \bigcup \big\{ \,]-\infty, f[\; : \, f \in F \big\} \\ &= \bigcup \big\{ \,]-\infty, r\, f[\; : \, r \in R,\, f \in F \big\} \\ &= \bigcup \big\{ \,]-\infty, (d + e)\, f[\; : \, d \in D,\, e \in E,\, f \in F \big\} \\ &= \bigcup \big\{ \,]-\infty, d\, f + e\, f[\; : \, d \in D,\, e \in E,\, f \in F \big\} \\ &= \bigcup \big\{ \,]-\infty, d\, f + e\, h[\; : \, d \in D,\, e \in E,\, f \in F,\, h \in F \big\} \\ &= \bigcup \big\{ \,]-\infty, s + t[\; : \, s \in S,\, t \in T \big\} \\ &= \bigcup \big\{ \,]-\infty, s[\; : \, s \in S \big\} + \bigcup \big\{ \,]-\infty, t[\; : \, t \in T \big\} \\ &= \alpha\, \gamma + \beta\, \gamma \end{aligned}$$

$\square$

Lemma 4.18

Let $g : \mathbb{D}_+ \longrightarrow \mathbb{R}_+$, $g(d) = \;]{-\infty}, d[\;$. Then g is injective. For every $d, e \in \mathbb{D}_+$ we have

(i) $g(\lfloor 0, 0 \rfloor) = \emptyset$

(ii) $d < e \iff g(d) < g(e)$

(iii) $g(d + e) = g(d) + g(e)$

(iv) $g(d \cdot e) = g(d) \cdot g(e)$

Proof. Exercise. □

Regarding the injection from the positive dyadic rationals to the positive reals in Lemma and Definition 4.18, the same comments apply as regarding the injection from the natural numbers to the positive dyadic rationals defined in Lemma 4.8. That is, it preserves the orderings $<$ and $\leq$ as well as the binary functions $+$ and $\cdot$. This, again, justifies the usage of the same symbols and allows one to write mixed expressions of positive dyadic rationals and positive reals, but also of natural numbers and positive reals. In the latter case the notation of both injections is then suppressed. For instance, $\alpha + d$, $m \cdot \alpha$, and $\alpha \leq m$ are valid expressions, where $m \in \mathbb{N}$, $d \in \mathbb{D}_+$, and $\alpha \in \mathbb{R}_+$. Occasionally, we may even write $\alpha \in \mathbb{N}$, or $A \subset \mathbb{N}$ although actually $A \subset \mathbb{R}_+$.

Proposition 4.19

For every $\delta, \varepsilon \in \mathbb{R}_+$ and $r \in \mathbb{D}_+$, the inequality $\delta \leq r$ implies $\delta + \varepsilon \leq r + \varepsilon$ and $\delta\,\varepsilon \leq r\,\varepsilon$.

Proof. Let $G, H \subset \mathbb{D}_+$, $G, H \neq \emptyset$, such that

$$\delta = \bigcup \{\,]{-\infty}, g[\;:\; g \in G \}\,, \qquad \varepsilon = \bigcup \{\,]{-\infty}, h[\;:\; h \in H \}$$

If $\delta \leq r$, then $g \leq r$ for every $g \in G$. It follows that

$$\begin{aligned} \delta + \varepsilon &= \bigcup \big\{ \left]-\infty, g+h\right[\; : \; g \in G,\, h \in H \big\} \\ &\leq \bigcup \big\{ \left]-\infty, r+h\right[\; : \; h \in H \big\} \;=\; r + \varepsilon \end{aligned}$$

and

$$\begin{aligned} \delta\,\varepsilon &= \bigcup \big\{ \left]-\infty, g\,h\right[\; : \; g \in G,\, h \in H \big\} \\ &\leq \bigcup \big\{ \left]-\infty, r\,h\right[\; : \; h \in H \big\} \;=\; r\,\varepsilon \end{aligned}$$

□

Lemma 4.20
Given $\alpha, \beta, \gamma \in \mathbb{R}_+$, the following implications hold:

$$\begin{aligned} \alpha < \beta \quad &\Longrightarrow \quad \alpha + \gamma < \beta + \gamma \\ (\alpha < \beta) \wedge (0 < \gamma) \quad &\Longrightarrow \quad \alpha\,\gamma < \beta\,\gamma \end{aligned}$$

Proof. Let $D, E, F \subset \mathbb{D}_+$ such that $D, E, F \neq \emptyset$ and

$$\alpha = \bigcup \big\{ \left]-\infty, d\right[\; : \; d \in D \big\}, \qquad \beta = \bigcup \big\{ \left]-\infty, e\right[\; : \; e \in E \big\},$$
$$\gamma = \bigcup \big\{ \left]-\infty, f\right[\; : \; f \in F \big\}$$

Assume that $\alpha < \beta$. We may choose $a, b \in \mathbb{D}_+$ such that $\alpha < a < b < \beta$ by Proposition 4.15. Then we have $d < a$ for every $d \in D$. Further there exists $e_0 \in E$ such that $b < e_0$.
Notice that the first implication is clear for $\gamma = 0$. Now assume that $0 < \gamma$. We may choose $m \in \mathbb{N}$ such that $a + \lfloor 1, m \rfloor < b$ by Proposition 4.6. Let k be the maximum of $\{n \in \mathbb{N} \, : \, \lfloor n, m \rfloor < \gamma\}$.

[There exists $\lfloor p,v\rfloor \in \mathbb{D}_+$ such that $\gamma < \lfloor p,v\rfloor$. We have

$$\lfloor p,v\rfloor = \lfloor p\,2^m, v+m\rfloor \le \lfloor p\,2^m, m\rfloor$$

Therefore the considered set is finite and non-empty since $0 < \gamma$. Thus it has a maximum by Lemma 3.60.]

There is $c \in F$ such that $\lfloor k,m\rfloor < c$. We have

$$\begin{aligned}
\alpha+\gamma \;\le\; a+\gamma \;\le\; a+\lfloor\sigma(k),m\rfloor \;=\; a+\lfloor k,m\rfloor+\lfloor 1,m\rfloor \;<\; b+\lfloor k,m\rfloor \\
<\; b+c \;\le\; \bigcup\big\{\,]-\infty, b+f[\;:\; f\in F\big\} \\
\le\; \bigcup\big\{\,]-\infty, e+f[\;:\; e\in E,\, f\in F\big\} \;=\; \beta+\gamma
\end{aligned}$$

where the first and second inequality follow by Proposition 4.19.
To prove the second implication, let $h \in \mathbb{D}_+$ such that $a+h=b$ by Proposition 4.6, and assume that $0<\gamma$. We have $0<h$, and thus $0<h\,\gamma$. It follows that

$$\begin{aligned}
\alpha\,\gamma \;\le\; a\,\gamma \;<\; a\,\gamma+h\,\gamma \;=\; b\,\gamma \;=\; \bigcup\big\{\,]-\infty, b\,f[\;:\; f\in F\big\} \\
\le\; \bigcup\big\{\,]-\infty, e\,f[\;:\; e\in E,\, f\in F\big\} \;=\; \beta\,\gamma
\end{aligned}$$

The first inequality follows by Proposition 4.19. The second inequality is a consequence of the inequality for sums. □

Lemma 4.21
Let $\alpha,\beta \in \mathbb{R}_+$. We have $\alpha+\beta = \{a+b \;:\; a\in\alpha,\; b\in\beta\}$.

