Set Theory

Frank Stephan

Semester I, Academic Year 2009-2010

Set Theory deals with the fundamental concepts of sets and functions used everywhere in mathematics. Cantor initiated the study of set theory with his investigations on the cardinality of sets of real numbers. In particular, he proved that there are different infinite cardinalities: the quantity of natural numbers is strictly smaller than the quantity of real numbers. Cantor formalized and studied the notions of ordinal and cardinal numbers. Set theory considers a universe of sets which is ordered by the membership or element relation $\in$. All other mathematical objects are coded into this universe and studied within this framework. In this way, set theory is one of the foundations of mathematics.

This text contains all information relevant for the exams. Furthermore, the exercises in this text are those which will be demonstrated in the tutorials. Each sheet of exercises contains some important ones marked with a star and some other ones. You have to hand in an exercise marked with a star in Weeks 3 to 6, Weeks 7 to 9 and Weeks 10 to 12; each of them gives one mark. Furthermore, you can hand in any further exercises, but they are only checked for correctness. There will be two mid term exams and a final exam; the mid term exams count 15 marks each and the final exam counts 67 marks.

Frank Stephan: Room S17#07-04
Departments of Mathematics and Computer Science, National University of Singapore
Singapore 117543, Republic of Singapore
Telephone 65162759
Email: fstephan@comp.nus.edu.sg
Homepage http://www.comp.nus.edu.sg/~fstephan/index.html

Thanks. These course notes were prepared by Frank Stephan. F. Stephan would like to thank Feng Qi and Klaus Gloede for their material used in these notes. Further thanks go to Johanna Franklin, Eric Martin, Yan Kuo, Yang Yue and several of his students for discussions, suggestions and improvements.

Contents

1 Foundations **3**

2 Basic Operations with Sets **7**

3 Functions **13**

4 Natural Numbers **18**

5 Recursive Definition **23**

6 Cardinality of Sets **28**

7 Finite and Hereditarily Finite Sets **30**

8 Countable Sets **35**

9 Graphs and Orderings **39**

10 Linear Ordering **43**

11 Well-Orderings **50**

12 Ordinals **55**

13 Transfinite Induction and Recursion **58**

14 The Rank of Sets **63**

15 Arithmetic on Ordinals **65**

16 Cardinals **69**

17 The Axiom of Choice **73**

18 The Set of Real Numbers **77**

19 The Continuum Hypothesis **82**

20 The Axioms of Zermelo and Fraenkel **86**

21 References **91**

1 Foundations

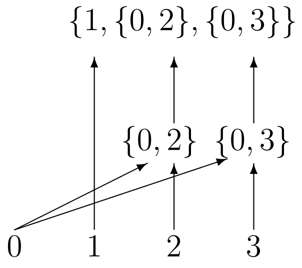
The two primitive notions of set theory are "*set*" and "*membership*". These two notions will not be defined. All other concepts are defined in terms of these two primitive notions.

In this course, all basic properties (Axioms) of set-theory are introduced and the various sets are defined one after the other. Nevertheless, for some examples and exercises, reference is made to well-known but later introduced objects in order to illustrate certain notions or results.

A set is intuitively a collection of objects. The objects in the collection are members (or elements) of the set. One important point is that the objects in a set can again be sets. Here an example.

Example 1.1. $\{1, \{0, 2\}, \{0, 3\}\}$ *is the set having three elements, namely* 1*,* $\{0, 2\}$ *and* $\{0, 3\}$*. The second and third elements are again sets containing* $0, 2$ *and* $0, 3$*, respectively.*

Graph Representation 1.2. One can consider every set to be a vertex of a large graph where there is a connection from x to y iff $x \in y$. The graph of the above example has the following 7 vertices and 7 directed edges:



The domain, that is the collection of all sets, is denoted by V and called the Von Neumann Universe.

Writing down a set is longer than writing down any element of it. Thus one can only finitely often go down from a set to a member since each time the description becomes shorter. Later, when infinite sets are introduced, it will no longer be possible to write them down in the above way and therefore the nonexistence of infinitely descending chains is no longer implicitly guaranteed. Therefore, this property is kept by making it explicitly an axiom.

The intuition is that there is no sequence $x_0, x_1, \ldots$ of vertices in V such that

$x_{n+1} \in x_n$ for $n = 0, 1, \ldots$, that is, there is no infinite descending chain. But the term "sequence" is not a primitive operation. Thus the following version of the axiom is given; it is based on the observation that whenever $x_{n+1} \in x_n$ then every element x_n of $X = \{x_0, x_1, \ldots\}$ has a common element with X, namely x_{n+1} is in x_n and in X. This is just ruled out by the Axiom of Foundation.

Axiom 1.3 (Foundation). Let $X \in V$ be a set which contains at least one element. Then there is an element $y \in X$ such that every $z \in X$ satisfies $z \notin y$.

Example 1.4. *Such a $y \in X$ is called a minimal element of X (with respect to $\in$). The set $A = \{0, \{0\}, \{1, 2\}, \{0, \{1, 2\}\}, \{\{0\}\}\}$ has two minimal elements, namely 0 and $\{1, 2\}$. 0 has no elements and the elements $1, 2$ of $\{1, 2\}$ are not in A. Thus a minimal element does not need to be unique. If a minimal element is contained in every other element of the set, is called a least element. The set $\{0\}$ is the least element of the set $\{\{0\}, \{\{0\}\}, \{\{0\}, \{0, \{0\}\}\}\}$.*

Remark 1.5. The main idea of the Axiom of Foundation is to enforce that the universe V of sets is build from bottom up and that no set is "cyclic" or "hanging down without a bottom from something". So the following things are permitted and forbidden:

- $\{\{\{x\}\}\}$: making sets of sets is legal;
- $\{x, \{x\}, \{x, \{x\}\}\}$: sets with comparable elements ($x \in \{x\}$) are legal;
- $\{x, y\}$ with $x \notin y$ and $y \notin x$: sets with incomparable sets are legal;
- $\{x, y\}$ with $x \in y$ and $y \in x$: cycles are illegal;
- $\{x_0, x_1, \ldots\}$ with $x_1 \in x_0, x_2 \in x_1, \ldots$: sets forming a descending sequence are illegal (see above).

Note that the Axiom of Foundation implies that a descending sequence is illegal whenever its elements form a set; but it cannot make descending sequences explicitly illegal without that condition (as this cannot be expressed in V).

Exercise 1.6. The property of being well-founded is an abstract property which applies also to some but not all directed graphs which are different from the universe of all sets. Here some examples of graphs. Which of the below graphs are well-founded? The answers should be proven by testing which of the below examples avoid the two negative criteria (cycles and descending chains) from Remark 1.5.

1. the set $\{0, 1, \{0\}, \{1\}, \{0, 1, \{0\}\}, \{\{1\}\}, \{\{\{1\}\}\}, 512\}$ with (a, b) being an edge iff $a \in b$;
2. the set $\{0, 1, 2, 3\}$ with the edges $(0, 1), (1, 0), (2, 3)$;
3. the set $\mathbb{N}$ of the natural numbers with every edge being of the form $(n, n + 1)$;
4. the set $\mathbb{Z}$ of the integers with the edges being the pairs $(n, n + 1)$ for all $n \in \mathbb{Z}$;
5. the set $\mathbb{Q}$ of rational numbers with the edges being the pairs $(q, 2q)$ for all $q \in \mathbb{Q}$;
6. the set $\mathbb{Q}$ of rational numbers with the edges being the pairs $(q, q + 1)$ for all $q \geq 0$ and $(q, q - 1)$ for all $q \leq 0$.

The Axiom of Foundation has an immediate application.

Theorem 1.7. *The collection V of all sets is not a set.*

Proof. Consider any set x. Then $\{x\}$ is also a set; actually this needs the Axiom of Pairs introduced below. By the Axiom of Foundation, the only element x of $\{x\}$ satisfies $y \notin x$ for all $y \in \{x\}$. As y takes the value than x, $x \notin x$. Thus $x \neq V$ since V contains the set x as a member. It follows that V cannot be a set. ▮

Property 1.8. No set contains itself as an element.

Convention 1.9. In set theory, there are two types of collections of objects, which are called "sets" and "classes". The members of V are called sets and the subsets of V (including V itself) are called classes. A class is something what behaves like a set but it is none. So it has subclasses and members in the same way as a set has subsets and members. The main objective of V is to tell which thing is a set and which not; due to this role, V cannot be a set itself.

Sets are the object of investigation. They are represented as members of the class of vertices of V and they are put into relation to each other by the element-relation $\in$. Intuitively, any set could be considered as a collection of other smaller sets. The Axiom of Foundation guarantees, that each set is determined uniquely by its members and helps to avoid contradictions. All elements of sets will be coded again as sets, so the usage of $0, 1, 2, 3$ in the example above will later be replaced by the usage of sets which code the numbers $0, 1, 2, 3$.

Classes are an auxiliary structure which enable to make statement about the collection of all sets or the collection of sets with certain property. The elements of a class are always sets, no class contains another one as a member.

All the variables will stand for sets, unless otherwise stated. The notation "$\in$"

denotes the membership of sets. For example, when writing "$x \in y$", this means that "set x is a member of set y". Similarly "$x \notin y$" means that "set x is not a member of y". The words "member" and "element" are synonyms, the symbol "$\in$" refers to the first letter of "element".

It is supposed that $\{0, 1, 2, 3, 4, 5, 6, 7\}$ and $\{0, 1, \ldots, 7\}$ denote the same set since both descriptions give the collections of the same elements. Further descriptions for this set can even be more indirect, for example as the set of all numbers which can be represented by up to three binary digits or the set of all integers on which the polynomial $4x^2 - 28x - 1$ takes negative values. These descriptions might come from different intentions, but they give the same extension, that is, the same list of elements. Therefore, the axiom of extensionality states that two sets are equal iff they have the same elements, whether or not they are generated from different descriptions does not matter at all.

Axiom 1.10 (Extensionality). A set A and a set B are *the same set*, that is, *the two sets A and B are equal*, denoted by $A = B$, if both sets have the same elements, that is, for every $x \in V$, $x \in A$ if and only if $x \in B$.

The importance of this identification of sets is that one only pays attention to the extensionality of sets; the different intentions are ignored when describing sets. This is very useful when the uniqueness of sets satisfying a certain property is determined.

Example 1.11. Equal sets have exactly the same members. Britain is a country consisting of four members: Britain = {England, Northern Ireland, Scotland, Wales}. These four members go directly into the European association in the case of soccer: UEFA = {Austria, Belgium, England, Northern Ireland, Scotland, Wales, Denmark, Estonina, ...}. But politically, Britain is the member of the EU and they are only members of this member, that is, belong indirectly to the union: EU = {Austria, Belgium, {England, Northern Ireland, Scotland, Wales}, Denmark, Estonia, ...}. Direct and indirect membership is not the same, thus UEFA $\neq$ EU. It can be noted that there is also other reasons for EU $\neq$ UEFA like Switzerland $\in$ UEFA $-$ EU, but this is not the point of this example.

Example 1.12. *There is at most one set which does not have any element.*

Exercise 1.13. Which of the following sets of natural numbers are equal? Well-known mathematical theorems can be applied without proving them.

1. $A = \{1, 2\}$;
2. $B = \{1, 2, 3\}$;

3. C is the set of all prime numbers;
4. $D = \{d \mid \exists a, b, c > 0\, (a^d + b^d = c^d)\}$;
5. $E = \{e \mid e > 0 \wedge \forall c \in C\, (e \leq c)\}$;
6. $F = \{f \mid \forall c \in C\, (f \geq c)\}$;
7. $G = \{g \mid g \geq 2 \wedge \forall a, b > 1\, (4g \neq (a+b)^2 - (a-b)^2)\}$;
8. $H = \{h \mid h > 0 \wedge h^2 = h^h\}$;
9. $I = \{i \mid i + i = i \cdot i\}$;
10. $J = \{j \mid (j+1)^2 = j^2 + 2j + 1\}$;
11. $K = \{k \mid 4k > k^2\}$;
12. $L = \{l \mid \exists c \in C\, (l < c)\}$;
13. $M = \{m \mid \exists c \in C\, (m = c^2)\}$;
14. $N = \{n \mid \exists c, d \in C\, (n = cd)\}$;
15. $O = \{o \mid o \text{ has exactly three prime factors}\}$;
16. $P = \{p \mid p, p+2 \in C\}$.

The question whether P is infinite is a famous open problem. Therefore it is still unknown whether

$$\mathbb{N} = \{n \mid \exists p \in P\, (n \leq p)\}.$$

So it is sometimes very difficult to decide whether two descriptions give the same set or not.

2 Basic Operations with Sets

In the previous section, many examples dealt with sets known from the every-day-life of a mathematician, but there was neither a proof nor an axiom given such that these sets really exist. Since "existence of a set" means that it is a member of V, one has to place an axiom such that it really occurs in V or one has to prove using other axioms that it is in V by these axioms.

Axiom 2.1 (Empty Set). There exists a set in V which has no members.

Notice that by the equality of sets, as in the first example, a set with no elements is unique. This set not having any element is denoted by $\emptyset$.

Subsets. One set could be a part of another in the sense that every element of the set is an element of the other. This is made more precise by introducing the concept of subsets.

Definition 2.2. A set A is a *subset of* B, denoted by $A \subseteq B$, iff for every set x, $x \in A$ implies that $x \in B$. And A is a *proper subset of* B, denoted by $A \subset B$, iff A is a subset of B and there exists one set y which is an element of B but not an element of A.

Thus, $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$. Notice that this gives a standard way to show that two given sets are equal. Namely, one checks that one set is a subset of the other, respectively. Notice also that $x \subseteq x$ for any set x.

Property 2.3. *$\emptyset$ is a subset of every set.*

Power Sets. Given a set X, one could collect all the subsets of X to form a new set. This procedure is called the power set operation.

Axiom 2.4 (Power Set). For every set x, there exists a (unique) set, called the *power set of* x, whose elements are exactly subsets of x. This set is denoted by $\mathcal{P}(x)$.

By Property 2.3, $\emptyset \in \mathcal{P}(x)$ for every set $x \in V$. Also $x \in \mathcal{P}(x)$. If $x = \{a, b, c, d\}$ then $\mathcal{P}(x) = \{\emptyset,\ \{a\}, \{b\}, \{c\}, \{d\},\ \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\},\ \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\},\ \{a, b, c, d\}\}$.

Exercise 2.5. Determine the power set of $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$. Is there any set X such that $\mathcal{P}(X)$ has exactly 9 elements?

In Exercise 1.13, many sets are defined by taking all those numbers which satisfy a certain property. This rule of forming subsets is made an axiom and is formally defined. It uses properties which are formally introduced in Definition 3.7 below.

Axiom 2.6 (Comprehension, defining subsets). Given a property $p(y)$ of sets, for any set A, there exists a (unique) set B such that $x \in B$ if and only if $x \in A$ and $p(x)$ holds.

Convention 2.7. The notation $\{x \in A \mid p(x)\}$ stands for the set of all $x \in A$ which satisfy $p(x)$.

Example 2.8. Let C be the set of all countries. Now define the set

$$L = \{c \in C \mid c \text{ has at least 25 subunits}\}$$

of all large countries. The European Union has since 01.05.2004 the necessary quantity of members. The United States has 50 states, Switzerland has 26 cantons and India has 25 states and 7 territories. But Australia has only 6 states and 2 main territories. Thus,

$$\begin{aligned} \text{European Union, India, Switzerland, United States} &\in L, \\ \text{The Commonwealth of Australia} &\notin L. \end{aligned}$$

Similarly, the set

$$E = \{e \in \text{European Union } \mid e \text{ has the Euro}\}$$

consists of all members of the European Union, which use the Euro as a currency. So France and Spain are in E, Britain is not be in E. Montenegro uses the Euro but is not in the European Union, therefore, Montenegro is not a member of E.

Exercise 2.9. Given the set $\mathbb{N}$ of natural numbers, establish properties to define the following subsets of $\mathbb{N}$ using the Axiom of Comprehension:

1. the set of all numbers with exactly three divisors,
2. the set $\{0, 2, 4, 6, \ldots\}$ of all even numbers,
3. the set of all square numbers,
4. the set of all numbers whose binary representation contains exactly four times a 1.

For example, the set of prime numbers can be defined as the set $\{x \in \mathbb{N} \mid \exists$ unique $y, z \in \mathbb{N}\,(y \cdot z = x \wedge z < y \leq x)\}$, that is the set of all natural numbers with exactly two divisors.