Proof. Let $D,E \subset \mathbb{D}_+$ with $D,E \neq \emptyset$ such that

$$\alpha \;=\; \bigcup\big\{\,]-\infty, d[\;:\; d\in D\big\}, \qquad \beta \;=\; \bigcup\big\{\,]-\infty, e[\;:\; e\in E\big\}$$

Now let $c \in \alpha + \beta$. There are $d \in D$ and $e \in E$ such that $c < d + e$. If $c < d$, then we have $c \in \alpha$. If $c < e$, then we have $c \in \beta$. If $c \geq d$ and $c \geq e$, then we define $g \in \mathbb{D}_+$ such that $c + g = d + e$ by Proposition 4.6. It follows that $0 < g$, $g \leq d \leq 2d$, and $g \leq e \leq 2e$. We further define $a \in \mathbb{D}_+$ such that $g + 2a = 2d$, as well as $b \in \mathbb{D}_+$ such that $g + 2b = 2e$. Thus we obtain $a \in \alpha$, $b \in \beta$, and $a + b = c$. The converse is clear. □

Proposition 4.22
Let $\alpha, \beta \in \mathbb{R}_+$. If $\alpha < \beta$, then there is $\gamma \in \mathbb{R}_+$ such that $\alpha + \gamma = \beta$.

Proof. Assume the condition. We define $\gamma = \sup\{\delta \in \mathbb{R}_+ \,:\, \alpha + \delta < \beta\}$.

[The supremum is well-defined by Lemmas 4.20 and 4.16 since $\alpha + \beta \geq \beta$.]

We may choose $A \subset \mathbb{D}_+$ with $A \neq \emptyset$ and, for every $\delta \in \mathbb{R}_+$, $D_\delta \subset \mathbb{D}_+$ with $D_\delta \neq \emptyset$ such that

$$\alpha = \bigcup\{\,]-\infty, a[\,:\, a \in A\}, \qquad \delta = \bigcup\{\,]-\infty, d[\,:\, d \in D_\delta\}$$

It follows that

$$\gamma = \bigcup\{\,]-\infty, d[\,:\, d \in D_\delta,\ \alpha + \delta < \beta\} = \bigcup\{\,]-\infty, d[\,:\, d \in D\}$$

where

$$D = \bigcup\{D_\delta \,:\, \alpha + \delta < \beta\}$$

Thus we obtain

$$\begin{aligned}\alpha + \gamma &= \bigcup\{\,]-\infty, a + d[\,:\, a \in A,\ d \in D\} \\ &= \bigcup\{\,]-\infty, a + d[\,:\, a \in A,\ d \in D_\delta,\ \alpha + \delta < \beta\} \leq \beta\end{aligned}$$

Now assume that $\alpha+\gamma < \beta$. Then there are $d, e \in \mathbb{D}_+$ such that $\alpha+\gamma < d < e < \beta$ by Lemma 4.15. Further there is $f \in \mathbb{D}_+$ with $f > 0$ such that $d + f = e$ by Proposition 4.6. It follows that $\alpha + \gamma + f < \beta$, which is a contradiction. □

Proposition 4.23
Let $\alpha, \beta \in \mathbb{R}_+ \backslash \{0\}$. There is $\gamma \in \mathbb{R}_+$ such that $0 < \alpha\,\gamma < \beta$.

Proof. We may choose $a, b \in \mathbb{D}_+$ such that $\alpha < a$ and $0 < b < \beta$ by Lemma 4.15. There is $c \in \mathbb{D}_+$ such that $0 < a\,c < b$ by Proposition 4.9. It follows that $0 < \alpha\,c < a\,c < \beta$. □

Proposition 4.24
Let $\alpha, \beta, \gamma, \delta \in \mathbb{R}_+$. If $\alpha < \beta$ and $\gamma < \delta$, then $\alpha\,\delta + \beta\,\gamma < \alpha\,\gamma + \beta\,\delta$.

Proof. This follows by the corresponding result for positive dyadic rationals, see Lemma 4.10.

[We may choose $A, B, C, D \subset \mathbb{D}_+$ such that

$$\alpha = \bigcup \{ \,]-\infty, a[\ : a \in A \}, \qquad \beta = \bigcup \{ \,]-\infty, b[\ : b \in B \},$$
$$\gamma = \bigcup \{ \,]-\infty, c[\ : c \in C \}, \qquad \delta = \bigcup \{ \,]-\infty, d[\ : d \in D \}$$

By definition of addition and multiplication we have

$$\alpha\,\delta + \beta\,\gamma = \bigcup \{ \,]-\infty, a\,d + b\,c[\ : a \in A,\ b \in B,\ c \in C,\ d \in D \},$$
$$\alpha\,\gamma + \beta\,\delta = \bigcup \{ \,]-\infty, a\,c + b\,d[\ : a \in A,\ b \in B,\ c \in C,\ d \in D \}$$

Now let $e \in \alpha\,\delta + \beta\,\gamma$. We may choose $a \in A$, $b_0 \in B$, $c \in C$, and $d_0 \in D$ such that $e < a\,d_0 + b_0\,c$. Further, there is $b \in B$ such that $a < b$ and $b_0 \leq b$. Similarly, there is $d \in D$ such that $c < d$ and $d_0 \leq d$. It follows that $e < a\,d + b\,c$. We have $a\,d + b\,c < a\,c + b\,d$ by Lemma 4.10. Hence $e \in \alpha\,\gamma + \beta\,\delta$.]

□

Proposition 4.25
Let $\alpha, \beta, \gamma \in \mathbb{R}_+$ with $\alpha\,\beta < \gamma$. There are $a, b \in \mathbb{D}_+$ such that $\alpha < a$, $\beta < b$, and $a\,b < \gamma$.

Proof. We may choose $c, d \in \mathbb{D}_+$ such that $\alpha\,\beta < c < d < \gamma$ by Lemma 4.15. There is $m \in \mathbb{N}$ with $0 < m$ such that $c + \lfloor 1, m \rfloor < d$ by Proposition 4.6. We may choose $k \in \mathbb{N}$ such that $(\alpha + \beta\ ) \cdot \lfloor 1, k \rfloor < \lfloor 1, m + 1 \rfloor$ by Proposition 4.9. Let $n = \sup \{m, k\}$. We may choose $a, b \in \mathbb{D}_+$ such that $\alpha < a < \alpha + \lfloor 1, n \rfloor$ and $\beta < b < \beta + \lfloor 1, n \rfloor$ by Lemma 4.15. Further, notice that $m + 1 \leq 2m$.

[The claim is clear for $m = 1$. Assume it holds for some $m \in \mathbb{N}$ with $m \geq 1$. Then we have $\sigma(m) + 1 \leq 2m + 1 < 2\,\sigma(m)$.]