Exercise 2.10. Show that every property p satisfies the following statements.

1. There are sets x, y such that $x \in y$ and either $p(x) \wedge p(y)$ or $\neg p(x) \wedge \neg p(y)$.
2. There is a set x with $x = \{y \in x \mid p(y)\}$.
3. There is a one-to-one function f such that $p(x)$ iff $p(y)$ for all $y \in f(x)$.

Before introducing the distinction between sets and classes, mathematicians believed that they can define things like "the set of all sets". Using comprehension, Russell split the set of all sets into the following two subsets: X is the set of all sets which are an element of itself, Y is the set of all sets which are not an element of itself.

Paradox 2.11 (Russell's Anatonomy of Naive Set Theory). If the above defined X, Y are sets, then there is a contradiction.

Proof. If $Y \in Y$, then $Y \in X$ since X contains all sets which are an element of itself. If $Y \notin Y$ then $Y \notin X$ again by the definition of X. Thus either Y is member of both sets X and Y or Y is not member of any of these two sets. This contradicts the fact that X, Y are obviously a partition of the sets of all sets. ∎

Comment 2.12. Russell's paradox forced the mathematicians to become a bit more cautious when dealing with sets. Mainly, the following two consequences where taken:

1. The distinction between classes and sets are introduced.

2. The postulate that V is well-founded.

So X and Y turn out to be classes and not sets. Furthermore, the postulate that V is well-founded makes it impossible for a set to be a member of itself. Therefore, Y is equal to the class V of all sets and X is the empty class.

Although a set cannot contain itself, there is for every set x also the set $\{x\}$ which is just the set containing the single element x as its member. Note that $x \neq \{x\}$ for the reasons given above, the existence of $\{x\}$ is provided by the next axiom.

In general, the intended property is that one can construct finite sets from a given finite list of elements. It is sufficient to postulate this for two given elements, note that taking the element x twice gives then the existence of $\{x\}$ by the Axiom of Extensionality.

Axiom 2.13 (Pair). Given any $x, y \in V$ then there exists also a set in V which contains exactly the elements x, y. This set is written as $\{x, y\}$; in the case that $x = y$ one can also just write $\{x\}$.

Union, Intersection and Difference. In the following, the basic operations with sets are defined.

Definition 2.14. Let A and B be sets.

1. The union of A and B is the set, denoted by $A \cup B$, whose elements are exactly those sets belonging to A or belonging to B.

2. The intersection of A and B is the set, denoted by $A \cap B$, whose elements are exactly those sets belonging to both A and B.

3. The (relative) difference of A with B is the set, denoted by $A - B$, whose elements are exactly those elements of A which do not belong to B.

For example, $\{0,1,2\} \cup \{4,5\} = \{0,1,2,4,5\}$, $\{0,1,2\} \cap \{1,2,3,4,5,6,7,8\} = \{1,2\}$ and $\{0,1,2\} - \{1,2,3\} = \{0\}$.

Remark 2.15. These operations can be specified as follows:

$$\begin{aligned}
A \cup B &= \{\, x \mid x \in A \text{ or } x \in B \,\};\\
A \cap B &= \{\, x \mid x \in A \text{ and } x \in B \,\};\\
A - B &= \{\, x \mid x \in A \text{ and } x \notin B \,\};\\
A \triangle B &= (A - B) \cup (B - A);\\
\bigcup A &= \{\, x \mid \exists\, y \in A\ (x \in y) \,\};\\
\bigcap A &= \{\, x \mid \forall\, y \in A\ (x \in y) \,\}.
\end{aligned}$$

Note that $\bigcap A$ is only defined for nonempty sets A. If $B \subseteq A$, then $A - B$ is called the complement of B in A. The set $A \triangle B$ is called the symmetric difference of A and B.

As an illustration, let $A = \{a,b,c\}$, $B = \{b,c,d\}$, $C = \{c,d,e\}$ and $D = \{\{a,b\}, \{b,c,d\}, \{b,d,e\}\}$. Then

$$\begin{aligned}
A \cup B &= \{a,b,c,d\};\\
A \cup C &= \{a,b,c,d,e\};\\
B \cup C &= \{b,c,d,e\};\\
A \cap B &= \{b,c\};\\
A \cap C &= \{c\};\\
B \cap C &= \{c,d\};\\
A - B &= \{a\};\\
A \triangle B &= \{a,d\};\\
A \triangle B \triangle C &= \{a,c,e\};\\
\bigcup \{A,B,C\} &= \{a,b,c,d,e\};\\
\bigcap \{A,B,C\} &= \{c\};\\
\bigcup D &= \{a,b,c,d,e\};\\
\bigcap D &= \{b\}.
\end{aligned}$$

The basis for these operations is the following axiom.

Axiom 2.16 (Union). For every $A \in V$ the union $\bigcup A$ of its elements is also a set and in V, where $\bigcup A = \{ \, x \mid \exists \, y \in A \; (x \in y) \, \}$.

Proposition 2.17. *Let $A, B \in V$. The sets constructed in Remark 2.15 exist, that is, are in V.*

Proof. This follows for $\bigcup A$ directly from the Axiom of Union, for the rest this is now shown. If A is not empty, the intersection is a member of V by the formula

$$\bigcap A = \{x \in \bigcup A \mid \forall y \in A \, (x \in y)\}$$

and using the Axiom of Comprehension. By the Axiom of Pair, $\{A, B\} \in V$. Thus $A \cup B = \bigcup\{A, B\}$ and $A \cap B = \bigcap\{A, B\}$ are in V. By the Axiom of Comprehension, the sets

$$\begin{aligned} A - B &= \{x \in A \cup B \mid x \notin B\} \text{ and} \\ A \triangle B &= \{x \in A \cup B \mid x \notin A \cap B\} \end{aligned}$$

exist and are in V as well.

Exercise 2.18. Prove that the symmetric difference is associative, that is, for all sets A, B, C, $(A \triangle B) \triangle C = A \triangle (B \triangle C)$. For this reason, one can just write $A \triangle B \triangle C$. Furthermore, prove that $A - B = A \cap (A \triangle B)$.

Exercise 2.19. Consider the sets *Apple*, *Pear*, *Strawberry*, *Cranberry*, *Blackberry*, *Banana*, *Blueberry* which consist of all fruits in the world usually designated by that name. Let *Fruits* be the union of these sets and *Red*, *Blue*, *Black* and *Yellow* be those elements of *Fruits* which have the corresponding colour. Which of the following expressions is the empty set?

1. $Apple - Red$,
2. $(Black \triangle Blueberry) \cap Blue$,
3. $Fruit - Red - Blue - Black - Yellow$,
4. $Red - Strawberry - Cranberry - Apple - Pear$,
5. $(Blueberry - Blue) \cup (Yellow - Apple - Pear - Banana)$,
6. $Banana - Yellow$,

7. $Banana \triangle Blueberry \triangle Strawberry \triangle Red$,

8. $(Strawberry \cup Blueberry \cup Cranberry) - Red$,

9. $(Apple \cup Pear) \cap (Strawberry \cup Blueberry)$,

10. $Fruit - \bigcup\{Apple, Pear, Strawberry, Cranberry, Blackberry, Banana\}$.

Give a set of three fruits which intersects those of the above sets which are not empty.

Property 2.20. *$A \subseteq B$ iff $A \cup B = B$ iff $A \cap B = A$.*

The next property states that the algebra of sets satisfies the following rules which it shares with any Boolean algebra.

Property 2.21. *All $A, B, C \in V$ satisfy the following laws.*

$$
\begin{array}{lrcl}
\text{Commutativity:} & A \cup B & = & B \cup A, \\
 & A \cap B & = & B \cap A; \\
\text{Associativity:} & (A \cup B) \cup C & = & A \cup (B \cup C), \\
 & (A \cap B) \cap C & = & A \cap (B \cap C); \\
\text{Distributivity:} & A \cup (B \cap C) & = & (A \cup B) \cap (A \cup C), \\
 & A \cap (B \cup C) & = & (A \cap B) \cup (A \cap C); \\
\text{De Morgan Laws:} & A - (B \cap C) & = & (A - B) \cup (A - C), \\
 & A - (B \cup C) & = & (A - B) \cap (A - C).
\end{array}
$$

Exercise 2.22. Many Boolean Algebras have a complementation operation, but here only the set difference is used in the De Morgan Laws. Why?

3 Functions

Graphs and functions are based on the notions of ordered pairs. For example, a graph consists of a set W of vertices and a set E of edges which is a set of ordered pairs of elements of W. These ordered pairs are constructed using the ordinary unordered pairs as follows.

Definition 3.1. $(x, y) = \{x, \{x, y\}\}$.

If $x \neq x'$ or $y \neq y'$ then $(x, y) \neq (x', y')$. This makes this definition suitable to introduce a representation for the Cartesian product of two sets.

Definition 3.2. For any sets A and B, the Cartesian product of A and B is the set $A \times B = \{\ (a, b) \mid a \in A \text{ and } b \in B\ \}$.

Example 3.3. *The Cartesian product of* $\{0, 1, 2\}$ *and* $\{3, 4\}$ *is* $\{\{0, \{0, 3\}\}, \{0, \{0, 4\}\}, \{1, \{1, 3\}\}, \{1, \{1, 4\}\}, \{2, \{2, 3\}\}, \{2, \{2, 4\}\}\}$. *The product of* $\{3, 4\}$ *and* $\{0, 1, 2\}$ *is* $\{\{3, \{0, 3\}\}, \{3, \{1, 3\}\}, \{3, \{2, 3\}\}, \{4, \{0, 4\}\}, \{4, \{1, 4\}\}, \{4, \{2, 4\}\}\}$. *Thus the Cartesian product is not commutative.*

Remark 3.4. Given $A, B \in V$, the existence of a Cartesian Product $A \times B$ is proven as follows. The sets $A \cup B$, $\mathcal{P}(A \cup B)$ and $C = \mathcal{P}(A \cup \mathcal{P}(A \cup B))$ are also in V. Now the following definition of the Cartesian Product is equivalent to the above one:

$$A \times B = \{c \in C \mid \exists a \in A \exists b \in B\, (c = \{a, \{a, b\}\})\}.$$

Thus $A \times B \in V$ by the Axiom of Comprehension.

A principal philosophy of set theory is that everything is coded or represented as a set, similarly to computer scientists who represent everything as finite sequences from $\{0, 1\}$. Thus the fundamental notions of relations and functions are defined in terms of sets. Thus, V does not only determine which sets exist, but also which functions, graphs and relations can be considered.

Definition 3.5. A *relation* is a subset of a Cartesian product of finitely many sets. A *graph* is given by a set G of vertices and a relation $E \subseteq G \times G$. A *function* $F : X \to Y$ is a graph, that is, a subset of $X \times Y$, such that for every $x \in X$ there is a unique $y \in Y$ with $(x, y) \in F$. This unique y will be denoted as $F(x)$, that is, $F(x) = y$ and $(x, y) \in F$ are equivalent notations for the same fact.

The set X is called the *domain* of F. It can be defined from F as $\{x \mid \exists (x', y') \in F\ (x = x')\}$. Y is a superset of the range $\{y \mid \exists (x', y') \in F\, (y = y')\}$. A function F is *one-to-one* (*or injective*) iff $(x, y) \in F$ and $(x', y) \in F$ implies that $x = x'$. A function $F : X \to Y$ is *onto* (*or surjective*) iff Y is equal to the range of F. F is a *bijective* function iff it is both injective and surjective.

Example 3.6. Let A, B be $\{0, 1, 2, 3, 4\}$ and f be given by $f(x) = x + 1$ for $x = 0, 1, 2, 3$ and $f(4) = 0$. Then the set representing f is $\{(0, 1), (1, 2), (2, 3), (3, 4), (4, 0)\}$; that is, f is identified with this set as in Definition 3.5.

Graphs and functions can also be defined on classes. So, in general, a function F from V to V would be a subclass of V whose members are of the form (x, y) with $x, y \in V$; that is, $F(x)$ would be the unique y with $(x, y) \in F$.

For this reason, it might be important to ask what classes exist and what classes do not exist. The answer will be the following: Subclasses of V exist iff they can be defined from other – already existing – subclasses of V using certain methods to make new classes. These methods will be introduced one after another, starting with basic expressions below.

Definition 3.7 (Expressions and Properties). An *basic expression* is either a variable-symbol or a constant which is a fixed member of V. For example, $\emptyset$ is a constant.

An *expression* is obtained from expressions by finitely often building new expressions from old ones. Given expressions $E_1, E_2, \ldots, E_n$, also the following constructions are expressions: $E_1 \cup E_2$, $E_1 \cap E_2$, $E_1 \triangle E_2$, $E_1 - E_2$, $\bigcup E_1$, $\bigcap E_1$, $\{E_1, E_2, \ldots, E_n\}$ (with $n \in \mathbb{N}$ and this expression being $\emptyset$ for $n = 0$), $\mathcal{P}(E_1)$.

A *basic property* is the comparison of two expressions with $\in$: $E_1 \in E_2$ is true iff E_1 is a member of E_2 and false otherwise; note that the truth-value can depend on variables built into E_1 and E_2.

A *property* is obtained from basic properties p_1, p_2 and expressions E by forming Boolean expressions and quantifying over variables: $\neg p_1$, $p_1 \wedge p_2$, $p_1 \vee p_2$, $\exists x \in E\ p_1(x)$ and $\forall x \in E\ p_1(x)$. Furthermore, one uses the symbols $\subseteq$ and $\subset$ as a shorthand:

$$\begin{aligned} E_1 \subseteq E_2 &\Leftrightarrow \forall x \in E_1\,(x \in E_2); \\ E_1 \subset E_2 &\Leftrightarrow (E_1 \subseteq E_2) \wedge \exists x \in E_2\,(\neg x \in E_1). \end{aligned}$$

For $\neg x \in E_1$ one can also write $x \notin E_1$.

One can also iterate the process by defining subexpressions using comprehension: namely if E is an expression and p a property then $\{x \in E \mid p(x)\}$ is also an expression.

All expressions and properties must be defined by finitely many iterations of the process outlined above.

Now every property defines a class. For example, if P is a property then $\{x \in V : P(x)\}$ defines a subclass of V. Furthermore, if E is an expression, then $\{(x, E(x)) : x \in V\}$ defines a subclass of V which is a function.

Example 3.8. The following are expressions where $a, b, c, \ldots$ stand for constants and u, v, w, x, y, z for variables; in some statements, the informal definition is given first and the correct formal one is the last in the chain of equations.

1. $\mathcal{P}(x)$, $\mathcal{P}(\mathcal{P}(\emptyset))$ and $\mathcal{P}(\bigcup (x \cup a))$;
2. $x \times y = \{z \in \mathcal{P}(x \cup y \cup \mathcal{P}(x \cup y)) \mid \exists v \in x \exists w \in y\,(z = \{v, \{v, w\}\})\}$;
3. $\{y \mid y \text{ is a graph on } x\} = \{(x, u) \mid u \subseteq x \times x\} = \{x\} \times \mathcal{P}(x \times x)$;

The following properties are either always true or always false.

1. $x = x$;
2. $x \in x$;
3. $x = \bigcup\{x\}$;
4. $\exists z \in \mathcal{P}(x)\,(z \in \mathcal{P}(y))$;
5. $x = \mathcal{P}(x)$;
6. $\forall y \in \mathcal{P}(x)\,\forall z \in y\,(z \in x)$;
7. $\exists y \in \mathcal{P}(\mathcal{P}(x))\,(y \not\subseteq \emptyset)$.

The property $\exists y \in \mathcal{P}(x)\,(y \not\subseteq \emptyset)$ is true iff $x \neq \emptyset$. So it is not always true and therefore the last property needed to be more complicated to go into the list of all always true properties. Properties which are always true are called tautologies.

Given the expression $\mathcal{P}(x)$, one can define the function $\{(x, \mathcal{P}(x)) : x \in V\}$ mapping x to $\mathcal{P}(x)$ as a subclass of V. Similarly, one can code the function $x, y \mapsto \mathcal{P}(x) \cup ((x \cap y) \times (x \cup y))$ as a class.

Set Theory following the axioms of Zermelo and Fraenkel does not have a formalized concept of classes. It is more the way that everything is a class which has a definition which can be written down using standard set-theoretic terminology, parameters from V (that is using sets) and also using the various recursion theorems explained in chapters to come. Somehow, to work this out formally, goes beyond what is mandatory for a student to learn; therefore, it will only be explained how to make new functions from old ones, but the whole mechanism is spared out. So a student should know these things for examinations about functions from V to V:

- A function f from V to V is a subclass of V which consists of ordered pairs (x, y) such that for every $x \in V$ there is a unique $y \in V$ with $(x, y) \in f$; one writes $f(x) = y$ for this pair (x, y).
- Functions from V to V can be concatenated and modified in the standard way.
- A function f from a set X to a set Y can be extended to a function from V to V by considering the subclass $\{(x, y) \mid x \in X \wedge y = f(x) \vee x \notin X \wedge y = \emptyset\}$ of V.
- Functions can be defined from other functions using the various types of recursion defined in Sections to come.
- Functions from V to V satisfy the below Axiom of Replacement.

This axiom is used in order to make sure that the range of a set under a function (from V to V) is again a set.

Axiom 3.9 (Replacement). Let f be a subclass of V which is a function from V to V, that is, f is a class of pairs (x, y) such that for each $x \in V$ there is a unique $y \in V$ with $(x, y) \in f$. Then, for every set $X \in V$, both $f(X)$ and $f[X] = \{f(y) : y \in X\}$ are sets again.

Exercise 3.10. Define (informally) functions f_n from $\mathbb{N}$ to $\mathbb{N}$ with the following properties:

1. f_1 is bijective and satisfies $f_1(x) \neq x$ but $f_1(f_1(x)) = x$ for all $x \in \mathbb{N}$;
2. f_2 is two-to-one: for every y there are exactly two elements $x, x' \in \mathbb{N}$ with $f(x) = f(x') = y$;
3. f_3 is dominating all polynomials, that is, for every polynomial p there is an x such that for all $y > x$, $f_3(y) > p(y)$;
4. f_4 satisfies $f_4(x+1) = f_4(x) + 2x + 1$ for all $x \in \mathbb{N}$;
5. $f_5(x) = \begin{cases} 0 & \text{if } x = 0; \\ f_5(x-1) & \text{if } x > 0 \text{ and } x \text{ is not a square number;} \\ f_5(x-1)+1 & \text{if } x > 0 \text{ and } x \text{ is a square number.} \end{cases}$

Determine the range of the function f_4.

Definition 3.11. Let $F : A \to B$ and $G : B \to C$. The composition of F and G is the function $G \circ F : A \to C$ defined by $G \circ F(a) = G(F(a))$ for every $a \in A$. That is, $G \circ F = \{\ (a, c) \mid \exists b \in B\, ((a, b) \in F \wedge (b, c) \in G)\}$.

Exercise 3.12. Let $A = \{0, 1, 2\}$ and $F = \{f : A \to A \mid f = f \circ f\}$. Show that F has exactly 10 members and determine these.

Definition 3.13. Let $f : A \to B$. Let $C \subseteq A$. Let $D \subseteq B$.

1. The *restriction* of f to C, denoted by $f{\restriction}C$, is the function $f \cap (C \times B)$.
2. $f[C]$ is the subset of B determined by

$$f[C] = \{\ b \in B \mid \exists\, a \in C\ ((a, b) \in f)\ \}.$$

   $f[C]$ is the *image* of C under f.

3. $f^{-1}[D]$ is the subset of A determined by

$$f^{-1}[D] = \{ \, a \in A \mid \exists \, b \in D \; ((a, b) \in f) \, \}.$$

$f^{-1}[D]$ is the *preimage* of D under f.

Example 3.14. *Let $f : n \mapsto n + 5$ be a function on the natural numbers. Then the set $f[\{3, 4, 5, 6\}] = \{8, 9, 10, 11\}$ is the image of $\{3, 4, 5, 6\}$ and $f^{-1}[\{5, 6, 7, 8\}] = \{0, 1, 2, 3\}$ is the preimage of $\{5, 6, 7, 8\}$ as $f[\{0, 1, 2, 3\}] = \{5, 6, 7, 8\}$. The preimage of $\{0, 1, 2, 3, 4\}$ is the empty set.*

Definition 3.15. Let A and B be sets. The set of all functions from A to B is denoted by B^A and $\{f : A \to B\}$.

For example, $\{0, 1\}^{\mathbb{N}}$ is the set of all functions from $\mathbb{N}$ to $\{0, 1\}$, that is, of all binary sequences.

Example 3.16. *For given nonempty sets A, B, C, let*

$$D = \{f \in C^A \mid \exists g \in B^A \exists h \in C^B \; (f = h \circ g)\}.$$

Then, depending on the choice of A, B, C, either $D \subset C^A$ or $D = C^A$.

Proof. If A or C has only one element, then A^C consists of all constant functions from A to C and so does D, hence $C^A = D$.

So let A, C be any sets with at least two elements. Obviously every function going from A to B and then from B to C is a function from A to C. Thus $D \subseteq C^A$.

If B has exactly one element, that is, $B = \{b\}$ for some b, then every function in D is constant: For all $a, a' \in A$, $f(a) = h(g(a)) = h(b) = h(g(a')) = f(a')$. But there is also a nonconstant function in C^A: By choice there are two distinct elements c, c' in C. Fix furthermore an element $a \in A$. Now define $f(a) = c$ and $f(a') = c'$ for all $a' \in A - \{a\}$. This function is not constant since A has at least two elements. Thus $f \notin D$ and $D \subset C^A$.

If $B = A$ and $f \in C^A$, then one can take g to be the identity and $h = f$. Now $h \circ g = f$ and thus $D = C^A$. ∎

4 Natural Numbers

In computing, everything is coded as a binary sequence. For example, the American Standard Code for Information Interchange (ASCII) represents the digit "0" as

00110000, the digit "9" as 00110101 and the letter "A" as 01000001. A consequence is that unsplittable objects like a letters and digits can now be split into subparts. So the digits have all the prefix 0011 followed by 0000, 0001, ..., 0101. Both parts can be split again into 4 binary digits 0 and 1 which are the only primitive objects.

In Set Theory, every object is represented as a set. Thus, primitive objects like the natural numbers have, due to this representation, elements of their own. Against the intuition, they are no longer unsplittable elements which cannot decomposed furthermore. Since one cannot avoid that a number has elements, one tries to represent them as natural as possible. That is, the number n has as elements those sets which represent the numbers m below n. Furthermore, the numbers are identified with their representation

$$0 = \emptyset, 1 = \{\emptyset\}, 2 = \{\emptyset, \{\emptyset\}\}$$

and this representation will be kept fixed. Note that $1 = \{0\} = 0 \cup \{0\}$, $2 = \{0, 1\} = 1 \cup \{1\}$, $3 = \{0, 1, 2\} = 2 \cup \{2\}$, $4 = \{0, 1, 2, 3\} = 3 \cup \{3\}$ and so on. This can be formalized using the notion of a successor S: $S(x) = x \cup \{x\}$.

Given a finite set X of natural numbers, one can easily see from the way they are coded that $\bigcup X$ is the maximum of X. So $S(\bigcup X)$ gives a new element of the natural numbers which is outside the set X. This means, that the set of natural numbers is provably infinite. Therefore, ensuring that the natural number are in V means to ensure that a provably infinite set is in V. If a set X shares the basic properties of the natural numbers that it contains 0 and for every n also the successor $S(n)$, then it is called *inductive*. Inductive sets are always infinite.

Definition 4.1. A set X is an *inductive set* iff $\emptyset \in X$ and X is closed under S: $\forall y \in X\, (S(y) \in X)$.

Axiom 4.2 (Infinity). There exists an inductive set.

Definition 4.3. The set $\emptyset$ is also called 0 and inductively the set $S(n)$ is called $n+1$ where S is the function given by $S(x) = x \cup \{x\}$. Let X be any inductive set. Then call $\mathbb{N} = \bigcap\{Y \subseteq X : Y \text{ is inductive}\}$.

One can show that the definition of $\mathbb{N}$ is unique, so it does not matter which inductive set one chooses to start. One can now ask, is $\mathbb{N}$ the set of all natural numbers? That is, is $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ true? From the axioms, one cannot prove this because one cannot prove that $\{0, 1, 2, 3, \ldots\}$ is a set. One can only prove the following things:

- $\mathbb{N}$ contains every natural number as every inductive set contains every natural number;

- $\mathbb{N}$ is inductive;
- every inductive subset X of $\mathbb{N}$ satisfies $\mathbb{N} = X$;
- if the collection $\{0, 1, 2, \ldots\}$ is actually a set then $\mathbb{N} = \{0, 1, 2, \ldots\}$;
- all the laws which one learns in lectures about $\{0, 1, 2, \ldots\}$ and which can be written down formally are true for $\mathbb{N}$.

For this reason, most books of set theory call "$\mathbb{N}$" just "the set of natural numbers". But this is not completely precise as will be shown in Section 20. Here some explanations.

The problem behind this is a problem of axioms: as long as formulas are finite and one writes down the axioms using the language of set theory by accessing V and $\in$ and using variables for members of V, quantifiers and Boolean connectives, one cannot come up with any set of finite formulas which define $\mathbb{N}$ better than done above. Such language of set theory would include a sentence like "there is a set X such that for all elements of $y \in X$, y contains an element z and z is not empty", but it does not contain an infinite formula like $\forall x \in \mathbb{N}\,(x = 0 \vee x = 1 \vee x = 2 \vee x = 3 \vee \ldots)$ which would be needed to make sure that $\mathbb{N} = \{0, 1, 2, \ldots\}$ and does not contain anything else. This situation is a bit unpleasant as everyone assumes that the existence of the set of natural numbers is self-evident, but unfortunately it is not. Nevertheless, $\mathbb{N}$ behaves like the set of natural numbers should behave and for that reason, in many set-theory books $\mathbb{N}$ is just called the set of natural numbers and assumed that $\mathbb{N} = \{0, 1, 2, \ldots\}$ would be guaranteed.

In the years 1945 – 1991, the Soviet Union was a member of the United Nations (UNO). The Soviet Union itself had 15 republics as its members: Soviet Union = {Belarus, Estonia, Latvia, Lithuania, Ukraine, ...}. Two of them, Belarus and Ukraine, were not only members of the Soviet Union but also of the UNO and represented in the general assembly. So one had the situation that Ukraine $\in$ Soviet Union $\in$ UNO and Ukraine $\in$ UNO at the same time. If this situation would not only occur at some but at all places, that is, if every member of a member of the UNO would be already be a direct member of the UNO, then one would call the UNO to be "transitive".

Definition 4.4. A set A is called *transitive* iff for all $a \in A$ and $b \in a$ it holds that $b \in A$ as well.

Example 4.5. *The set $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \{\{\{\{\emptyset\}\}\}\}\}$ is transitive, but its elements are not. Furthermore, $\mathbb{N}$ is transitive.*

Proof of Second Statement. Given $\mathbb{N}$, let $A = \{X \in \mathbb{N} \mid X \subseteq \mathbb{N}\}$. Clearly $\emptyset \in A$. Now consider any $X \in A$. By definition of A, $X \in \mathbb{N}$ and $X \subseteq \mathbb{N}$. Thus $S(X) =$

$X \cup \{X\} \subseteq \mathbb{N}$. Furthermore, $S(X) \in \mathbb{N}$ as $\mathbb{N}$ is inductive. So $X \in A \Rightarrow S(X) \in A$ for all $X \in A$. It follows that A is an inductive subset of $\mathbb{N}$; as $\mathbb{N}$ is the smallest inductive set, $A = \mathbb{N}$ and every $X \in \mathbb{N}$ satisfies $X \subseteq \mathbb{N}$. Hence, $\mathbb{N}$ is transitive. ▮

Exercise 4.6. Which of the following sets is transitive and which is inductive?

1. $A = \{\emptyset, \{\emptyset\}\}$,
2. $B = \{\emptyset, \{\{\{\emptyset\}\}\}\}$,
3. $C = \{x \mid \forall y \in x \, \forall z \in y \, (z = \emptyset)\}$,
4. D is the closure of $\{\emptyset, \mathbb{N} \times \mathbb{N}\}$ under the successor operation $x \mapsto S(x)$,
5. E is the set of even numbers,
6. F is the set of all natural numbers which can be written down with at most 256 decimal digits,
7. G is the set of all finite subsets of $\mathbb{N}$,
8. $H = \mathcal{P}(G)$.

Exercise 4.7. Show that the following statements are equivalent for any inductive set X.

1. $X = \mathbb{N}$;
2. X has no proper inductive subset;
3. X is a subset of every inductive set;
4. $\forall x \in X \, (x = 0 \vee \exists y \in X \, (x = S(y)))$;
5. $X = N(Y)$ for every inductive set Y where $N(Y)$ is the subset of those $y \in Y$ which are in every inductive subset of Y;
6. $X = N(Y)$ for some inductive set Y where $N(Y)$ is defined as in the previous item.

Theorem 4.8 (Mathematical Induction Principle). *Let $\phi(x)$ be a property. Assume that*

1. *$\phi(0)$ holds and*

2. *for all* $n \in \mathbb{N}$, $\phi(n)$ *implies* $\phi(n+1)$.

Then $\phi(n)$ *holds for every natural number* $n \in \mathbb{N}$.

Proof. Let $A = \{n \in \mathbb{N} \mid \phi(n) \text{ holds}\}$. Then A is an inductive set. Hence $\mathbb{N} \subseteq A$. In particular, $\phi(n)$ is true for all $n \in \mathbb{N}$. ▮

Exercise 4.9. Assume that a property p satisfies

$$p(1) \text{ and } \forall x\,(p(x) \Rightarrow p(S(S(x)))).$$

Show that $p(x)$ is true for all odd numbers. Assume now that one chooses the property p to identify the odd numbers, that is,

$$p(x) \Leftrightarrow \exists y \in \mathbb{N}\,(x = y + y + 1).$$

Show that p then satisfies the above condition.

The following properties follow from the fact that the natural numbers are defined such that $n = \{m \in \mathbb{N} \mid m < n\}$. This means in particular $n < m$ iff $n \in m$. So, for example, the third item is then just the transitivity of $<$ on $\mathbb{N}$.

Property 4.10. *Assume that* $m, n, k \in \mathbb{N}$.

1. *Either* $0 = n$ *or* $0 \in n$.
2. $k \in n + 1$ *if and only if* $k \in n$ *or* $k = n$.
3. *If* $k \in m$ *and* $m \in n$, *then* $k \in n$.
4. *If* $n \in m$, *then* $m = n + 1$ *or* $n + 1 \in m$.
5. *Either* $m \in n$ *or* $m = n$ *or* $n \in m$.
6. $n \neq n + 1$.
7. *If* $n \neq m$ *then* $n + 1 \neq m + 1$.
8. $n \subset \mathbb{N}$.

Proposition 4.11. *Let* A *be a transitive nonempty set. Then* $A \subseteq \mathbb{N}$ *iff* 0 *is the unique element which is not the successor of any other element of* A.

Proof. First consider the case $A \subseteq \mathbb{N}$. By the Axiom of Foundation there is an $y \in A$ such that every $z \in A$ satisfies $z \notin y$. On the other hand, every $z \in y$ satisfies $z \in A$.

Thus y has no elements and $y = 0$. Hence $0 \in A$.

Now consider any transitive A and assume now that $x \in A - \{0\}$ and $x \neq S(y)$ for all $y \in A$. If x is an inductive set then $\mathbb{N} \subseteq x$ and $x \notin \mathbb{N}$ by the Axiom of Foundation. If x is not an inductive set then there is a $y \in x$ with $S(y) \notin x$. Note that $y \in A$ since A is transitive. If $y \notin \mathbb{N}$ then $x \notin \mathbb{N}$ as $\mathbb{N}$ is transitive. If $y \in \mathbb{N}$ then $x \neq S(y)$ by choice of x. But $S(y)$ is the only natural number z such that $y \in z \wedge S(y) \notin z$. It follows that $x \notin \mathbb{N}$ again. So whenever $A \subseteq \mathbb{N}$ and A not empty then 0 is the unique element of A which is not the successor of any other element of A.

Last consider the case that $A \not\subseteq \mathbb{N}$ and let $B = A - \mathbb{N}$. By the Axiom of Foundation there is an element $x \in B$ with $y \notin x$ for all $y \in B$. Let z be any element of A. If $z \in \mathbb{N}$ then $S(z) \in \mathbb{N}$ and $S(z) \neq x$. If $z \in B$ then $z \in S(z)$ and thus again $S(z) \neq x$. So the element x of A is different from 0 and not the successor of any other element of A. ▮

Corollary 4.12. *A set A is equal to $\mathbb{N}$ iff A is transitive, inductive and 0 is the unique element of A which is not the successor of any other element of A.*

Remark 4.13. *Let X be a nonempty set and $n \in \mathbb{N}$.*

1. *If $X \subseteq \mathbb{N}$, then there is a unique $m \in X$ such that $m \cap X = \emptyset$.*

2. *If $X \subseteq n$ then there is a unique $m \in X$ such that $m \cap X = \emptyset$.*

The proof of this is immediate by Axiom 1.3 (Foundation); indeed, the statement is literally almost the same. But the interpretation of this statement is that every nonempty of the set of natural numbers has a smallest element, which is m above, as $m \cap X = \emptyset$ implies that there is no $n < m$ in X. For this reason, the two statements are explicitly listed here.