It follows that

$$\begin{aligned} a\,b \;&<\; \alpha\,\beta + (\alpha + \beta\,) \cdot \lfloor 1, n \rfloor + \lfloor 1, 2n \rfloor \\ &<\; \alpha\,\beta + \lfloor 1, m+1 \rfloor + \lfloor 1, 2m \rfloor \\ &\le\; \alpha\,\beta + \lfloor 1, m+1 \rfloor + \lfloor 1, m+1 \rfloor \;=\; \alpha\,\beta + \lfloor 1, m \rfloor \end{aligned}$$

□

Theorem 4.26

The triple $(\mathbb{R}_+ \backslash \{0\}\,,\,\cdot\,, 1)$ is an Abelian group. Let $\alpha, \beta \in \mathbb{R}_+ \backslash \{0\}$. The inverse of α with respect to multiplication is denoted by α^{-1}, $(1/\alpha)$ or $\left(\dfrac{1}{\alpha}\right)$. We also write (α/β) or $\left(\dfrac{\alpha}{\beta}\right)$ for $\alpha \cdot (1/\beta)$. In the absence of brackets the superscript "-1" is evaluated before sums and products.

Proof. Multiplication is associative and commutative by Lemma and Definition 4.17.

Moreover, we have $\alpha \cdot 1 = \alpha$ for every $\alpha \in \mathbb{R}_+$ by definition.

Now let $\alpha \in \mathbb{R}_+ \backslash \{0\}$. We define $A = \{\beta \in \mathbb{R}_+ \backslash \{0\} \,:\, \beta\,\alpha < 1\}$. It follows that $A \neq \emptyset$ by Proposition 4.23. Moreover, A has an upper bound by Proposition 4.9. Hence the supremum of A is well-defined by Lemma 4.16, and we have $\sup A = \bigcup A$. We define $1/\alpha = \sup A$. We may choose $D \subset \mathbb{D}_+$ with $D \neq \emptyset$ and, for every $\beta \in \mathbb{R}_+ \backslash \{0\}$, $E_\beta \subset \mathbb{D}_+$ with $E_\beta \neq \emptyset$ such that

$$\alpha \;=\; \bigcup \big\{\,]{-\infty}, d[\;:\; d \in D \big\}, \qquad \beta \;=\; \bigcup \big\{\,]{-\infty}, e[\;:\; e \in E_\beta \big\}$$

It follows that

$$1/\alpha \;=\; \bigcup \big\{\,]{-\infty}, e[\;:\; e \in E_\beta,\ \beta \in \mathbb{R}_+ \backslash \{0\}\,,\ \beta\,\alpha < 1 \big\}$$

and hence

$$\alpha \cdot (1/\alpha) \;=\; \bigcup \big\{\,]{-\infty}, d\,e[\;:\; d \in D,\ e \in E_\beta,\ \beta \in \mathbb{R}_+ \backslash \{0\}\,,\ \beta\,\alpha < 1 \big\} \;\le\; 1$$

Now assume that $\alpha \cdot (1/\alpha) < 1$. Then there is $f \in \mathbb{D}_+$ such that $1/\alpha < f$ and $\alpha\, f < 1$ by Proposition 4.25, which is a contradiction.
Thus we have $\alpha \cdot (1/\alpha) = 1$. □

Corollary 4.27
Let $\alpha, \beta \in \mathbb{R}_+$ with $0 < \alpha < \beta$. Then we have $\beta^{-1} < \alpha^{-1}$.

Proof. The inequality $\alpha < \beta$ implies $1 < \beta\, \alpha^{-1}$. The claim follows. □

Remark 4.28
Let $\alpha, \beta \in \mathbb{R}_+ \backslash \{0\}$. We have $(\alpha\, \beta)^{-1} = \alpha^{-1}\, \beta^{-1}$. ■

Corollary 4.29
Let $\alpha, \beta \in \mathbb{R}_+ \backslash \{0\}$. There is $\gamma \in \mathbb{R}_+$ such that $\beta < \alpha\, \gamma$.

Proof. There is $\delta \in \mathbb{R}_+$ such that $\alpha^{-1}\, \delta < \beta^{-1}$ by Proposition 4.23. It follows that $\beta < \alpha\, \delta^{-1}$ by Corollary 4.27 and Remark 4.28. □

We continue by defining exponentiation on the positive reals where the exponent is a natural number.

Lemma and Definition 4.30
We define a function $h : \mathbb{R}_+ \times \mathbb{N} \longrightarrow \mathbb{R}_+$ recursively by

(i) $h(\alpha, 0) = 1$

(ii) $h(\alpha, \sigma(m)) = h(\alpha, m) \cdot \alpha$

for every $\alpha \in \mathbb{R}_+$ and every $m \in \mathbb{N}$. This function is called **exponentiation on** $\mathbb{R}_+$. We also write α^m for $h(\alpha, m)$ and call α the **base** and m the **exponent** or **power**. In the absence of brackets we define the following priorities:

$$\alpha^{m+n} = \alpha^{(m+n)}\,, \quad \alpha^{m\cdot n} = \alpha^{(m\cdot n)}\,, \quad \alpha + \beta^m = \alpha + (\beta^m)\ , \quad \alpha\,\beta^m = \alpha\ (\beta^m)$$

We have, for every $\alpha, \beta \in \mathbb{R}_+$ and $m, n \in \mathbb{N}$,

$$\alpha^{m+n} = \alpha^m\,\alpha^n, \qquad (\alpha^m)^n = \alpha^{m\cdot n}, \qquad (\alpha\,\beta)^m = \alpha^m\,\beta^m$$

and the implications

$$\begin{aligned} (\alpha < \beta)\ \wedge\ (0 < m) \quad &\Longrightarrow \quad \alpha^m < \beta^m \\ (0 < \alpha < 1)\ \wedge\ (m < n) \quad &\Longrightarrow \quad \alpha^n < \alpha^m \\ (1 < \alpha)\ \wedge\ (m < n) \quad &\Longrightarrow \quad \alpha^m < \alpha^n \end{aligned}$$

Furthermore we define $\alpha^{-m} = \left(\alpha^{-1}\right)^m$ for $\alpha \in \mathbb{R}_+\backslash\{0\}$ and $m \in \mathbb{N}$.
Given $\alpha \in \mathbb{R}_+\backslash\{0\}$ and $m, n \in \mathbb{N}$, we have $(\alpha^m)^{-1} = \alpha^{-m}$. If $m \leq n$, then $\alpha^p = \alpha^n\,\alpha^{-m}$ where $p \in \mathbb{N}$ such that $m + p = n$. If $m > n$, then $\alpha^{-p} = \alpha^n\,\alpha^{-m}$ where $p \in \mathbb{N}$ such that $n + p = m$.

Proof. The existence and uniqueness of the function follow by Theorem 3.13.
The three equations follow by the Induction principle.
We now show the three implications, again by means of the Induction principle.
To see the first implication, assume that $\alpha < \beta$. The implication clearly holds for $m = 1$. Now assume that it holds for some $m \in \mathbb{N}$ with $0 < m$. Then we have

$\alpha^{\sigma(m)} = \alpha^m\,\alpha < \beta^m\,\beta = \beta^{\sigma(m)}$.
To see the second implication, let $m \in \mathbb{N}$ and $0 < \alpha < 1$. We have $\alpha^{\sigma(m)} = \alpha^m\,\alpha < \alpha^m$.

> [The second inequality follows by Lemma 4.20 since $0 < \alpha^m$, which in turn is proven by the Induction principle.]

Now assume that the implication holds for some $n \in \mathbb{N}$ with $n \geq \sigma(m)$. It follows that $\alpha^{\sigma(n)} = \alpha^n\,\alpha < \alpha^m\,\alpha < \alpha^m$.
To see the third implication, let $m \in \mathbb{N}$ and $1 < \alpha$. We have $\alpha^m < \alpha^m\,\alpha < \alpha^{\sigma(m)}$.
Now assume that the implication holds for some $n \in \mathbb{N}$ where $n \geq \sigma(m)$. It follows that $\alpha^m < \alpha^m\,\alpha < \alpha^n\,\alpha < \alpha^{\sigma(n)}$.
The equation $(\alpha^m)^{-1} = \alpha^{-m}$ clearly holds for every $\alpha \in \mathbb{R}_+ \backslash \{0\}$ and $m = 0$. Now assume that it holds for every $\alpha \in \mathbb{R}_+ \backslash \{0\}$ and some $m \in \mathbb{N}$. We have

$$\left(\alpha^{\sigma(m)}\right)^{-1} = (\alpha^m\,\alpha)^{-1} = (\alpha^m)^{-1}\,\alpha^{-1} = \left(\alpha^{-1}\right)^m\,\alpha^{-1} = \left(\alpha^{-1}\right)^{\sigma(m)}$$

Finally let $\alpha \in \mathbb{R}_+ \backslash \{0\}$ and $m, n \in \mathbb{N}$. If $m \leq n$, then we have

$$\alpha^n\,\alpha^{-m} = \alpha^{m+p}\,\alpha^{-m} = \alpha^m\,\alpha^p\,\alpha^{-m} = \alpha^p$$

If $m > n$, then we have

$$\alpha^p\,\alpha^n\,\alpha^{-m} = \alpha^{p+n}\,\alpha^{-m} = 1$$

Thus $\alpha^n\,\alpha^{-m}$ is the inverse of α^p. □

Lemma 4.31
For every $\lfloor m, u \rfloor \in \mathbb{D}_+$ and $n \in \mathbb{N}$, we have $\left(g\left(\lfloor m, u \rfloor\right)\right)^n = g\left(\lfloor m^n, n\,u \rfloor\right)$ where g is defined in Lemma and Definition 4.17.

Proof. This follows by the Induction principle. □

Corollary 4.32
Let $g : \mathbb{N} \longrightarrow \mathbb{R}_+$, $g(m) = \,]-\infty, \lfloor m, 0 \rfloor[$. Then g is injective, and we have for every $m, n \in \mathbb{N}$:

$$g(m^n) = g(m)^n$$

Proof. The map g is injective as it is the composition of two injections by Lemma 4.8 and Lemma and Definition 4.17.
The equation follows by Lemma 4.31. □

Corollary 4.32 shows that exponentiation on $\mathbb{N}$ as defined in Lemma and Definition 3.25 and exponentiation on $\mathbb{R}_+$ as defined in Lemma and Definition 4.30 are in agreement with the injection from $\mathbb{N}$ to $\mathbb{R}_+$. Therefore we may also use mixed notation even when exponentiation occurs.

Corollary 4.33
Using the notation of negative exponents in Lemma and Definition 4.30, we have $\lfloor m, u \rfloor = m\, 2^{-u}$ for $m, u \in \mathbb{N}$.

Proof. We have $\lfloor 1, u \rfloor \cdot 2^u = \lfloor 2^u, u \rfloor = 1$. Therefore $\lfloor 1, u \rfloor$ is the inverse of 2^u, i.e. $\lfloor 1, u \rfloor = 2^{-u}$. □

Lemma 4.34
$\mathbb{R}_+ \backslash \mathbb{D}_0$ is $<$-dense in $\mathbb{R}_+$.

Proof. We first show that $1/3 \notin \mathbb{D}_+$. Assume there are $m, u \in \mathbb{N}$ such that $3 \cdot m\, 2^{-u} = 1$. It follows that $3m = 2^u$. This is clearly false for every $m \in \mathbb{N}$ and $u = 0$. Assume it is false for every $m \in \mathbb{N}$ and some $u \in \mathbb{N}$. If m is even, then there is $n \in \mathbb{N}$ such that $m = 2n$, and therefore $3m = 2^{u+1}$ implies $3n = 2^u$, which is a contradiction. If m is odd, then there is $n \in \mathbb{N}$ such that $m = 2n + 1$, and thus $3m = 2^{u+1}$ implies $3 \cdot 2n + 3 = 2^{u+1}$. The left hand side of the last equation is odd whereas the right hand side is even, which is again a contradiction.

Now let $\alpha, \beta \in \mathbb{R}_+$ with $0 < \alpha < \beta$. There are $a, b \in \mathbb{D}_+$ such that $\alpha < a < b < \beta$ by Lemma 4.15. We may choose $m, u \in \mathbb{N}$ such that $a = \lfloor m, u \rfloor$ and $\lfloor m+1, u \rfloor < b$. We define $\gamma = (m + 1/3)\ 2^{-u}$. Assume that $\gamma = n\, 2^{-v}$ for some $n, v \in \mathbb{N}$. It follows that $m + 1/3 = n\, 2^u\, 2^{-v}$, which is a contradiction to the first part of the proof. □

4.3 Real numbers

In this Section it remains to construct the full number systems, i.e. those containing positive and negative numbers. Since natural numbers and positive dyadic rationals can be identified with a subset of the positive reals as shown above, it is enough to construct the system of positive and negative real numbers, its orderings, as well as addition and multiplication on the reals.

Lemma and Definition 4.35
Let P be the equivalence relation on $\mathbb{R}_+^2$ defined by

$$\big((\alpha,\beta),(\gamma,\delta)\big) \in P \quad\Longleftrightarrow\quad \alpha+\delta=\gamma+\beta$$

and $\mathbb{R} = \mathbb{R}_+/P$. The equivalence classes are called **real numbers**. For every $\alpha,\beta \in \mathbb{R}_+$, the equivalence class of (α,β) is denoted by $\langle\alpha,\beta\rangle$. We define a total ordering in the sense of "$<$" on $\mathbb{R}$ by

$$\langle\alpha,\beta\rangle < \langle\gamma,\delta\rangle \quad\Longleftrightarrow\quad \alpha+\delta<\gamma+\beta$$

It is called the **standard ordering in the sense of** "$<$". Moreover, we define $\leq$ to be the total ordering in the sense of "$\leq$" on $\mathbb{R}$ obtained from the ordering $<$ by the method of Lemma 2.19. It is called the **standard ordering in the sense of** "$\leq$".