5 Recursive Definition

Functions with domain the set of natural numbers can be defined recursively.

Example 5.1. One can define $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ as follows: For each $m \in \mathbb{N}$, define $f_m : \{m\} \times \mathbb{N} \to \mathbb{N}$ by

$$f_m(m, 0) = m, \; f_m(m, S(n)) = S(f_m(m, n)).$$

Then $\bigcup\{f_m \mid m \in \mathbb{N}\} = \{(m, n, f_m(n)) \mid m, n \in \mathbb{N}\}$ represents the addition, that is, $m + n = f_m(n)$ for all $m, n \in \mathbb{N}$.

Similarly, one can also define $\cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ as follows: For each $m \in \mathbb{N}$, define

$g_m : \{m\} \times \mathbb{N} \to \mathbb{N}$ by

$$g_m(m, 0) = 0,\ g_m(m, S(n)) = g_m(m, n) + m.$$

Then $\bigcup\{g_m \mid m \in \mathbb{N}\} = \{(m, n, g_m(n)) \mid m, n \in \mathbb{N}\}$ represents the multiplication, that is, $m \cdot n = g_m(n)$ for all $m, n \in \mathbb{N}$.

Furthermore the factorial $m!$ can be defined by $0! = 1$ and $S(n)! = n! \cdot S(n)$. Similarly, 2^n can be defined by $2^0 = 1$ and $2^{S(n)} = 2 \cdot 2^n$.

Theorem 5.2 (The Recursion Theorem). *Assume that $g : \mathbb{N} \times V \to V$ is a function and a a value. Then there exists a unique function $f : \mathbb{N} \to V$ such that*

1. *$f(0) = a$ and*
2. *$f(S(n)) = g(n, f(n))$ for all $n \in \mathbb{N}$.*

Proof. Let C be the class of all functions with domain $\mathbb{N}$ such that $h(0) = a$ and for all $b \in \mathbb{N}$ with $h(S(b))$ being defined, $h(b)$ is defined as well and $h(S(b)) = g(b, h(b))$. Define the class

$$f = \{(b, c) \mid b \in \mathbb{N} \wedge \exists h \in C\ ((b, c) \in h)\}$$

It is now shown that f is actually a function from $\mathbb{N}$ to V. This is done by considering the following two subclasses of $\mathbb{N}$:

$$\begin{aligned} D &= \{d \in \mathbb{N} \mid \exists c \in V\ ((d, c) \in f)\}; \\ E &= \{e \in D \mid \forall c, \tilde{c} \in V\ ((e, c), (e, \tilde{c}) \in f \Rightarrow c = \tilde{c})\}. \end{aligned}$$

By the Axiom of Comprehension, they are subsets of $\mathbb{N}$. Both sets contain 0 as there is a function with domain $\{0\}$ and value a and as every member of C contains only the pair $(0, a)$ but not any other pair.

Assume that $D \neq \mathbb{N}$ then there is a minimum $d \in \mathbb{N} - D$. As $0 \in D$, $d = S(b)$ for some b. Let h be a member of C with b in its domain, such a member exists by the choice of f, C and D. As $h(d)$ is undefined, one can look at the function $\tilde{h} = h \cup \{(d, g(b, h(b)))\}$. This function is also in the class C, hence $(d, \tilde{h}(b)) \in f$ in contradiction to the assumption on C. This contradiction gives $D = \mathbb{N}$.

Assume that $E \neq \mathbb{N}$. Then there is a minimum $e \in \mathbb{N} - E$. By choice there are two functions $h, \tilde{h} \in C$ such that $h(e) \neq \tilde{h}(e)$. As $e > 0$, there is a number b such that $e = S(b)$. Furthermore, $h(b) = \tilde{h}(b)$. By definition of C, $h(e) = g(b, h(b)) = g(b, \tilde{h}(b)) = \tilde{h}(e)$, in contradiction to the assumption on e. Hence, e cannot exist and $E = \mathbb{N}$.

In summary, f is a function from $\mathbb{N}$ to V. By the Axiom of Replacement, $f[\mathbb{N}]$

is a set and f actually a function from a set to a set. Furthermore, $f(0) = a$ and $f(S(n)) = g(n, f(n))$ for all $n \in \mathbb{N}$. If any function $\tilde{f} : \mathbb{N} \to V$ satisfies conditions 1. and 2., then the restriction of $\tilde{f}$ to each domain $n \in \mathbb{N}$ is a member in C and that member is extended by f; hence $\tilde{f} = f$. So f is the only function satisfying the recursive definition. ∎

This theorem has some applications. The first one is that for each set X there is a transitive closure $\mathcal{TC}(X)$ of X. Although it is not stated in the next proposition, one can even show that the operation $X \mapsto \mathcal{TC}(X)$ is a function from V to V which is represented as a subclass of V. Note that the concept of a transitive closure is quite natural: it contains the elements plus element of elements plus elements of elements of elements plus ... of a set. One could also define it as the intersection of all transitive supersets:

$$\mathcal{TC}(X) = \bigcap \{Y : \{X\} \subseteq Y \subseteq Z \text{ and } Y \text{ is transitive}\},$$

where Z is any transitive set containing X to start with, in the same way as the natural numbers were defined as the intersection of all inductive subsets of some given inductive set. The definition is not sensitive to which set Z is chosen.

Proposition 5.3. *For every set X there is a set $\mathcal{TC}(X)$ which is the smallest transitive set containing X as an element.*

Proof. For and any $n \in \mathbb{N}$ and $y \in V$, let

$$g(n, y) = y \cup \bigcup y = y \cup \{x \mid \exists z\, (x \in z \wedge z \in y)\}$$

consist of all elements of y plus all elements of elements of y. The function g corresponds clearly to a class as it is written down as an expression. Thus there is a function $f : \mathbb{N} \to V$ with $f(0) = \{X\}$ and $f(S(n)) = g(n, f(n))$ for all $n \in \mathbb{N}$. Now $f[\mathbb{N}]$ is a set which coincides with the union of all $f(n)$.

Now define $\mathcal{TC}(X) = \bigcup f[\mathbb{N}]$. As $f(0) = \{X\}$, $X \subseteq \mathcal{TC}(X)$.

The set $\mathcal{TC}(X)$ is transitive: if $z \in \bigcup f[\mathbb{N}]$ and $x \in z$ then there is an $n \in \mathbb{N}$ with $z \in f(n)$. It follows from the definition that $x \in f(S(n))$ and $x \in \bigcup f[\mathbb{N}]$.

If $Y \supseteq X$ and Y is transitive then $Y \supseteq \mathcal{TC}(X)$; that is, $\mathcal{TC}(X)$ is the smallest transitive superset of $\{X\}$: To see this, let Y be any transitive superset of $\{X\}$. Clearly $f(0) \subseteq Y$. Furthermore, if $f(n) \subseteq Y$ then $f(S(n))$ contains only elements which are either in $f(n)$ or elements of elements of $f(n)$. As $f(n) \subseteq Y$, the elements and also the elements of elements of $f(n)$ are also elements and elements of elements of Y, respectively. As Y is transitive, $f(S(n)) \subseteq Y$. It follows from the induction principle that every set $f(n)$ with $n \in \mathbb{N}$ is a subset of Y. Then $\mathcal{TC}(X) = \bigcup f[\mathbb{N}]$ is also a subset of Y. ∎

Proposition 5.4. *Every set X is a subset of an inductive set.*

Proof. Using Recursion, there is a function $f : \mathbb{N} \rightarrow V$ with $f(0) = X \cup \{\emptyset\}$ and $f(S(n)) = \{S(y) \mid y \in f(n)\}$ for all $n \in \mathbb{N}$. Now $Y = \bigcup f[\mathbb{N}]$ is an inductive set: it contains $\emptyset$ as $\emptyset \in f(0)$; for every $y \in Y$ there is an $n \in \mathbb{N}$ with $y \in f(n)$ and $S(y) \in f(n+1)$, hence $S(y) \in Y$ as well. Furthermore, $X \subseteq Y$. ▮

Example 5.5. For every function $h : \mathbb{N} \rightarrow \mathbb{N}$ there exists a function f such that $f(n) = h(0) + h(1) + \ldots + h(n)$. This is a consequence of the Recursion Theorem which takes f to be the function satisfying $f(n+1) = g(n, f(n))$ where $g(n, f(n)) = f(n) + h(n)$. Here some examples for such recursively defined functions f and the functions h on which they are based.

Function h	Sum f, $f(n) = h(0) + h(1) + \ldots + h(n)$
1	$n+1$
n	$n^2/2 + n/2$
n^2	$n^3/3 + n^2/2 + n/6$
n^3	$n^4/4 + n^3/2 + n^2/4$
3^n	$(3^{n+1} - 1)/2$

Exercise 5.6. Determine the functions f_n given by the following recursive equations:

1. $f_1(0) = 0$, $f_1(S(n)) = f_1(n) + 2^n$,
2. $f_2(0) = 1$, $f_2(1) = 1$, $f_2(n+2) = f_2(n) \cdot (n^2 + 3n + 2)$,
3. $f_3(n) = 1$ for $n = 0, 1, \ldots, 9$, $f_3(10n + m) = f_3(n) + 1$ for $n = 1, 2, \ldots$ and $m = 0, 1, \ldots, 9$,
4. $f_4(0) = 0$, $f_4(1) = 0$, $f_4(2) = 0$, $f_4(3) = 1$, $f_4(S(n)) = f_4(n) + \frac{1}{2}(n^2 - n)$ for $n > 2$,
5. $f_5(n) = 1$, $f_5(S(n)) = 256 \cdot f_5(n)$.

Give informal explanations what these functions compute, for example, consider f_6 given by $f_6(0) = 0$, $f_6(1) = 0$ and $f_6(S(n)) = f_6(n) + 2n$ for $n \geq 1$. Then $f_6(n) = n(n-1)$. One explanation would be to assume that there is a soccer league with n teams. Then there are $f_6(n)$ games per season, each pair $\{A, B\}$ of two different teams plays once at A's place and once at B's place. Another example would be coming from set theory; let the function $f_7(n)$ say how many subsets of $n = \{0, l, \ldots, n-1\}$ have an even number of elements. Then f_7 is given by $f_7(0) = 1$, $f_7(1) = 1$ and $f_7(n+1) = 2 \cdot f_7(n)$ for all $n > 0$.

Recall the definition of domination from Exercise 3.10: A function $h : \mathbb{N} \to \mathbb{N}$ *dominates* a function $f : \mathbb{N} \to \mathbb{N}$ if there is an n such that $\forall m \geq n\,(f(m) < h(m))$. A function $f : \mathbb{N} \to \mathbb{N}$ is *unbounded* if it dominates every constant function $c : \mathbb{N} \to \mathbb{N}$ satisfying $\forall n, m\,(c(m) = c(n))$. A function $f : \mathbb{N} \to \mathbb{N}$ is *increasing* if $\forall n\,(f(n) \leq f(S(n)))$.

Theorem 5.7. *Given a function $H : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that for every $m \in \mathbb{N}$ the function h_m given as $h_m(n) = H(m, n)$ is increasing and unbounded, there is an unbounded and increasing function f such that every h_m dominates f.*

Proof. This is proven using the Recursion Theorem with a function g defined as

$$g(n,k) = \begin{cases} S(k) & \text{if } S(k) < h_m(S(n)) \text{ for all } m \in k; \\ k & \text{otherwise.} \end{cases}$$

This definition can be realized as an expression since one can access $H(m, S(n))$ in order to get $h_m(S(n))$. So the function g can be defined using only one parameter from V, namely H. This is necessary because the set $\{h_0, h_1, \ldots\}$ might be outside V if H would not be there.

There is a function f such that $f(0) = 0$ and $\forall n \in \mathbb{N}\,(f(S(n)) = g(n, f(n)))$. Now it is verified that f satisfies all necessary requirements.

1. *f is increasing.* This follows directly from the definition of g: for all n, $f(S(n)) \in \{f(n), S(f(n))\}$.

2. *f is unbounded.* Let k be the value of f at some place i. For every $m \in k$ there is a value n_m such that $h_m(n_m) > S(k)$ since h_m is unbounded. Let $j = 2 + \max\{i, n_0, n_1, \ldots, n_{k-1}\}$. Then either $f(j) > k$ or $f(j) = k \wedge g(j, f(j)) = S(k) \wedge f(S(j)) = S(k)$. Thus f is going up at infinitely many places and unbounded.

3. *Every h_m dominates f.* Assume by contradiction that h_m does not dominate f. Since f is unbounded and increasing, there is a first value n such that $f(S(n)) > S(m)$ and $f(S(n)) \geq h_m(S(n))$. Since $f(0) = 0$, one can conclude that either $f(n) \leq S(m)$ or $f(n) < h_m(n)$. Since h_m is increasing, one has in both cases that $f(S(n)) = S(f(n))$ and it follows from the definition of g that $S(f(n)) < h_{m'}(S(n))$ for all $m' \in f(n)$. Since $f(S(n)) > S(m)$, one has $f(n) \geq S(m)$ and $m \in f(n)$. Thus $f(S(n)) < h_m(S(n))$ in contradiction to the choice of n. So the n chosen cannot exist and h_m dominates f.

So f is unbounded, increasing and dominated by every h_m with $m \in \mathbb{N}$. ∎

Exercise 5.8. Let $H : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be a function and $h_m : \mathbb{N} \to \mathbb{N}$ be given by $h_m(n) = H(m, n)$ for all n. Show that there is a function f dominating every h_m; that is, show that there is a function f such that for every m there is an n with $f(k) > h_m(k)$ for all $k > n$.

6 Cardinality of Sets

The cardinality of a set is the number of its elements. For example, the set {Adelaide, Brisbane, Canberra, Melbourne, Perth, Sydney} of Australia's largest towns has six members, that is, its cardinality is 6. This is established by counting, Adelaide is the first, Brisbane the second, Canberra the third, Melbourne the fourth, Perth is the fifth and Sydney the sixth town. Mathematically, this can be viewed as a bijective mapping from the set of the largest Australian towns to the set {first, second, third, fourth, fifth, sixth} representing 6. When working out the foundations of set theory, Cantor defined that two sets have the same size iff there is a bijective function between them.

Definition 6.1. Let A and B be two sets. The sets A and B have *the same cardinality*, denoted by $|A| = |B|$, iff there exists a bijective function from A to B.

Property 6.2. *Having the same cardinality is a transitive and reflexive equivalence relation. That is, the following holds for all sets A, B, C:*

1. $|A| = |A|$;
2. *if* $|A| = |B|$, *then* $|B| = |A|$;
3. *if* $|A| = |B|$ *and* $|B| = |C|$, *then* $|A| = |C|$.

Example 6.3. *Let $E \subset \mathbb{N}$ be the set of all even natural numbers. Then $|E| = |\mathbb{N}|$ which is witnessed by the bijection $x \to 2 \cdot x$ from $\mathbb{N}$ to E.*

Proposition 6.4. $|\mathcal{P}(X)| = |\{0,1\}^X|$.

Proof. For each $A \subseteq X$, let $f(A)$ be the characteristic function of A, defined by

$$f(A)(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Then f is one-to-one and onto. ∎

Proposition 6.5. *If $C \subseteq B \subseteq A$ and $|C| = |A|$ then $|A| = |B|$.*

Proof. Let $f : A \to C$ be a bijective function. One recursively defines two sequences $A_0, A_1, \ldots, A_n, \ldots$ and $B_0, B_1, \ldots, B_n, \ldots$ of sets as follows. Let $A_0 = A$ and $B_0 = B$ and for each $n \in \mathbb{N}$, let

$$A_{n+1} = f[A_n], \quad B_{n+1} = f[B_n].$$

By induction, $A_{n+1} \subseteq B_n \subseteq A_n$ for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, let $E_n = A_n - B_n$. Note that $f[E_n] = f[A_n] - f[B_n] = A_{n+1} - B_{n+1} = E_{n+1}$ which can be proven by induction. Thus there is a set E such that

$$E = \bigcup\{E_n \mid n \in \mathbb{N}\}.$$

The idea is now to use that all the E_n are disjoint and that $E_0 = A - B$. The new function g will be a one-to-one mapping from each E_n to E_{n+1} where it follows f and will be the identity otherwise:

$$g(x) = \begin{cases} f(x) & \text{if } x \in E; \\ x & \text{if } x \in A - E. \end{cases}$$

Then g is one-to-one and $g[A] = \bigcup\{E_n \mid n \geq 1\} \cup (A - E) = A - E_0 = B$. ∎

Definition 6.6. Let A and B be two sets. *The cardinality of A is less than or equal to the cardinality of B*, denoted by $|A| \leq |B|$, iff there exists a one-to-one function $f : A \to B$. *The cardinality of A is less than the cardinality of B*, denoted by $|A| < |B|$, if $|A| \leq |B|$ and there is no one-to-one function from B into A.