Proof. Exercise. $\square$

Definition 4.36
We adopt the convention that all notions related to orderings on $\mathbb{R}$, in particular intervals, refer to the standard ordering in the sense of "$<$" as defined in Lemma and Definition 4.35 unless otherwise specified.
Furthermore, for every $x, y \in \mathbb{R}$ with $x < y$ we define

$$\begin{aligned}]-\infty, y] &= \ ]-\infty, y[\ \cup \{y\}, & [x, \infty[&= \ ]x, \infty[\ \cup \{x\}, \\ ]x, y] &= \ ]x, y[\ \cup \{y\}, & [x, y[&= \ ]x, y[\ \cup \{x\}, \\ [x, y] &= \ ]x, y[\ \cup \{x, y\} \end{aligned}$$

The set $[x, y]$ with $x, y \in \mathbb{R}$ is called **closed interval**. Notice that it is a proper interval with respect to the ordering $\leq$.
We further agree that the sets $[0, \infty[$ and $]0, \infty[$ always refer to subsets of $\mathbb{R}$, or equivalently to the sets $\mathbb{R}_+$ and $\mathbb{R}_+ \backslash \{0\}$ unless otherwise specified. ■

This convention is in agreement with the ones adopted in the context of natural numbers, Definition 3.5, and positive dyadic rational numbers, Definition 4.2. We remark again that, apart from the definition of intervals, it is mostly irrelevant whether the ordering $<$ or the ordering $\leq$ on $\mathbb{R}$ is considered, cf. Lemmas 2.34, 2.46, and 2.78.

Lemma and Definition 4.37
We define two binary functions $+$ (called **addition**) and $\cdot$ (called **multiplication**) on $\mathbb{R}$ by

$$\begin{aligned} \langle \alpha, \beta \rangle + \langle \gamma, \delta \rangle &= \langle \alpha + \gamma, \ \beta + \delta \rangle \\ \langle \alpha, \beta \rangle \cdot \langle \gamma, \delta \rangle &= \langle \alpha\, \gamma + \beta\, \delta, \ \alpha\, \delta + \beta\, \gamma \rangle \end{aligned}$$

$\langle \alpha, \beta \rangle + \langle \gamma, \delta \rangle$ and $\langle \alpha, \beta \rangle \cdot \langle \gamma, \delta \rangle$ are called the **sum** and the **product of** $\langle \alpha, \beta \rangle$

and $\langle \gamma, \delta \rangle$, respectively. For every $x, y \in \mathbb{R}$ we also write $x\,y$ for $x \cdot y$. We further define that in the absence of brackets products are evaluated before sums. Both addition and multiplication are commutative and associative, and the distributive law

$$(x + y) \cdot z \;=\; (x \cdot z) + (y \cdot z)$$

as well as the implications

$$\begin{aligned} x < y \quad &\Longrightarrow \quad x + z < y + z \\ x < y \;\;\wedge\;\; \langle 0, 0 \rangle < z \quad &\Longrightarrow \quad x\,z < y\,z \\ x < y \;\;\wedge\;\; z < \langle 0, 0 \rangle \quad &\Longrightarrow \quad y\,z < x\,z \end{aligned}$$

hold for $x, y, z \in \mathbb{R}$. For every $x \in \mathbb{R}$ with $x > \langle 0, 0 \rangle$ there is $\alpha \in \mathbb{R}_+$ such that $x = \langle \alpha, 0 \rangle$. Furthermore, for every $x \in \mathbb{R}$ with $x < \langle 0, 0 \rangle$ there is $\alpha \in \mathbb{R}_+$ such that $x = \langle 0, \alpha \rangle$.
Further, let $g : \mathbb{R}_+ \longrightarrow \mathbb{R}$, $g(\alpha) = \langle \alpha, 0 \rangle$, and $h : \mathbb{R}_+ \longrightarrow \mathbb{R}$, $h(\beta) = \langle 0, \beta \rangle$. Then we have

$$g\,[\mathbb{R}_+] = \{ x \in \mathbb{R} \,:\, x \geq \langle 0, 0 \rangle \} \;, \qquad h\,[\mathbb{R}_+] = \{ x \in \mathbb{R} \,:\, x \leq \langle 0, 0 \rangle \}$$

The functions g and h are injective. We have

(i) $\alpha < \beta \iff g(\alpha) < g(\beta) \iff h(\alpha) > h(\beta)$

(ii) $g(\alpha + \beta) = g(\alpha) + g(\beta)$

(iii) $h(\alpha + \beta) = h(\alpha) + h(\beta)$

(iv) $g(\alpha\,\beta) = g(\alpha)\,g(\beta)$

Furthermore, we define the **exponentiation** on the positive subset by

$$f : g\,[\mathbb{R}_+] \times \mathbb{N} \longrightarrow g\,[\mathbb{R}_+] \;, \quad f(\langle \alpha, 0 \rangle, m) = \langle \alpha^m, 0 \rangle$$

For every $m \in \mathbb{N}$ and $x \in \mathbb{R}$ with $x \geq \langle 0, 0\rangle$, we also write x^m for $f(x, m)$. We define the same rules regarding the order of evaluation as for $\mathbb{R}_+$. We have

$$g\left(\alpha^m\right) = (g(\alpha))^m$$

for every $m \in \mathbb{N}$ and $\alpha \in \mathbb{R}_+$.

Proof. The proofs that addition and multiplication are well-defined, and that they are commutative and associative, as well as the proof of the distributive law are left as exercise.

To see the three implications, let $\langle \alpha, \beta\rangle$, $\langle \gamma, \delta\rangle$, $\langle \chi, \psi\rangle \in \mathbb{R}$ with $\langle \alpha, \beta\rangle < \langle \gamma, \delta\rangle$. Hence we have $\alpha + \delta < \beta + \gamma$.

To show the first implication, notice that $\alpha + \delta + \chi + \psi < \beta + \gamma + \chi + \psi$. It follows that $\langle \alpha, \beta\rangle + \langle \chi, \psi\rangle = \langle \alpha + \chi,\ \beta + \psi\rangle < \langle \gamma + \chi,\ \delta + \psi\rangle = \langle \gamma, \delta\rangle + \langle \chi, \psi\rangle$.

To show the second implication, assume that $\langle \chi, \psi\rangle > \langle 0, 0\rangle$. This implies $\psi < \chi$. By Proposition 4.24 we obtain

$$(\alpha + \delta)\,\chi + (\beta + \gamma)\,\psi \;<\; (\alpha + \delta)\,\psi + (\beta + \gamma)\,\chi$$

and thus

$$\alpha\,\chi + \beta\,\psi + \gamma\,\psi + \delta\,\chi \;<\; \alpha\,\psi + \beta\,\chi + \gamma\,\chi + \delta\,\psi$$

It follows that $\langle \alpha, \beta\rangle \cdot \langle \chi, \psi\rangle = \langle \alpha\,\chi + \beta\,\psi,\ \alpha\,\psi + \beta\,\chi\rangle < \langle \gamma\,\chi + \delta\,\psi,\ \gamma\,\psi + \delta\,\chi\rangle = \langle \gamma, \delta\rangle \cdot \langle \chi, \psi\rangle$.

To see the third implication, assume that $\langle \chi, \psi\rangle < \langle 0, 0\rangle$. This implies $\psi > \chi$. In this case we obtain

$$\alpha\,\chi + \beta\,\psi + \gamma\,\psi + \delta\,\chi \;>\; \alpha\,\psi + \beta\,\chi + \gamma\,\chi + \delta\,\psi$$

by Proposition 4.24. This implies $\langle \alpha, \beta\rangle \cdot \langle \chi, \psi\rangle > \langle \gamma, \delta\rangle \cdot \langle \chi, \psi\rangle$.