Exercise 6.7. Prove by giving a one-to-one function that the set {Auckland, Christchurch, Dunedin, Wellington} of New Zealand's largest towns has a cardinality which is less than the one of the set of Australian towns given above. Furthermore, prove that it is not less or equal than the cardinality of the set {Singapore}.

Property 6.8. *Each $A, B, C \in V$ satisfy the following:*

1. $|A| \leq |A|$*;*
2. *if* $|A| = |B|$ *then* $|A| \leq |B|$*;*
3. *if* $|A| \leq |B|$ *and* $|B| \leq |C|$ *then* $|A| \leq |C|$*.*

Theorem 6.9 (Cantor-Bernstein Theorem). *If $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$.*

Proof. Let $f : X \to Y$ and $g : Y \to X$ be one-to-one functions. Then $g[f[X]] \subseteq g[Y] \subseteq X$. Since $|X| = |g[f[X]]|$, by Proposition 6.5, $|X| = |g[Y]|$. Since $|Y| = |g[Y]|$, $|X| = |Y|$. ∎

Example 6.10. *The sets $\{0,1\}^{\mathbb{N}}$ and $\{0,1,2\}^{\mathbb{N}}$ have the same cardinality.*

Proof. Every $\{0,1\}$-valued function is also a $\{0,1,2\}$-valued function, thus $\{0,1\}^{\mathbb{N}} \subseteq \{0,1,2\}^{\mathbb{N}}$ and $|\{0,1\}^{\mathbb{N}}| \leq |\{0,1,2\}^{\mathbb{N}}|$. It remains to show the other direction. Consider the function $F : \{0,1,2\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ given as

$$\begin{aligned} F(f)(2n) &= \min\{f(n),1\}, \\ F(f)(2n+1) &= \min\{2-f(n),1\}. \end{aligned}$$

Assume that $f(n)=0$, $g(n)=1$, $h(n)=2$. Then $F(f)(2n)=0$ and $F(f)(2n+1)=1$, $F(g)(2n)=1$ and $F(g)(2n+1)=1$, $F(h)(2n)=1$ and $F(h)(2n+1)=0$. Thus if two functions are different at n, their image is different either at $2n$ or at $2n+1$. It follows that F is one-to-one and $|\{0,1,2\}^{\mathbb{N}}| \leq |\{0,1\}^{\mathbb{N}}|$. By Theorem 6.9, both sets have the same cardinality. ∎

Exercise 6.11. Show that if $|X| = |X \times \mathbb{N}|$ then $|\{0,1\}^X| = |\mathbb{N}^X|$.

Theorem 6.12 (Cantor). $|X| < |\mathcal{P}(X)|$.

Proof. The function $f(x) = \{x\}$ is a one-to-one function. Hence $|X| \leq |\mathcal{P}(X)|$.

To show that $|X| \neq |\mathcal{P}(X)|$, consider any function f from X to $\mathcal{P}(X)$. One has to show that X is not onto. For this, one defines the subset $A \subseteq X$ by

$$A = \{x \in X \mid x \notin f(x)\}$$

and shows that A is not in the range of f. This is done by verifying that $f(a) \neq A$ for any $a \in X$.

Actually this comes directly from the definition of A: for given a, the definition states that $a \in A$ iff $a \notin f(a)$. Thus A and $f(a)$ differ with respect to the membership of a and $f(a) \neq A$. This shows that f is not onto. ∎

7 Finite and Hereditarily Finite Sets

Example 6.3 considers the set of even numbers which is a proper subset of $\mathbb{N}$ but still has the same cardinality of $\mathbb{N}$. This is no longer possible for finite set. A proper subset of a finite set is strictly smaller. For example, for the set $\{0,1,2,3,4\}$ identified with the natural number 5, one has $|\{0,2,3,4\}| < |\{0,1,2,3,4\}|$.

Theorem 7.1. *Let $n \in \mathbb{N}$. If $f : n \to n$ is a one-to-one function, then f is onto. Furthermore, if $n \in \mathbb{N}$ and $u \subseteq n$ then $|u| = |m|$ for some $m \in \{0,1,2,\ldots,n\}$.*

Proof. This is proven by induction. When $n = 0$, the statement holds trivially. Let $f : S(n) \to S(n)$ be a one-to-one function.

Either there is no $m \in n$ with $f(m) = n$. Then $f[n] \subseteq n$ and $f[n] = n$ by induction hypothesis, that is, every $m < n$ is in the range of $f[n]$. Thus $f(n) \leq n$ but not $f(n) < n$. Therefore $f(n) = n$ and f is onto.

Or there is some $m \in n$ with $f(m) = n$. Now let $\tilde{f}(m) = f(n)$, $\tilde{f}(n) = n$ and $\tilde{f}(k) = f(k)$ for all $k \in S(n) - \{m, n\}$. $\tilde{f}$ and f have the same range since $\tilde{f}$ was obtained from f by interchanging the values at m and n. $\tilde{f}$ satisfies now the case "Either" above and is onto. Then also f is onto.

For the second statement, assume that u, n are given with $n \in \mathbb{N}$ and $u \subseteq n$. Now define by recursion a function g with $g(0) = 0$ and

$$g(S(k)) = \begin{cases} g(k) & \text{if } k \notin u; \\ S(g(k)) & \text{if } k \in u. \end{cases}$$

Assume now that $x, y \in u$ and x, y are different elements. Then either $x < y$ or $y < x$, say the first. Thus $S(x) \leq y$ and thus $g(y) \geq g(S(x)) > g(x)$. So g is one-to-one on u (although not outside u). Now let $m = g(n)$. If $\ell \in m$ then there is a minimal k with $g(S(k)) > \ell$. It follows that $g(S(k)) > g(k)$, $g(S(k)) = S(g(k))$, $k \in u$ and $g(k) = \ell$. Thus $m = g[u]$ and g is one-to-one on u, so $|u| = |m|$. ∎

Definition 7.2. A set X is finite iff there is some $n \in \mathbb{N}$ such that $|X| = |n|$. A set X is infinite iff X is not finite.

So the finite subsets of $\mathbb{N}$ are intuitively those which can be completely listed, for example $\{2, 3, 5\}$, $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and $\{1, 5, 25, 125, 625, 3125\}$. Infinite sets can never be written down completely as the examples $\{0, 1, 2, 3, \ldots\}$ of the natural numbers and $\{0, 2, 4, 6, 8, \ldots\}$ of the even numbers show. From a practical point of view, this is also true for some finite sets like $\{0, 1, 2, 3, \ldots, 1723907238947, 1723907238948, 1723907238949\}$.

Theorem 7.3. *Let X, Y, Z be finite sets and f be a function.*

1. *Every subset of X is finite.*
2. *If $X \subset Y$ then $|X| < |Y|$.*
3. *$f[X]$ is finite.*
4. *$X \cup Y$ is finite.*
5. *If every element of Z is finite then $\bigcup Z$ is finite.*
6. *$\mathcal{P}(X)$ is finite.*

Proof. 1. Let $A \subseteq X$. By assumption, there is a one-to-one and onto mapping f from X to some $n \in \mathbb{N}$. Let $u = f[A]$. By Theorem 7.1 there is a one-to-one mapping g from u onto some $m \in \mathbb{N}$. Now one defines for all $b \in A$ the mapping $b \mapsto g(f(b))$ which then is a one-to-one mapping from A onto m. Thus $|A| = |m|$ and A is finite.

2. Since Y is finite, there is a natural number n and a bijection $g : n \to Y$. Assume furthermore that $X \subseteq Y$ and that $|X| = |Y|$. The second property implies that there is a bijection $h : Y \to X$. Now let h' map every $m \in n$ to $g^{-1}(h(g(m)))$. h' is a bijection and maps n to a subset of n. By Theorem 7.1, h' is onto and $h'[n] = n$. Thus $X = h[g[n]] = g[g^{-1}[h[g[n]]]] = g[h'[n]] = g[n] = Y$. So $|X| = |Y|$ does not hold for a proper subset of Y.

3. There is a bijection $g : n \to X$ for some natural number n. Now let $h(k) = f(g(k))$ for all $k \in n$ and define $u = \{k \in n : h(k) \notin h[k]\}$ as the set of all numbers k for which $h(k)$ takes a value not taken by any $h(\ell)$ with $\ell < k$. Then $h[u] = h[n] = f[X]$ and h is one-to-one on u. By Theorem 7.1, $|u| = |m|$ for some $m \in \mathbb{N}$ and it follows that $|f[X]| = |h[u]| = |u| = |m|$. That is, $f[X]$ is a finite set.

4. As X, Y are finite, there are natural numbers n, m and mapping h_X, h_Y such that $X = h_X[n]$ and $Y = h_Y[m]$. Now define $g(k) = h_X(k)$ for $k = 0, 1, \ldots, n-1$ and $g(k) = h_Y(k-n)$ for $k = n, n+1, \ldots, n+m-1$. Then the finite set $\{0, 1, \ldots, n+m-1\}$ is the domain of g and $X \cup Y$ the range of g. It follows that $X \cup Y$ is finite.

5. As Z is finite, $Z = \{z_0, z_1, \ldots, z_{n-1}\}$ for some finite index set n. By assumption, $z_0, z_1, \ldots, z_{n-1}$ are all finite sets. Now define $u_0 = \emptyset$ and for every $k \in n$ inductively $u_{k+1} = u_k \cup z_k$. Clearly u_0 is finite and by induction over k it follows that u_{k+1} is finite as it is the union of two finite sets. Thus $u_n = \bigcup Z$ is finite.

6. It is verified by induction that $\mathcal{P}(X)$ has 2^n elements and is finite whenever X has n elements, that is, can bijectively mapped to n. This is obviously true for $X = \emptyset$ where $\mathcal{P}(X) = \{\emptyset\}$ contains $1 = 2^0$ elements. Assume that the inductive hypothesis for a set X having n elements is proven. Furthermore, $2^n \in \mathbb{N}$ and every set of 2^n elements is finite. Now let $Y = X \cup \{x\}$ have $S(n)$ elements, that is, $x \notin X$. For every subset Z of X there are two subsets $Z, Z \cup \{x\}$ of Y, thus the quantity of subsets of Y is two times as large as the quantity of subsets of X. It follows that $|\mathcal{P}(Y)| = 2 \cdot |\mathcal{P}(X)|$. Thus $|\mathcal{P}(Y)|$ has $2^{S(n)} = 2^{|Y|}$ elements and is finite again. ∎

Exercise 7.4. Let X be finite. Prove that the set of all functions from X to X is finite.

The set $\{\{\mathbb{N}\}\}$ is finite since it contains only the set $\{\mathbb{N}\}$. That set is finite again, but it contains the infinite set $\mathbb{N}$. That is, going along the iterated membership relation of $\{\{\mathbb{N}\}\}$ end up in an infinite set.

A set is called hereditarily finite if this does not happen. That is, if A is hereditarily finite, then not only A itself but also all of its elements, all of the elements of its elements and so on are finite. Most easily, this is defined by looking at the transitive closure.

Definition 7.5. A set A is called *hereditarily finite* iff every element of $\mathcal{TC}(A)$ is finite. Furthermore, let $V_\omega = \{x \in V \mid x \text{ is hereditarily finite}\}$.

Theorem 7.6. *A is hereditarily finite iff $\mathcal{TC}(A)$ is finite.*

Proof. If $\mathcal{TC}(A)$ is finite then A is also hereditarily finite. Assume now that $\mathcal{TC}(A)$ is infinite. Let $B = \{x \in \mathcal{TC}(A) \mid \mathcal{TC}(x) \text{ is finite}\}$. Clearly $B \subset \mathcal{TC}(A)$ as $A \notin B$. By the Axiom of Foundation there is $X \in \mathcal{TC}(A) - B$ such that no element of X is in $\mathcal{TC}(A) - B$, thus $X \subseteq B$. Now $\mathcal{TC}(Y)$ is finite for all $Y \in X$ and thus $\{\mathcal{TC}(Y) \mid Y \in X\}$ is a set of finite sets. Thus $\mathcal{TC}(X) = \{X\} \cup \bigcup\{\mathcal{TC}(Y) \mid Y \in X\}$ is infinite only if X is infinite, hence X is infinite and A not hereditarily finite. ∎

The natural numbers had been introduced by iterations of the operation $x \mapsto x \cup \{x\}$ starting with the empty set. This permits to write down every natural number explicitly, for example, 3 is $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$. The intuition behind this is to call a set hereditarily finite iff it can be written down explicitly on paper. In practice, one encounters of course the problem that already sets corresponding to moderately sized numbers like 275 require more symbols to be written down explicitly than there are atoms in the universe – it is estimated that there are between 10^{78} and 10^{82} atoms in the universe and between 10^{48} and 10^{52} atoms on the Earth.

Example 7.7. *The set $\mathbb{N}$ is infinite but all its elements are hereditarily finite. The set $\{1, \{0, 2\}, \{0, 3\}\}$ is hereditarily finite. The set $\{\{\mathbb{N}\}\}$ is finite but not hereditarily finite.*

Property 7.8. *A finite set x is hereditarily finite iff every $y \in x$ is hereditarily finite.*

Theorem 7.9. *The class V_ω is actually a set and coincides with the smallest set W satisfying*

1. *$\emptyset \in W$;*
2. *if $v \in W$ then $\{v\} \in W$;*
3. *if $v, w \in W$ then $v \cup w \in W$.*

Proof. Clearly $\emptyset$ is hereditarily finite. If v is hereditarily finite, so is $\{v\}$. If v, w are hereditarily finite then they are finite sets of hereditarily finite sets. It follows that $v \cup w$ is a finite set and that all its members are hereditarily finite since they are members of either v or w (or both). Thus V_ω is closed under the three operations given.

The next step is to construct a function f from $\mathbb{N}$ onto V_ω in order to show that it is a set. This function is first constructed as a function from $\mathbb{N}$ into V and the properties are shown later. The inductive construction goes as follows:

- $f(0) = \emptyset$;
- $f(2^n + m) = f(m) \cup \{f(n)\}$ for $n \in \mathbb{N}$ and $m \in \{0, 1, \ldots, 2^n - 1\}$.

One can show that $f[\mathbb{N}] \subseteq V_\omega$: Assume that k would be the least number with $f(k) \notin V_\omega$. Then $k > 0$ as $f(0) = \emptyset$ and $\emptyset \in V_\omega$. So there are $n, m \in \mathbb{N}$ with $m < 2^n \leq 2^n + m = k$. $f(n)$ and $f(m)$ are both in V_ω. Thus $\{f(n)\}$ and $\{f(n)\} \cup f(m)$ are also both in V_ω. It follows that $f(k) \in V_\omega$ in contradiction to the assumption. So one actually has $f[\mathbb{N}] \subseteq V_\omega$. Note that this argumentation also works with every set W satisfying 1., 2. and 3.; thus one has that $f[\mathbb{N}] \subseteq W$ for such a set.

Now consider any set $z \notin f[\mathbb{N}]$. It follows from the Axiom of Foundation that there is a set $x \in \mathcal{TC}(z) - f[\mathbb{N}]$ with $x \cap (\mathcal{TC}(z) - f[\mathbb{N}]) = \emptyset$. As $x \subset \mathcal{TC}(z)$ one can conclude that $x \subseteq f[\mathbb{N}]$. Now define inductively a function g such that $f(g(n)) = f[n] \cap x$ as follows: $g(0) = 0$ and

$$g(S(n)) = \begin{cases} g(n) + 2^n & \text{if } f(n) \in x; \\ g(n) & \text{if } f(n) \notin x. \end{cases}$$

One can show inductively that $0 \leq g(n) < 2^n$ for all $n \in \mathbb{N}$. Furthermore, $f(g(0)) = f[0] \cap x$ as $g(0) = 0$ and $f[0] = f[\emptyset] = \emptyset$. Inductively, if $f[n] \cap x = f(g(n))$ and $f(n) \notin x$ then $g(n+1) = g(n)$ and $f(g(n+1)) = f(g(n)) = f[n] \cap x = f[n+1] \cap x$; if $f[n] \cap x = f(g(n))$ and $f(n) \in x$ then $g(n+1) = g(n) + 2^n$ and $f(g(n+1)) = \{f(n)\} \cup f(g(n)) = \{f(n)\} \cup (f[n] \cap x) = f[n+1] \cap x$. It follows that $x = \bigcup f[g[\mathbb{N}]]$. Since $x \notin f[\mathbb{N}]$, $f(g(m)) \neq x$ for all $m \in \mathbb{N}$ and for every $m \in \mathbb{N}$ there is an $n \in \mathbb{N}$ with $n > m$ and $g(n) > g(m)$. In particular, $f(g(n))$ is a proper superset of $f(g(m))$. It follows that x is infinite and $x \notin V_\omega$. As $x \in \mathcal{TC}(z)$, $z \notin V_\omega$ as well. As a consequence, $V_\omega = f[\mathbb{N}]$ and V_ω is a set. Furthermore, $V_\omega \subseteq W$ for all sets W satisfying the conditions 1., 2., 3. as this had been proven above for $f[\mathbb{N}]$ in place of V_ω. ∎

Exercise 7.10. Prove that V_ω satisfies the following property: if $x \in V_\omega$ and $y \subseteq x$ or $y \in x$, then $y \in V_\omega$. Show that $\mathbb{N}$ does not satisfy this property, but that some proper infinite subset of V_ω does.

Exercise 7.11. Determine all $x_0 \in V$ which satisfy that there are no $x_1, x_2, x_3, x_4 \in V$ with $x_1 \in x_0, x_2 \in x_1, x_3 \in x_2, x_4 \in x_3$. The set $\{\{\emptyset\}\}$ is such an x_0, although $x_1 = \{\emptyset\}$ and $x_2 = \emptyset$ exist, x_3 and x_4 do not exist. The set $\{\{\emptyset, \{\{\emptyset\}\}\}\}$ does not qualify.

8 Countable Sets

Since $\mathbb{N}$ and $\mathcal{P}(\mathbb{N})$ do not have the same cardinality, there are several different cardinalities of infinite sets. The set $\mathbb{N}$ has the least infinite cardinality and this cardinality is called "countable". Countable sets are the only infinite sets where one — theoretically — can name every element by a name; for example, every natural number can be written down as a finite sequence of digits. The same applies to all other types of objects which can be coded with a finite alphabet. For example, the set of all possible novels is countable since one can write down each novel using an alphabet plus punctuation symbols and special characters. The same is true for the set of all computer programs. The word "countable" itself comes from the fact that one can count one by one all the things which can be written down explicitly. In the case of words (or strings) over the English alphabet, "a" is the first word, "b" the second, "z" the twenty-sixth, "aa" the twenty-seventh, "az" the fifty-second, "ba" the fifty-third, "zx" the sevenhundredth, "zz" the sevenhundred-second and "aaa" the sevenhundred-third word. So there is a surjective mapping from the natural numbers to all finite nonempty strings over the English alphabet which counts these strings. This fact motivates the following definition; in order to avoid to say "infinite and countable" all the time, countable sets are defined to be infinite.

Definition 8.1. A set X is at most countable iff either there is a surjective function $f : \mathbb{N} \to X$ or $X = \emptyset$. X is countable iff X is at most countable and infinite. X is uncountable iff $|X| > |\mathbb{N}|$.

Remark 8.2. *The sets of natural numbers and integers are examples of countable sets. Any finite set is at most countable but not countable.*

If g is a function and X at most countable, then $g[X]$ is at most countable: By definition, there is a surjective function $f : \mathbb{N} \to X$ which witnesses that X is at most countable. The concatenation $n \to g(f(n))$ then witnesses that $g[X]$ is at most countable, too.

Definition 8.1 reflects the property that one can enumerate the elements of a countable set, that is, that there is a surjective function from $\mathbb{N}$ to X. But the condition "$|X| \leq |\mathbb{N}|$" is defined the other way round: there has to be a one-to-one function from X to $\mathbb{N}$. The next result shows that both ways to define "at most countable" and "countable" are equivalent.

Proposition 8.3. *A set X is at most countable iff $|X| \leq |\mathbb{N}|$. An infinite set X is at most countable iff $|X| \leq |\mathbb{N}|$ iff $|X| = |\mathbb{N}|$.*

Proof. Let X be at most countable. Then there is a function $f : \mathbb{N} \to X$ which is surjective. For every $x \in X$ let $g(x)$ be the minimum of $f^{-1}(x) = \{y \in \mathbb{N} \mid f(y) = x\}$. The value $g(x)$ exists because f is surjective. Furthermore, if $g(x) = g(y)$ then $x = f(g(x))$, $y = f(g(y))$ and $x = y$, thus g is injective. So $|X| \leq |\mathbb{N}|$.

Assume now that X is infinite and that f is as above. For every number n there is a first natural number m such that $|S(n)| \leq |f[S(m)]|$; let $h(n)$ denote this number m for given n. Due to the definition of h one has that $|f[S(h(n))]| > |f[h(n)]|$. In particular, $f(h(n)) \in f[h(S(n))] - f[h(n)]$. It follows that the mapping $n \mapsto f(h(n))$ is one-to-one and witnesses that $|\mathbb{N}| \leq |X|$. Thus $|X| = |\mathbb{N}|$ by Theorem 6.9. ▮

Since the identity restricted to a subset Y of a set X is a one-to-one mapping from Y to X, one has the following corollary.

Corollary 8.4. *If X is countable and $Y \subseteq X$ then Y is at most countable; if Y is infinite then $|Y| = |X|$.*

Example 8.5. *The set $\mathbb{Q}$ of all rationals is countable.*

Proof. For every rational q there are unique numbers $n(q), m(q), k(q)$ such that the following conditions hold:

1. if $q = 0$ then $n(q) = 0$, $m(q) = 0$ and $k(q) = 5$;
2. if $q > 0$ then $q = \frac{n(q)}{m(q)}$, $k(q) = 7$ and $n(q), m(q)$ do not have a common prime factor;
3. if $q < 0$ then $q = -\frac{n(q)}{m(q)}$, $k(q) = 11$ and $n(q), m(q)$ do not have a common prime factor.

Let $f(q) = 2^{n(q)} \cdot 3^{m(q)} \cdot k(q)$. It is easy to see that f is a one-to-one mapping from $\mathbb{Q}$ to $\mathbb{N}$. Thus $|\mathbb{Q}| \leq |\mathbb{N}|$. Since $\mathbb{N} \subseteq \mathbb{Q}$, the cardinality of both sets is the same and $\mathbb{Q}$ is countable. ▮

Proposition 8.6. *If X and Y are at most countable so is $X \times Y$.*

Proof. Since X, Y are at most countable, there are one-to-one mappings $f : X \to \mathbb{N}$ and $g : Y \to \mathbb{N}$. Now let $h(x, y) = 2^{f(x)} \cdot 3^{g(y)}$. The function h is a one-to-one mapping from $X \times Y$ to $\mathbb{N}$. Thus $|X \times Y| \leq |\mathbb{N}|$. ▮

Remark 8.7. If X and Y are infinite, then one can take h even such that h is bijective. Cantor constructed an explicit bijection between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$. He used the function p mapping m, n to $\frac{1}{2} \cdot (m+n) \cdot (m+n+1) + m$; so $p(0,0) = 0$, $p(0,1) = 1$, $p(1,0) = 2$, $p(0,2) = 3$, $p(1,1) = 4$, $p(2,0) = 5$, $p(0,3) = 6$ and so on. The verification that p really is a bijection is left to the reader.

Proposition 8.6 can be adapted to show that $X \times Y$ and $X \cup Y$ are countable whenever X and Y are countable sets.

For the next result, consider just the function f build in Theorem 7.9.

Property 8.8. *The set V_ω of all hereditarily finite sets is countable.*

Exercise 8.9. Let $D = \{f : \mathbb{N} \to \mathbb{N} \mid \forall n\, (f(S(n)) \leq f(n))\}$ be the set of all decreasing functions. Show that D is countable.

Theorem 8.10. *Let A be nonempty and at most countable and A^* denote the set of finite sequences of members in A. Then A^* is countable. Furthermore, the set of all finite subsets of A is at most countable.*

Proof. Let $g : \mathbb{N} \to A$ be a surjective function and consider the following set B: $x \in B \Leftrightarrow x$ is of the form $\{(0, b_0), (1, b_1), (2, b_2), \ldots, (n-1, b_{n-1})\}$ for some natural numbers n and $b_0, b_1, \ldots, b_{n-1}$ where the latter are used to code elements of A. As ordered pairs are just coded sets, one can easily set that $B \subseteq V_\omega$. Now mapping the empty set to the empty sequence and $\{(0, b_0), (1, b_1), (2, b_2), \ldots, (n-1, b_{n-1})\}$ to $g(b_0)\, g(b_1)\, g(b_2)\, \ldots\, g(b_{n-1})$ shows that A^* is at most countable. As A^* contains for each $n \in \mathbb{N}$ a sequence of length n, A^* is infinite and countable.

Let E be the set of all finite subsets of $\mathbb{N}$. The set E is at most countable as V_ω is countable. Furthermore, mapping each $x \in E$ to $g[x]$ produces a surjective mapping from E onto all finite subsets of A and thus the set of all finite subsets of A is at most countable. ∎

Exercise 8.11. Let A contain four elements, the symbol $\emptyset$, the bracket $\{$, the bracket $\}$ and the comma in this ordering; that is, the symbol $\emptyset$ comes first and the comma comes last. Let $<_{ll}$ be the length-lexicographic ordering of the set A^* of all strings over A: if v is shorter than w then $v <_{ll} w$; if v, w have the same length then $v <_{ll} w \Leftrightarrow v <_{lex} w$. Now let $f : V_\omega \to A^*$ map every set x in V_ω to first expression describing x; for example,

- $f(\{1,2\}) =$"$\{\{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$";
- $f(\{0,3\}) =$"$\{\emptyset, \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}\}$".

As $\emptyset <_{ll} \{\}$, the symbol "$\{\}$" is never used to describe the empty set; this convention is also applied in this text. Check which of the following facts are true:

1. the length of $f(x)$ is odd for every $x \in V_\omega$;
2. the length of $f(\mathcal{P}(x))$ is the product of the length of $f(x)$ and the cardinality of $\mathcal{P}(x)$ plus 1;
3. if $f(x) = \{y\}$ then $f(S(x)) = \{y, \{y\}\}$;
4. $f(2) = \{\emptyset, \{\emptyset\}\}$.

Furthermore, find a formula giving the length of $f(n)$ for every n and determine which of the following numbers is the length of $f(10)$: 42, 100, 1000, 1001, 1022, 1023, 1024, 2047, 4096, 256^2, $10^{10} - 1$, 2^{256}, 1023^{1023}.

If the length of $f(x)$ is n and $f(y)$ is m, what is the length of $f((x, y))$ for the ordered pair (x, y)?

Exercise 8.12. Let $\mathbb{A}$ be the set of algebraic real numbers, that is, the set of all $r \in \mathbb{R}$ for which there are $n \in \mathbb{N}$ and $z_0, z_1, \ldots, z_n \in \mathbb{Z}$ such that $z_n \neq 0$ and $z_0 + z_1 r + z_2 r^2 + \ldots + z_n r^n = 0$. Note that such a polynomial of degree n can have up to n places r which are mapped to 0. Show that $\mathbb{A}$ is countable by giving a one-to-one mapping from $\mathbb{A}$ into $\mathbb{N}$.

Example 8.13. *The set F of all continuous functions $f : \mathbb{Q} \to \mathbb{Q}$ is uncountable.*

Proof. Given $A \subseteq \mathbb{Z}$, one can map A to the continuous function $f \in F$ given as

$$f(q) = \begin{cases} 0 & \text{if } z, z+1 \notin A; \\ q - z & \text{if } z \notin A, z+1 \in A; \\ z + 1 - q & \text{if } z \in A, z+1 \notin A; \\ 1 & \text{if } z, z+1 \in A; \end{cases}$$

where z is the unique integers with $z \leq q < z + 1$. Since this mapping is one-to-one, $|\mathcal{P}(\mathbb{Z})| \leq |F|$. ∎

Exercise 8.14. Call a set A *hereditarily at most countable* iff for every $B \in \mathcal{TC}(A)$ it holds that B is at most countable. For example, $\mathbb{N}$ and V_ω are hereditarily at most countable. Now assume that X, Y are hereditarily at most countable. Show that the following sets are hereditarily at most countable as well: $X \cup Y$, $X \cap Y$, $X - Y$ and $\{X\}$. Furthermore, show that whenever $\mathcal{TC}(Z)$ is at most countable then Z is hereditarily at most countable.

9 Graphs and Orderings

Consider the following list which relates words with the same meaning to each other, say English words to their Spanisch counterparts: (dog, el perro), (cat, el gato), (cat, la gata), (boy, el niño), (girl, la niña), (cow, la vaca), (English, el Inglés), (Spain, España), (blue, azul). Since the relation is not unique, "cat" can be "el gato" (if male) or "la gata" (if female), one cannot represent the structure as a function. The most common more general notion studied by mathematicians is that of directed graphs. It has a set of vertices, in this example the set of all words in either English or Spanish language. Furthermore, it has a set of pairs of vertices, the edges, in this example the edges consist of one English and one Spanish word whose meaning is related to each other in the way that there is an object or concept which can be denoted by the first word in English and the second word in Spanish: there exists a male cat and thus (cat, el gato) is in the set of edges, furthermore there exists a female cat and thus (cat, la gata) is in the set of edges. These pairs are ordered, the English word is always the first one. For the reader's convenience, the formal introduction of graphs is repeated from Definition 3.5.

Definition 9.1. A (*directed*) *graph* is pair (G, E) such that G is a set and $E \subseteq G \times G$. The members of G are called *vertices* and the members of E are called *edges*.

Example 9.2. *Every function $f : X \to Y$ can be represented as the graph*

$$(X \cup Y, \{(x, y) \in (X \cup Y) \times (X \cup Y) \mid x \in X \wedge y \in Y \wedge y = f(x)\}).$$

One can also consider graphs with a class as their domain: $(V, \in)$ is a graph and $(V_\omega, \in)$ is a graph where in the latter case "$\in$" is restricted to the domain V_ω. Also $(V, \{(x, S(x)) \mid x \in V\})$ is a graph.

Exercise 9.3. A graph (G, E) is called *bipartite* if there are two subsets X, Y of G such that $X \cap Y = \emptyset$ and every pair $(x, y) \in E$ is actually in $X \times Y \cup Y \times X$. An English-Russian dictionary is a bipartite graph because one can take X to be the words written in the Latin alphabet and Y to be the words written in the Cyrillic alphabet. An English-Spanish dictionary is not bipartite, for example place names like "Los Angeles" appear in the same spelling in both languages. By the way, the mentioned name is of Spanish origin and has the English translation "the angels". Which of the following graphs are bipartite? The set of vertices is $\mathbb{N}$ and the set E_n of edges is specified below, note that $E_n \subseteq \mathbb{N} \times \mathbb{N}$.

1. $(x, y) \in E_1 \Leftrightarrow x = y$,

2. $(x, y) \in E_2 \Leftrightarrow x < y$,

3. $(x, y) \in E_3 \Leftrightarrow 12 < x + y < 18$,

4. $(x, y) \in E_4 \Leftrightarrow x > 4 \wedge (y = x^2 \vee y = x^4)$,

5. $(x, y) \in E_5 \Leftrightarrow \exists z > 0 \, (x = 2^z \wedge y = 3^z)$,

6. $(x, y) \in E_6 \Leftrightarrow \exists z > 0 \, (x \in \{2^z, 3^z\} \wedge y \in \{5^z, 7^z\})$,

7. $(x, y) \in E_7 \Leftrightarrow y = S(x) \wedge x$ is even.

If one considers the graph $(V, \subset)$ instead of $(V, \in)$, one has additional properties not present at $(V, \in)$. The main additional property is transitivity. On the other hand, there are still incomparable elements of V; for example $\{\{\{0\}\}\}$ and $\{0, 1\}$. This is captured by the definition of a partially ordered set.