Now let $\alpha, \beta \in \mathbb{R}_+$. If $\langle \alpha, \beta\rangle > \langle 0, 0\rangle$, then we have $\beta < \alpha$. Thus there exists $\gamma \in \mathbb{R}_+$ such that $\beta + \gamma = \alpha$ by Proposition 4.22. Therefore we have $\langle \alpha, \beta\rangle = \langle \gamma, 0\rangle$.

If $\langle \alpha, \beta \rangle < \langle 0, 0 \rangle$, then there is $\gamma \in \mathbb{R}_+$ such that $\alpha + \gamma = \beta$ by Proposition 4.22. Thus $\langle \alpha, \beta \rangle = \langle 0, \gamma \rangle$.
Finally the results (i) to (iv) clearly follow by definition. $\square$

Here as in the previous cases the fact that the injection from the positive reals to the reals preserves the orderings and binary functions explains the usage of the same symbols and allows us to deliberately mix the different kinds of numbers in expressions, such as $x + m$, $x \cdot \lfloor m, u \rfloor$, or $x \cdot \alpha$, where $m, u \in \mathbb{N}$, $\alpha \in \mathbb{R}_+$, and $x \in \mathbb{R}$.
The next Proposition, on which the subsequent Lemma is based, is almost obvious though its derivation is a bit lengthy.

Proposition 4.38
Let g and h be defined as in Lemma and Definition 4.37, $B_+ \subset \mathbb{R}_+$ with $B_+ \neq \emptyset$, and $\mathbb{R}_- = \{x \in \mathbb{R} : x \leq 0\}$. The following statements hold:

(i) If B_+ has a minimum (maximum), say α, then $g(\alpha)$ is a minimum (maximum) of $g[B_+]$ and $h(\alpha)$ is a maximum (minimum) of $h[B_+]$.

(ii) Let L_+ and U_+ be the sets of all lower and upper bounds of B_+, respectively, U_g the set of all upper bounds of $g[B_+]$, and U_h the set of all upper bounds of $h[B_+]$, i.e.

$$\begin{aligned} L_+ &= \big\{\alpha \in \mathbb{R}_+ : \forall \beta \in B_+ \backslash \{\alpha\} \quad \alpha < \beta\big\} \\ U_+ &= \big\{\alpha \in \mathbb{R}_+ : \forall \beta \in B_+ \backslash \{\alpha\} \quad \beta < \alpha\big\} \\ U_g &= \big\{x \in \mathbb{R} : \forall y \in g[B_+] \backslash \{x\} \quad y < x\big\} \\ U_h &= \big\{x \in \mathbb{R} : \forall y \in h[B_+] \backslash \{x\} \quad y < x\big\} \end{aligned}$$

Then we have $g[U_+] = U_g$ and $h[L_+] = U_h \cap \mathbb{R}_-$.

(iii) If B_+ has a supremum, then $g[B_+]$ has a supremum and we have $\sup g[B_+] = g(\sup B_+)$.

(iv) B_+ has an infimum, $h[B_+]$ has a supremum, and we have $\sup h[B_+] = h(\inf B_+)$.

Proof. (i) and (ii) are consequences of Lemma and Definition 4.37 (i).

[If α is the minimum of B_+, then we have $\alpha \in B_+$ and $\alpha < \beta$ for every $\beta \in B_+ \backslash \{\alpha\}$. It follows that $g(\alpha) \in g[B_+]$ and $h(\alpha) \in h[B_+]$. Moreover, we have $g(\alpha) < g(\beta)$ and $h(\beta) < h(\alpha)$ for every $\beta \in B_+ \backslash \{\alpha\}$. Therefore we have $g(\alpha) < x$ for every $x \in g[B_+ \backslash \{\alpha\}] = g[B_+] \backslash \{g(\alpha)\}$. Additionally, we have $x < h(\alpha)$ for every $x \in h[B_+ \backslash \{\alpha\}] = h[B_+] \backslash \{h(\alpha)\}$. The claim in (i) that is stated in brackets is shown similarly.

Further, we obtain (ii) as follows:

$$\begin{aligned}
g[U_+] &= g\big[\big\{\alpha \in \mathbb{R}_+ : \forall \beta \in B_+ \backslash \{\alpha\} \ \ \beta < \alpha\big\}\big] \\
&= \big\{g(\alpha) : \alpha \in \mathbb{R}_+ , \forall \beta \in B_+ \backslash \{\alpha\} \ \ \beta < \alpha\big\} \\
&= \big\{g(\alpha) : \alpha \in \mathbb{R}_+ , \forall \beta \in B_+ \backslash \{\alpha\} \ \ g(\beta) < g(\alpha)\big\} \\
&= \big\{g(\alpha) : \alpha \in \mathbb{R}_+ , \forall y \in g[B_+ \backslash \{\alpha\}] \ \ y < g(\alpha)\big\} \\
&= \big\{g(\alpha) : \alpha \in \mathbb{R}_+ , \forall y \in g[B_+] \backslash \{g(\alpha)\} \ \ y < g(\alpha)\big\} \\
&= \big\{x \in \mathbb{R} : x \geq 0, \forall y \in g[B_+] \backslash \{x\} \ \ y < x\big\} \\
&= \big\{x \in \mathbb{R} : \forall y \in g[B_+] \backslash \{x\} \ \ y < x\big\} = U_g
\end{aligned}$$

$$\begin{aligned}
h[L_+] &= h\big[\big\{\alpha \in \mathbb{R}_+ : \forall \beta \in B_+ \backslash \{\alpha\} \ \ \alpha < \beta\big\}\big] \\
&= \big\{h(\alpha) : \alpha \in \mathbb{R}_+ , \forall \beta \in B_+ \backslash \{\alpha\} \ \ \alpha < \beta\big\} \\
&= \big\{h(\alpha) : \alpha \in \mathbb{R}_+ , \forall \beta \in B_+ \backslash \{\alpha\} \ \ h(\beta) < h(\alpha)\big\} \\
&= \big\{h(\alpha) : \alpha \in \mathbb{R}_+ , \forall y \in h[B_+ \backslash \{\alpha\}] \ \ y < h(\alpha)\big\} \\
&= \big\{h(\alpha) : \alpha \in \mathbb{R}_+ , \forall y \in h[B_+] \backslash \{h(\alpha)\} \ \ y < h(\alpha)\big\} \\
&= \big\{x \in \mathbb{R} : \forall y \in h[B_+] \backslash \{x\} \ \ y < x\big\} \cap \mathbb{R}_- = U_h \cap \mathbb{R}_-
\end{aligned}$$

]

Now (iii) and (iv) follow by (i) and (ii).

[If the condition of (iii) is satisfied, then $\sup B_+$ is the minimum of U_+ as defined in (ii). Hence, by (i), $g\,(\sup B_+)$ is the minimum of $g\,[U_+]$, which is the supremum of $g\,[B_+]$ by (ii).
The infimum of B_+ exists by Lemma 4.16 and Theorem 2.49. Moreover, $\inf B_+$ is the maximum of L_+ as defined in (ii). Hence, by (i), $h\,(\inf B_+)$ is the minimum of $h\,[L_+]$. This in turn is the minimum of U_h, which is the supremum of $h\,[B_+]$ by (ii).]