Definition 9.4. A set G is partially ordered by a relation $<$ iff this relation is antisymmetric and transitive. These two properties are defined as follows:

1. the relation $<$ is antisymmetric iff there are no $x, y \in G$ with $x < y$ and $y < x$;

2. the relation $<$ is transitive iff $x < z$ for all $x, y, z \in G$ with $x < y$ and $y < z$.

Note that a transitive relation $<$ is antisymmetric iff it is antireflexive, that is, iff there is no $x \in G$ with $x < x$.

Convention 9.5. If G is partially ordered by $<$ then $<$ is called a *partial ordering* on G and $(G, <)$ is called a *partially ordered set*. One usually writes $a < b$ if the ordering is denoted by a symbol of the type $<$ and $(a, b) \in R$ if the ordering is denoted by a letter like R.

Furthermore, the notation $a \leq b$ stands for $a < b \vee a = b$. Similarly $A \subseteq B$ stands for $A \subset B \vee A = B$.

If $a \leq b$, then one says that a is less than or equal to b. If $a < b$, then one says that a is less than b or a is smaller than b.

Example 9.6. *Assume that $G = \{a, b, c, d\}$ and $a < b$, $b < c$, $a < c$ and $d < c$. Then G is a partially ordered set where the elements a, d are incomparable. Furthermore, $c > b$ and $b < c$ mean the same. The relation $a < c$ is needed since the ordering $<$ would otherwise not be transitive. Graphically, the ordering looks like this:*

```
      c
     / \
    b   d
   /
  a
```

Exercise 9.7. *Let* $A = \mathbb{N} - \{0,1\} = \{2,3,4,\ldots\}$ *and let* $<_{div}$ *be given by* $x <_{div} y \Leftrightarrow \exists z \in A\,(x \cdot z = y)$. *That is,* $x <_{div} y$ *iff* x *is a proper divisor of* y, *so* $2 <_{div} 8$ *but* $2 \not<_{div} 2$ *and* $2 \not<_{div} 5$. *Prove that* $(A, <_{div})$ *is a partially ordered set.*

Example 9.8. *The subset-relation* $\subset$ *is a partial ordering of* V. *Furthermore, the relation* R *on* V *defined as* $(x,y) \in R \Leftrightarrow |x| < |y|$ *is a partial ordering of* V. *The same holds for* R' *given by* $(x,y) \in R' \Leftrightarrow |\mathcal{P}(x)| \leq |y|$. *These two partial orderings are different since* $(\{a,b\},\{a,b,c\})$ *is in* R *but not in* R'.

Due to coding, the element-relation $\in$ *is an ordering of* $\mathbb{N}$ *where it coincides with the natural less-than relation.*

If $|G| \geq 2$, *then* $(G, \neq)$ *is not a partially ordered set. Let* a, b *be distinct elements of* G. *Then* $a \neq b$ *and* $b \neq a$, *but* $a = a$, *thus the inequality is not transitive.*

Example 9.9. *The following relation* $\sqsubset$ *is a partial ordering of the set* $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$: $(x,y,z) \sqsubset (x',y',z') \Leftrightarrow x \leq x' \wedge y \leq y' \wedge z \leq z' \wedge x+y+z < x'+y'+z'$.

Proof. If $(x,y,z) \sqsubset (x',y',z')$ then $x+y+z < x'+y'+z'$ and it cannot be that $x'+y'+z' < x+y+z$. Thus $\sqsubset$ is antisymmetric.

To see that $\sqsubset$ is transitive, consider any three triples $(x,y,z), (x',y',z'), (x'',y'',z'')$ such that $(x,y,z) \sqsubset (x',y',z')$ and $(x',y',z') \sqsubset (x'',y'',z'')$. By the transitivity of $\leq$ and $<$ one has that $x \leq x''$, $y \leq y''$, $z \leq z''$ and $x+y+z < x''+y''+z''$. Thus $(x,y,z) \sqsubset (x'',y'',z'')$ and $\sqsubset$ is transitive.

Note that the triples $(0,0,1)$ and $(0,1,0)$ are incomparable with respect to $\sqsubset$. By using the term "partial ordering" this is explicitly permitted although it is not mandatory. ∎

Exercise 9.10. Prove that the following relations are partial orderings on $\mathbb{N}^{\mathbb{N}}$:

- $f \sqsubset_1 g \Leftrightarrow \exists n\, \forall m > n\,(f(m) < g(m))$;
- $f \sqsubset_2 g \Leftrightarrow \forall n\,(f(n) \leq g(n)) \wedge \exists m\,(f(m) < g(m))$;
- $f \sqsubset_3 g \Leftrightarrow \forall n\,(f(n) \leq g(n)) \wedge \exists n\,(f(n) < g(n)) \wedge \exists n\, \forall m > n\,(f(m) = g(m))$;
- $f \sqsubset_4 g \Leftrightarrow f(0) < g(0)$.

Determine for every ordering a pair of incomparable elements f, g such that neither $f \sqsubset_m g$ nor $g \sqsubset_m f$ nor $f = g$. For which of these orderings is it possible to choose the f of this pair (f,g) of examples such that $f(n) = 0$ for all n?

Remark 9.11 (Preordering). A relation $\sqsubseteq$ on a set G is called preordering iff it is transitive; it can also be reflexive, but this is not required here although some authors require it in other books. The ordering $<$ on G defined as

$$x < y \Leftrightarrow x \sqsubseteq y \text{ and not } y \sqsubseteq x$$

is a partial ordering generated from $\sqsubseteq$. Note that the symbol $\leq$ derived from $<$ can be more restrictive than $\sqsubseteq$: For example, the preordering $\sqsubseteq$ on V defined by

$$x \sqsubseteq y \Leftrightarrow |x| \leq |y|$$

defines a partial ordering $<$ on V such that

$$x < y \Leftrightarrow |x| < |y|$$

but the derived relation $\leq$ is then

$$x \leq y \Leftrightarrow (x = y \vee |x| < |y|)$$

and it would be that $\mathbb{N} - \{0\} \sqsubseteq \mathbb{N}$ is true but $\mathbb{N} - \{0\} \leq \mathbb{N}$ is false with respect to the just defined relations $\sqsubseteq, <, \leq$.

Proposition 9.12. *Given a graph (G, E), one can define a preordering $\leq_E$ by $x \leq_E y$ iff there is a natural number n and a function $f : S(n) \to G$ with $f(0) = x$, $f(n) = y$ and $(f(m), f(S(m))) \in E$ for all $m \in n$.*

Furthermore, if (G, E) is cycle-free, that is, iff there is no $n \in \mathbb{N} - \{0\}$ and no function $f : S(n) \to G$ with $f(0) = f(n)$ and $\forall m \in n\,((f(m), f(S(m))) \in E)$ then the partial ordering $<_E$ generated from $\leq_E$ satisfies $x \leq_E y \Leftrightarrow x <_E y \vee x = y$ as desired.

Proof. The preordering $\leq_E$ can formally be defined as follows: For all $x, y \in G$, $x \leq_E y$ iff $\exists n \in \mathbb{N} \exists f \in G^{S(n)}\,(f(0) = x \wedge f(n) = y \wedge \forall m \in n\,((f(m), f(S(m))) \in E))$.

The transitivity is easy to see. Assume that f with domain $S(n)$ witnesses $x \leq_E y$ and g with domain $S(m)$ witnesses $y \leq_E z$. Then define $h : S(n+m) \to G$ with $h(k) = f(k)$ for $k \in S(n)$ and $h(n+k) = g(k)$ for $k \in S(m)$. As $f(n) = g(0) = y$, this definition is not contradictory. Furthermore, for $k \in \{0, 1, \ldots, n-1\}$ it holds that $(h(k), h(S(k))) = (f(k), f(S(k))) \in E$ and for $k \in \{n, n+1, \ldots, n+m-1\}$ it holds that $(h(k), h(S(k))) = (g(k-n), g(S(k-n))) \in E$. As $h(0) = x$ and $h(n+m) = z$, one has $x \leq_E z$.

Assume that $x \neq y$ and $x \leq_E y$. If (G, E) is cycle-free then it cannot be that $y \leq_E x$ by the transitivity shown before, thus $x <_E y$ in the way as $<_E$ is derived from the preordering $\leq_E$. Furthermore, $x \leq_E x$ as one could take $n = 0$ and consider the function f with domain 0 and $f(0) = x$. Thus one has that $x \leq_E y \Leftrightarrow x <_E y \vee x = y$. ▮

Remark 9.13. *Note that in a cycle-free graph also $(x,x) \notin E$ for all x. This property gives then the additional property that $x <_E y$ iff there is an $n \in \mathbb{N}-\{0\}$ and a function $f : S(n) \to G$ with $f(0) = x$, $f(n) = y$ and $\forall m \in n\,((f(m), f(S(m))) \in E)$ so that $<_E$ can be directly defined from (G,E). Therefore the case of cycle-free graphs is the most desirable one.*

If a graph is not cycle-free but has some cycle of n different nodes $x_0, x_1, \ldots, x_{n-1} \in G$ with $(x_m, x_{m+1}) \in E$ for all $m < n-1$ and $(x_{n-1}, x_0) \in E$ as well, then one has $x_i \leq_E x_j$ but not $x_i <_E x_j$ for all $i, j \in n$.

Well-founded graphs are cycle-free but not vice versa as the graph $(\mathbb{Z}, E)$ with $E = \{(z, z+1) \mid z \in \mathbb{Z}\}$ is cycle-free but not well-founded. The ordering $<_E$ is then the natural ordering on all integers.

One can define $<_\in$ on the whole universe V and obtains that $\mathcal{TC}(x) = \{y \in V \mid y = x \vee y <_\in x\}$ and $y <_\in x \Leftrightarrow y \in \mathcal{TC}(x) - \{x\}$. So these two notions stand in a very close correspondence.

Exercise 9.14. Let $A = \mathbb{N} - \{0,1\}$ and $<_{div}$ given by $x <_{div} y \Leftrightarrow \exists z \in A\,(x \cdot z = y)$ as in Exercise 9.7. Define a relation E on $A \times A$ by putting (x,y) into E iff there is a prime number z with $x \cdot z = y$. So $(2,4) \in E$, $(2,6) \in E$, $(2,10) \in E$ but $(2,7) \notin E$, $(2,8) \notin E$ and $(2,20) \notin E$. Show that $(A, <_E)$ and $(A, <_{div})$ are identical partially ordered sets.

10 Linear Ordering

A linear ordering is a partial ordering with the additional property that any two different elements are comparable.

Definition 10.1. Let A be a set and $R \subseteq A \times A$. One says that R is a *linear ordering* of A iff R satisfies the following properties for all $a, b, c \in A$:

antisymmetric: if $(a,b) \in R$ then $(b,a) \notin R$;

transitive: if $(a,b) \in R$ and $(b,c) \in R$ then $(a,c) \in R$;

comparable: either $a = b$ or $(a,b) \in R$ or $(b,a) \in R$.

The pair (A,R) is called a linearly ordered set iff R is a linear ordering of A.

Example 10.2. *Recall that the natural ordering $<$ on $\mathbb{N}$ coincides with the $\in$-relation due to the way the natural numbers are coded into V; that is, for all $m, n \in \mathbb{N}$,*

$$n = \{0, 1, \ldots, n-1\} \text{ and } m < n \Leftrightarrow m \in n.$$

It is easy to see that $(\mathbb{N}, <)$ *is a linearly ordered set, that is, that* $<$ *is antisymmetric, transitive and comparable on* $\mathbb{N}$.

Example 10.3. *The two orderings*

$$(0,0) <_1 (0,1) <_1 (0,2) <_1 (0,3) <_1 \ \ldots\ <_1 (1,3) <_1 (1,2) <_1 (1,1) <_1 (1,0) \text{ and}$$
$$(0,0) <_2 (0,1) <_2 (0,2) <_2 (0,3) <_2 \ \ldots\ <_2 (1,0) <_2 (1,1) <_2 (1,2) <_2 (1,3) <_2 \ldots$$

on the set $\{0,1\} \times \mathbb{N}$ *are linear.*

Example 10.4. *Let* $A = \mathbb{N} \times \mathbb{N}$. *Define* $<_{lex}$ *on* A *by*

$$(m,n) <_{lex} (i,j) \iff (m < i) \vee (m = i \wedge n < j)$$

for all $m, n, i, j \in \mathbb{N}$. *Then* $(\mathbb{N} \times \mathbb{N}, <_{lex})$ *is a linearly ordered set. This ordering is called the lexicographic ordering of* $\mathbb{N} \times \mathbb{N}$. *Note that* $<_2$ *from the previous example is the restriction of* $<_{lex}$ *to the domain* $\{0,1\} \times \mathbb{N}$.

Example 10.5. *Let* $(A, <)$ *be a linearly ordered set. Recall that* A^* *contains every* A*-valued function whose domain can be represented by a natural number. For every* $f, g \in A^*$ *with domain* m, n, *respectively, define*

$$f <_{KB} g \Leftrightarrow \exists k \in m \cap S(n)\, ((f\restriction k = g\restriction k) \wedge (k = n \vee (k < n \wedge f(k) < g(k)))).$$

This is an ordering which is called the Kleene-Brouwer-ordering.

Proof. It is shown that $(A, <_{KB})$ satisfies the properties necessary to be a linearly ordered set.

Antireflexiveness. Assume that $f = g$ and $k \in m$. Then $k < n$ and $f(k) = g(k)$, thus it cannot be that $f <_{KB} g$.

Transitivity. Assume that $f <_{KB} g$ and $g <_{KB} h$. Let m, n, o be the corresponding domains and k be the minimum of the parameters of the same name in order to establish that $f <_{KB} g$ and $g <_{KB} h$. Note that $f\restriction k = g\restriction k = h\restriction k$. The fact that $f <_{KB} g$ gives that $k < m$. Similarly $g <_{KB} h$ gives that $k < n$. If $k = o$ then $f <_{KB} h$ and transitivity holds. If $k < o$ then one has $f(k) \leq g(k) \leq h(k)$ and one of these must be proper since k is the parameter for either $f <_{KB} g$ or $g <_{KB} h$. Thus $f(k) < h(k)$ and $f <_{KB} h$. Again transitivity holds.

Comparability. Assume that $f \neq g$. Then there is a minimal number k such that either $k \notin n$ or $k \notin m$ or $f(k) \neq g(k)$. Since $f\restriction k = g\restriction k$ but $f \neq g$ it cannot happen that $k \notin n \cup m$, that is, $k = n = m$. So one has exactly one of the following four cases.

1. $k \in n \cap m$ and $f(k) < g(k)$. Then $f <_{KB} g$.

2. $k \in n \cap m$ and $g(k) < f(k)$. Then $g <_{KB} f$.

3. $k = n$ and $k < m$. Then $f <_{KB} g$.

4. $k = m$ and $k < n$. Then $g <_{KB} f$.

So it follows from this case distinction that $<_{KB}$ is indeed a linear ordering. ∎

Exercise 10.6. Let $(A, <)$ be a linearly ordered set and $B = A^{\mathbb{N}}$. Define

$$f <_{lex} g \Leftrightarrow \exists k \in \mathbb{N}\, (f \restriction k = g \restriction k \wedge f(k) < g(k)).$$

Furthermore, let $C = A^*$. The lexicographic ordering on A^* differs from the Kleene-Brouwer ordering in the sense that it is reverted iff f, g coincide on the intersection of their domains m, n. That is,

$$f <_{lex} g \Leftrightarrow \exists k \in S(m) \cap n\, ((f \restriction k = g \restriction k) \wedge (k = m \vee (k < m \wedge f(k) < g(k)))).$$

Show that $(B, <_{lex})$ and $(C, <_{lex})$ are linearly ordered sets. Assuming that $A = \{0, 1, 2, \ldots, 9\}$ with the usual ordering, put the following elements of C into lexicographic order: 120, 88, 512, 500, 5, 121, 900, 0, 76543210, 15, 7, 007, 00.

Example 10.7. *Let $\mathbb{R}$ be the set of reals and $<$ be the natural ordering of the reals. Then $(\mathbb{R}, <)$ is a linearly ordered set. This ordering can be inherited to the subsets $\mathbb{Z}$ of integers and $\mathbb{Q}$ of rationals, thus $(\mathbb{Z}, <)$ and $(\mathbb{Q}, <)$ are linearly ordered sets as well.*

Definition 10.8. Let $(A, <)$ be a linearly ordered set. Recall that $a \leq b$ stands for $a = b \vee a < b$. Let B be a nonempty subset of A, $a \in A$ and $b \in B$.