□

Lemma 4.39
The ordered space $(\mathbb{R}, <)$ has the least upper bound property.

Proof. Let $A \subset \mathbb{R}$ such that A has an upper bound and $A \neq \emptyset$. Further let U be the set of all upper bounds of A.
First assume there exists $y \in A$ with $y \geq 0$. We define $B = \{y \in A : y \geq 0\}$ and $B_+ = g^{-1}\,[B]$ where g is defined as in Lemma and Definition 4.37. Then $B_+ \neq \emptyset$ and $g\,[B_+] = B$. Moreover, U is the set of all upper bounds of B. Thus also B_+ has an upper bound by Proposition 4.38 (ii). Therefore $\sup B_+$ exists by Lemma 4.16, and $g(\sup B_+) = \sup B = \sup A$ by Proposition 4.38 (iii).
Now assume that $y < 0$ for every $y \in A$. We define $A_+ = h^{-1}\,[A]$ where h is defined as in Lemma and Definition 4.37. Then we have $A_+ \neq \emptyset$. By Proposition 4.38 (iii), the infimum of A_+ and the supremum of $h\,[A_+]$ exist, and we have $h(\inf A_+) = \sup h\,[A_+] = \sup A$. □

Lemma 4.40
Given $x \in \mathbb{R}$, the function $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(y) = y + x$, is strictly increasing. If $x > 0$, then the function $g : \mathbb{R} \longrightarrow \mathbb{R}$, $g(y) = y\,x$, is strictly increasing. If $x < 0$, g is strictly decreasing. For every $m \in \mathbb{N}$, $m \geq 1$, the function $h_m : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$, $h_m(y) = y^m$, is strictly increasing. For every $x \in \mathbb{R}$ and $m \in \mathbb{N}$, $m \geq 1$, the functions f, g, and h_m are unbounded.

Proof. The fact that f is strictly increasing, and the fact that g is strictly increasing or decreasing under the respective conditions follows by Lemma and Definition 4.37. The fact that h_m is strictly increasing for every $m \in \mathbb{N}$, $m \geq 1$, follows by Lemma and Definition 4.30.

To see that the functions are unbounded, let $\gamma, \delta, \psi, \chi \in \mathbb{R}_+$ such that $x = \langle \gamma, \delta \rangle$. For the case of f, we define $\alpha = \psi + \delta + 1$ and $\beta = \chi + \gamma$. Then we have $\langle \alpha, \beta \rangle > \langle \psi, \chi \rangle + \langle \delta, \gamma \rangle$. It follows that $\langle \alpha, \beta \rangle + \langle \gamma, \delta \rangle > \langle \psi, \chi \rangle$.

For the case of g, we assume that $\psi > \chi$. If $\gamma > \delta$, then there is $\varepsilon \in \mathbb{R}_+$ such that $\gamma = \delta + \varepsilon$ by Proposition 4.22. We may choose $\zeta \in \mathbb{R}_+$ such that $\varepsilon \zeta > \psi$ by Corollary 4.29. We define $\alpha = \gamma + \zeta$ and $\beta = \gamma$. It follows that

$$\begin{aligned}
& \alpha \gamma + \beta \delta + \chi \\
= \;& \gamma \gamma + \zeta \gamma + \gamma \delta + \chi \\
= \;& \gamma \gamma + \zeta \delta + \zeta \varepsilon + \gamma \delta + \chi \\
> \;& \gamma \gamma + \zeta \delta + \psi + \gamma \delta + \chi \\
\geq \;& \gamma \gamma + \zeta \delta + \gamma \delta + \psi \\
= \;& \alpha \delta + \beta \gamma + \psi
\end{aligned}$$

Hence $\langle \alpha, \beta \rangle \cdot \langle \gamma, \delta \rangle = \langle \alpha \gamma + \beta \delta,\ \alpha \delta + \beta \gamma \rangle > \langle \psi, \chi \rangle$. If $\gamma < \delta$, then there is $\varepsilon \in \mathbb{R}_+$ such that $\delta = \gamma + \varepsilon$. Then we may again choose $\zeta \in \mathbb{R}_+$ such that $\varepsilon \zeta > \psi$. We define $\alpha = \gamma$ and $\beta = \gamma + \zeta$. It follows that

$$\begin{aligned}
& \alpha \gamma + \beta \delta + \chi \\
= \;& \gamma \gamma + \gamma \delta + \zeta \delta + \chi \\
= \;& \gamma \gamma + \gamma \delta + \zeta \gamma + \zeta \varepsilon + \chi \\
> \;& \gamma \gamma + \gamma \delta + \zeta \gamma + \psi \\
= \;& \alpha \delta + \beta \gamma + \psi
\end{aligned}$$

also in this case.

Let $m \in \mathbb{N}$ with $m \geq 1$. To see that h_m is unbounded, let $\beta \in \mathbb{R}_+$. We may choose $p \in \mathbb{N}$ such that $p > 1$ and $p > \beta$. It follows that $p^m \geq p > \beta$ by Lemma and Definition 3.25. □

Lemma and Definition 4.41

The triple $(\mathbb{R}, +, 0)$ is an Abelian group. Let $x, y \in \mathbb{R}$ and $\alpha, \beta \in \mathbb{R}_+$ such that $x = \langle \alpha, \beta \rangle$. The inverse of x with respect to addition is given by $\langle \beta, \alpha \rangle$ and denoted by $-x$. We also write $y - x$ for $y + (-x)$. In the absence of brackets we define the following priorities:

$$-x + y = (-x) + y, \quad -x\,y = -(x\,y), \quad -x^m = -\left(x^m\right)$$

Proof. The addition is associative and commutative by Lemma and Definition 4.37. The other assertions are clear. □

Remark 4.42

Let $\alpha, \beta \in \mathbb{R}_+$ and $x, y \in \mathbb{R}$. We have

(i) $\langle \alpha, \beta \rangle = \alpha - \beta$

(ii) If $\alpha > \beta$, then $(\alpha - \beta) \in \mathbb{R}_+$.

(iii) $0 - x = -x$

(iv) $(-1) \cdot x = -x$

(v) $x < y \implies -y < -x$

(vi) $(x > 0) \wedge (y > 0) \implies x\,y > 0$

(vii) $(x > 0) \wedge (y < 0) \implies x\,y < 0$

■

Lemma and Definition 4.43
The triple $(\mathbb{R}\backslash\{0\}, \cdot, 1)$ is an Abelian group. Let $x, y \in \mathbb{R}\backslash\{0\}$. The inverse of x with respect to multiplication is denoted by x^{-1}, $(1/x)$, or $\left(\frac{1}{x}\right)$. We also write (x/y) or $\left(\frac{x}{y}\right)$ for $x \cdot (1/y)$. In the absence of brackets the superscript "-1" is evaluated before sums and products, and we define $-x^{-1} = -\left(x^{-1}\right)$. We have $(-x)^{-1} = -\left(x^{-1}\right)$.

Proof. The multiplication is associative and commutative by Lemma and Definition 4.37. For every $\alpha, \beta \in \mathbb{R}_+$ we have

$$\langle \alpha, \beta \rangle \cdot \langle 1, 0 \rangle = \langle \alpha, \beta \rangle$$

For $x \in \mathbb{R}\backslash\{0\}$ we define

$$1/x = \begin{cases} \left\langle \dfrac{1}{\alpha - \beta}, 0 \right\rangle & \text{if} \quad \alpha > \beta \\ \left\langle 0, \dfrac{1}{\beta - \alpha} \right\rangle & \text{if} \quad \beta > \alpha \end{cases}$$

where $\alpha, \beta \in \mathbb{R}_+$ such that $x = \langle \alpha, \beta \rangle$. The inverses on the right hand side are defined according to Theorem 4.26. Notice that this definition is independent of the specific choice of α and β. Then $x \cdot (1/x) = 1$, that is $(1/x)$ is the inverse of x. The last claim is clear. □

Definition 4.44
The members of the set

$$\mathbb{D} = \left\{ x \in \mathbb{R} \,:\, x \in \mathbb{D}_+ \vee -x \in \mathbb{D}_+ \right\}$$

are called **dyadic rational numbers**. ■

Remark 4.45
$\mathbb{D}$ is countable by Corollary 4.3 and Lemma 3.70. ■

Lemma 4.46
We have

(i) $\mathbb{D}$ is $<$-dense in $\mathbb{R}$.

(ii) $\mathbb{R} \backslash \mathbb{D}$ is $<$-dense in $\mathbb{R}$.

(iii) $\mathbb{R}$ is $<$-dense.

Proof. (i) follows by Lemma 4.15.
(ii) follows by Lemma 4.34.
(iii) is a consequence of (i). □

Definition 4.47
We define the function $b : \mathbb{R} \longrightarrow \mathbb{R}_+$ by

$$b(x) = \begin{cases} x & \text{if} \quad x \geq 0 \\ -x & \text{if} \quad x < 0 \end{cases}$$

We also write $|x|$ for $b(x)$. $|x|$ is called **absolute value of** x. ■

Remark 4.48
The function b is clearly surjective. Moreover, we have

$$|x|^2 = x^2, \qquad |x + y| \leq |x| + |y|\,, \qquad |x\,y| = |x|\,|y|$$

for every $x, y \in \mathbb{R}$. ■

The following result is applied in the proof of Lemma 6.62.

Lemma 4.49

Let X be a set, $f : X \longrightarrow \mathbb{R}$ a map, and $c \in \mathbb{R}$. Then

$$\inf_{x \in X} \big(f(x) + c \big) \;=\; \inf_{x \in X} f(x) \,+\, c$$

Proof. Let a be a lower bound of $\{f(x) + c \,:\, x \in X\}$, i.e. we have $a \leq f(x) + c$ for every $x \in X$. It follows that $a - c \leq f(x)$ for every $x \in X$, i.e. $(a - c)$ is a lower bound of $\{f(x) \,:\, x \in X\}$. Thus we have $a - c \leq \inf_{x \in X} f(x)$, and hence

$$\inf_{x \in X} \big(f(x) + c \big) \;\leq\; \inf_{x \in X} f(x) \,+\, c$$

Applying this result to $-c$ instead of c, and to the function

$$g : X \longrightarrow \mathbb{R}, \quad g(x) = f(x) + c$$

we obtain the reverse inequality. □

We conclude this Section with some examples of orderings and functions involving the real numbers.

Example 4.50

We recall Example 2.82: Let (X_i, R_i) $(i \in I)$ be pre-ordered spaces, where I is an index set, and $X = \bigtimes_{i \in I} X_i$. Then $\mathcal{R} = \left\{ p_i^{-1} [R_i] \,:\, i \in I \right\}$ is a system of pre-orderings on X.

Now, if $(X_i, R_i) = (\mathbb{R}, <)$ $(i \in I)$, then the members of $\mathcal{R}$ are orderings in the sense of "$<$". However, they are not total orderings unless I is a singleton. Clearly $\mathcal{R}$ is independent. ■

Example 4.51

Let (X_i, R_i) $(i \in I)$ be pre-ordered spaces, where I is an index set, and $X = \bigtimes_{i \in I} X_i$. Then $S = \bigcap \left\{ p_i^{-1}[R_i] : i \in I \right\}$ is a pre-ordering on X (cf. Example 2.84). Now let $n \in \mathbb{N}$, $n > 0$, and $I = \sigma(n) \backslash \{0\}$. If $(X_k, R_k) = (\mathbb{R}, <)$ $(k \in \mathbb{N},\ 1 \leq k \leq n)$, then S is an ordering in the sense of "$<$". However, it is not a total ordering unless $n = 1$. For $x, y \in \mathbb{R}^n$ we have $x < y$ iff $x_k < y_k$ $(1 \leq k \leq n)$. For the same I, if $(X_k, R_k) = (\mathbb{R}, \leq)$ $(1 \leq k \leq n)$, then S is an ordering in the sense of "$\leq$" since the ordering $\leq$ on $\mathbb{R}$ is antisymmetric and $\{p_i : i \in I\}$ distinguishes points. However the ordering $\leq$ on $\mathbb{R}^n$ is not a total ordering unless $n = 1$. For $x, y \in \mathbb{R}^n$ we have $x \leq y$ iff $x_k \leq y_k$ $(1 \leq k \leq n)$. Both $<$ and $\leq$ on $\mathbb{R}^n$ have full range and full domain. They are not connective unless $n = 1$. Since $\mathbb{D}$ is dense in $\mathbb{R}$, $\mathbb{D}^n$ is dense in $\mathbb{R}^n$ with respect to both orderings. ■

Example 4.52

The pair $(\mathbb{R}, \leq)$, where $\leq$ denotes the standard ordering in the sense of "$\leq$", is a pre-ordered space. The function $f : \mathbb{R} \longrightarrow \mathbb{R}$ that maps every real number x to the smallest integer greater or equal than x is $\leq$-increasing and projective. Similarly, the function $g : \mathbb{R} \longrightarrow \mathbb{R}$ that maps every real number x to the smallest *even* integer greater or equal than x is $\leq$-increasing and projective. Thus also the composition $g \circ f$ is $\leq$-increasing. ■

Bibliography

[Bernays] P. Bernays. *Axiomatic Set Theory.* North-Holland, Amsterdam, 1958.

[Ebbinghaus] H.-D. Ebbinghaus. *Einführung in die Mengenlehre.* BI-Wissenschaftsverlag, Mannheim, third edition, 1994.

[Gaal] S. A. Gaal. *Point Set Topology.* Dover, Mineola, 2009.

[Jech] T. Jech. *Set Theory.* Springer, Berlin, third edition, 2002.

[Kelley] J. L. Kelley. *General Topology.* Sam Sloan and Ishi Press International, New York, 2008.

[Nagel] F. Nagel. *Set Theory and Topology. An Introduction to the Foundations of Analysis.* Parts II to V, in preparation.

[von Querenburg] B. v. Querenburg. *Mengentheoretische Topologie.* Springer, Berlin, third edition, 2001.

[Steen] L. A. Steen, J. A. Seebach, Jr. *Counterexamples in Topology.* Dover, New York, 1995.

[Suppes] P. Suppes. *Axiomatic Set Theory.* D. van Nostrand Company, Princeton, 1960.

[Wilansky] A. Wilansky. *Topology for Analysis.* Dover, Mineola, 2008.

Index