1. a is a lower bound of B iff $a \leq c$ for all $c \in B$.

2. a is an upper bound of B iff $c \leq a$ for all $c \in B$.

3. b is the least element of B (with respect to $<$) iff b is a lower bound for B.

4. b is the greatest element of B (with respect to $<$) iff b is an upper bound of B.

5. a is the infimum of B iff a is the greatest lower bound of B.

6. a is the supremum of B iff a is the least upper bound of B.

7. B is bounded from above in $(A, <)$ iff B has an upper bound in A.

8. B is bounded from below in $(A, <)$ iff B has a lower bound in A.

9. B is bounded iff B is bounded both from above and from below.

Example 10.9. *Consider the following subsets of* $(\mathbb{R}, <)$*:*

$$\begin{array}{rcl} X & = & \{x \in \mathbb{R} \mid -2 \leq x < 5\}, \\ Y & = & \{y \in \mathbb{R} \mid y \geq 5\}. \end{array}$$

The set X is bounded in $(\mathbb{R}, <)$, for example -1024 is a lower and 1024 is an upper bound. Furthermore, $0 \in X$ and thus X is not empty. Therefore X has an infimum and a supremum. The infimum of X is -2 and the supremum of X is 5. Since $-2 \in X$, X has a least element, namely -2. But $5 \notin X$. Thus X does not have a greatest element.

The set Y has the infimum 5 which is also a lower bound. But Y has no upper bound in $(\mathbb{R}, <)$. Thus Y is unbounded and has no supremum.

Exercise 10.10. Determine which of the following subsets of the real numbers $\mathbb{R}$ have a lower and upper bound. If so, determine the infimum and supremum and check whether these are even the least and greatest element of these sets.

1. $A = \{a \in \mathbb{R} \mid \exists b \in \mathbb{R}\,(a^2 + b^2 = 1)\}$;

2. $B = \{b \in \mathbb{R} \mid b^3 - 4 \cdot b < 0\}$;

3. $C = \{c \in \mathbb{R} \mid \sin(c) > 0\}$;

4. $D = \{d \in \mathbb{R} \mid d^2 < \pi^3\}$;

5. $E = \{e \in \mathbb{R} \mid \sin(\frac{\pi}{2} \cdot e) = \frac{e}{101}\}$.

Definition 10.11. Two partial ordered sets $(A, <_1)$ and $(B, <_2)$ are *isomorphic*, denoted by $(A, <_1) \cong (B, <_2)$, iff there is a bijection $f : A \to B$ such that for all $a, b \in A$, $a <_1 b$ if and only if $f(a) <_2 f(b)$. Such functions are called *isomorphisms*.

For partial ordered sets $(A, <_1)$ and $(B, <_2)$, a function $f : A \to B$ is *order preserving* iff the implication $a <_1 b \Rightarrow f(a) <_2 f(b)$ is true for all $a, b \in A$.

Assume that $(A, <)$ is a linear ordered set and $f : A \to B$ is order-preserving. Then f is an isomorphism iff f is surjective. There is an order-preserving mapping from $\mathbb{Z}$ into $\mathbb{Q}$ but no isomorphism. The next result is of similar nature.

Example 10.12. *There is no order-preserving function from $\mathbb{Z}$ into $\mathbb{N}$.*

Proof. Assume by way of contradiction that there is an order-preserving function $f : \mathbb{Z} \to \mathbb{N}$. Then $f(0) = n$ for some n. It follows that $f(-1) < n$ and thus $f(-1) \le n-1$, $f(-2) < n-1$ and thus $f(-2) \le n-2$. By induction one can show that $f(-m) \le n-m$ and $f(-n) \le 0$. But then $f(-n-1) < 0$ what is impossible. So f cannot exist. ▮

Exercise 10.13. Consider the ordering $\sqsubset$ given by

$$\begin{array}{rcl} (m,n) \sqsubset (i,j) & \Leftrightarrow & (m < i) \\ & \vee & (m = i \wedge m \text{ is even} \wedge n < j) \\ & \vee & (m = i \wedge m \text{ is odd} \wedge n > j) \end{array}$$

on $A = \{0,1,2,3,4,5\} \times \mathbb{N}$. Construct an order-preserving mapping from $(\mathbb{Z}, <)$ into $(A, \sqsubset)$ where $<$ is the natural ordering of $\mathbb{Z}$.

The set $(\mathbb{Z}, <)$ there are nontrivial isomorphisms onto itself, that is, isomorphism different from the identity. For example, $z \mapsto z + 8$. Does $(A, \sqsubset)$ also have nontrivial isomorphisms onto itself? If so, is there any element which is always mapped to itself?

Proposition 10.14. *If $(A, <)$ is a finite linearly ordered set and $A \neq \emptyset$ then A has a greatest and a least element with respect to $<$.*

Proof. This is proven by induction. The proposition holds for orderings having one element since this unique element is the least and greatest element with respect to the given ordering at the same time.

Assume now that $n \ge 1$ and the proposition holds for all nonempty finite linearly ordered sets of cardinality up to n. Let $(A, <)$ be a linearly ordered set of cardinality $S(n)$. Let $a \in A$ and $B = A - \{a\}$. Then $(B, <)$ is a linearly ordered set of cardinality n. By induction hypothesis, B has a least element b_1 and a greatest element b_2. There are three cases:

1. $b_2 < a$. Then b_1 is the least and a the greatest element of A.

2. $a < b_1$. Then a is the least and b_2 the greatest element of A.

3. $b_1 < a < b_2$. Then b_1 is the least and b_2 the greatest element of A.

These three cases cover every possibility since $b_1 \le b_2$. Thus it follows from case-distinction that A has a least and a greatest element. This completes the inductive step. ▮

Theorem 10.15. *If $(A, <)$ is a finite linearly ordered set and $n = |A|$, then $(A, <) \cong (n, \in)$.*

Proof. The theorem is proven by induction on n. If $A = \emptyset$ then $(A, <) \cong (0, \in)$ since both are empty sets and the ordering of an empty set is irrelevant.

Assume that the theorem hold for all linearly ordered sets of size n. Let $(A, <)$ be a linearly ordered set of size $S(n)$. Let $a \in A$ be the greatest element of A, given by Proposition 10.14. Let $B = A - \{a\}$. Then $(B, <) \cong (n, \in)$ by induction hypothesis. Let $g : B \to n$ be the isomorphism. Define $f : A \to S(n)$ by $f = g \cup \{(a, n)\}$. Then f is an isomorphism since a is the greatest element of A and n is the greatest element of $S(n)$ with respect to $\in$ and g is an isomorphism. This finishes the proof of the inductive step and the whole theorem. ▮

The theorem says that, up to isomorphism, the finite linearly ordered sets are sets of the form $n = \{m \in \mathbb{N} \mid m < n\}$ with the natural ordering. The next result is that every countable linear ordering is isomorphic to a subset of $\mathbb{Q}$ with the standard ordering. So, the linear ordering of the set of all rational numbers is really a universal linear ordering of countable sets.

Definition 10.16. A linearly ordered set $(A, <)$ is *dense* iff A has at least two elements and for any pair $a, b \in A$ with $a < b$ there is $c \in A$ such that $a < c < b$. A subset $B \subseteq A$ is *dense in a linearly ordered set* $(A, <)$ iff for every $a, b \in A$ with $a < b$ there is a $c \in B$ such that $a < c < b$. A linearly ordered set $(A, <)$ *has no end points* iff for all $a \in A$ there are $b, c \in A$ such that $b < a < c$.

Example 10.17. $(\mathbb{Q}, <)$ *and* $(\mathbb{R}, <)$ *are dense linearly ordered sets without end points.* $\mathbb{Q}$ *is also dense in* $(\mathbb{R}, <)$. $(\mathbb{Z}, <)$ *is not dense.* $(\{0, 1, 2, 3\}, <)$ *has end points* 0 *and* 3. *The set* $\mathbb{D} = \{m \cdot 2^{-n} \mid n \in \mathbb{N} \wedge m \in \{0, 1, 2, \ldots, 2^n\}\}$ *of all dyadic numbers between* 0 *and* 1 *is dense and has end points* 0 *and* 1. $\mathbb{D}$ *is a subset of* $\mathbb{Q}$.

Theorem 10.18. *Every countable dense linear order* $(A, \sqsubset)$ *with end points* a_0, a_1 *is isomorphic to* $\mathbb{D}$.

Proof. As A is countable, $A = \{a_0, a_1, a_2, \ldots\}$ for some enumeration $a_0, a_1, a_2, \ldots$ which, formally, is an one-to-one function $n \mapsto a_n$ from $\mathbb{N}$ onto A.

Now one defines a function $f : \mathbb{D} \to A$ by recursion as follows: $f(0) = a_0$ and $f(1) = a_1$. After the values $f(m \cdot 2^{-n})$ have been defined for all $m \in \{0, 1, \ldots, 2^n\}$, one defines $f((2m+1) \cdot 2^{-S(n)})$ to be a_ℓ for the first ℓ with $f(m \cdot 2^{-n}) \sqsubset a_\ell \sqsubset f((m+1) \cdot 2^{-n})$. The search terminates as $(A, \sqsubset)$ is dense.

One can easily show by induction that f is order-preserving. This is true on the domain $\{0, 1\}$ by $a_0 \sqsubset a_1$. If f is order-preserving on the domain $\{m \cdot 2^{-n} \mid m \in \{0, 1, \ldots, 2^n\}\}$ then it is also order-preserving on the extended domain $\{m \cdot 2^{-S(n)} \mid m \in \{0, 1, \ldots, 2^{S(n)}\}\}$ as the new values are inserted such that the order is preserved. Thus for any $p, q \in \mathbb{D}$, $p < q \Rightarrow f(p) \sqsubset f(q)$.

As f is order-preserving, f is also one-to-one. Now one shows that f is onto. This is done by showing that

$$\forall n \in \mathbb{N}\, (a_n \in f[\{m \cdot 2^{-n} \mid m \in \{0, 1, \ldots, 2^n\}\}]).$$

This is true for a_0 as $a_0 = f(0)$. Assume now that it is true for $a_0, a_1, \ldots, a_n$. Let m be the smallest element of $\{0, 1, \ldots, 2^n\}$ with $a_{S(n)} \sqsubseteq f(m \cdot 2^{-n})$. Note that $m > 0$ as $f(0) = a_0 \sqsubset a_{S(n)}$. If $f(m \cdot 2^{-n}) = a_{S(n)}$ then there is nothing to prove. Otherwise the first index ℓ with $f((m-1) \cdot 2^{-n}) \sqsubset a_\ell \sqsubset f(m \cdot 2^{-n})$ is equal to $S(n)$ as $a_{S(n)}$ is between these two values but $a_0, a_1, \ldots, a_n$ are by induction hypothesis all of the form $f(q)$ for some q with either $q \leq (m-1) \cdot 2^{-n}$ or $q \geq m \cdot 2^{-n}$. Thus $a_{S(n)}$ is in the set $f[\{m \cdot 2^{-S(n)} \mid m \in \{0, 1, \ldots, 2^{S(n)}\}\}]$. It follows that $A \subseteq f[\mathbb{D}]$ and f is onto. ∎

Corollary 10.19. Two countable and dense linearly ordered sets are order-isomorphic iff either both sets have no end points or both sets have a minimum but no maximum or both sets have a maximum but no minimum or both sets have both end points. Furthermore, if $(A, \sqsubset)$ is an at most countable linearly ordered set then there is an order-preserving mapping from A into $(\mathbb{D}, <)$.

Proof. Let $(A, \sqsubset)$ be a countable dense linearly ordered set. It is shown that $(A, \sqsubset)$ is isomorphic to exactly one of the following four sets: $(\mathbb{D}, <)$, $(\mathbb{D} - \{0\}, <)$, $(\mathbb{D} - \{1\}, <)$, $(\mathbb{D} - \{0, 1\}, <)$.

If A has end points a_0, a_1, then this follows from Theorem 10.18. If A has no end points then one can modify A to considering new elements $-\infty, +\infty \notin A$ with $-\infty \sqsubset q \sqsubset +\infty$ for all $a \in A$. Then $(A \cup \{-\infty, +\infty\}, \sqsubset)$ is isomorphic to $(\mathbb{D}, <)$ and thus $(A, \sqsubset)$ is isomorphic to $(\mathbb{D} - \{0, 1\}, <)$. Similarly one handles the case if A has only one of the end points, that is, either a minimum or a maximum but not both.

The four sets $(\mathbb{D}, <)$, $(\mathbb{D} - \{0\}, <)$, $(\mathbb{D} - \{1\}, <)$, $(\mathbb{D} - \{0, 1\}, <)$ are not isomorphic to each other. For example, if B is either $\mathbb{D}$ or $\mathbb{D} - \{1\}$ and C is either $\mathbb{D} - \{0\}$ or $\mathbb{D} - \{0, 1\}$ and $f : B \to C$ is order-preserving then there is an element $y \in C$ with $y < f(0)$ as C has no minimum. Thus $f[B] \subset C$ and f is not an order-isomorphism. Furthermore, if B is either $\mathbb{D}$ or $\mathbb{D} - \{0\}$ and C is either $\mathbb{D} - \{1\}$ or $\mathbb{D} - \{0, 1\}$ and $f : B \to C$ is order-preserving then there is an element $y \in C$ with $y > f(1)$ as C has no maximum. Again $f[B] \subset C$ and f is not an order-isomorphism. This shows that none of these four sets are order-isomorphic to each other.

If $(A, \sqsubset)$ and $(A', \sqsubset')$ are both isomorphic to the same set $(B, <)$, then $(A, \sqsubset)$ is isomorphic to $(A', \sqsubset')$ as isomorphisms can be inverted and concatenated.

For the last statement, assume that A is a countable linearly ordered set. Then $(A \times \mathbb{Q}, <_{lex})$ is a dense linearly ordered set which is order-isomorphic via some function

g to $(\mathbb{D} - \{0,1\}, <)$. Now define $f(a) = g((a,0))$ for all $a \in A$ and one obtains an order-preserving function f from A into $\mathbb{D}$. ▮

The last result of this section is the characterization of the real line as a linearly ordered set.

Definition 10.20. A linearly ordered set $(A, <)$ is complete iff every nonempty subset of A bounded from above has a supremum in A.

Theorem 10.21. *The real line* $(\mathbb{R}, <)$ *is the unique* (*up to isomorphism*) *complete linearly ordered set without end points that has a countable subset dense in it.*

Proof. Assume that the ordered set $(A, \sqsubset)$ is a complete linearly ordered set without end points and has a countable subset B which is dense in it. The set B has no endpoints and thus there is a bijection $f : \mathbb{Q} \to B$. This function f is extended to $\mathbb{R}$ by defining

$$f(r) = \sup_{\sqsubset}\{f(q) \mid q \in \mathbb{Q} \wedge q < r\} \text{ for all } r \in \mathbb{R} - \mathbb{Q}.$$

If $r, r' \in \mathbb{R}$ are distinct than one is strictly smaller than the other, say $r < r'$. Since $\mathbb{Q}$ is a dense subset there are two rationals q, q' in between: $r < q < q' < r'$. It follows $f(r) = \sup_{\sqsubset}\{f(q'') \mid q'' \in \mathbb{Q} \wedge q'' < r\} \sqsubseteq f(q) \sqsubset f(q') \sqsubseteq \sup_{\sqsubset}\{f(q'') \mid q'' \in \mathbb{Q} \wedge q'' < r'\} = f(r')$ and thus $f(r) \sqsubset f(r')$. So f is order-preserving and one-to-one.

Assume by way of contradiction that $a \in A - f[\mathbb{R}]$. Since B has no end points, there are members $q, q' \in \mathbb{Q}$ with $f(q) \sqsubset a \sqsubset f(q')$. Now let $b = \sup_{\sqsubset}\{f(q'') \mid q'' \in \mathbb{Q} \wedge f(q'') \sqsubset a\}$. By choice of a, $b \sqsubset a$. Since B is dense in A there is a $c \in B$ in between, this is, $b \sqsubset c \sqsubset a$ and $c = f(q''')$ for some $q''' \in \mathbb{Q}$. But this contradicts to the definition of b which imposes either $f(q''') \sqsubseteq b$ or $a \sqsubset f(q''')$. Thus a cannot exist and $A = f[\mathbb{R}]$. ▮

Exercise 10.22. Assume that $(A, <)$ is linearly ordered, has no end-points, is dense and satisfies that every nonempty subset $B \subset A$ which is bounded from below has an infimum. Show that $(A, <)$ is a complete ordered set.

11 Well-Orderings

Linear orderings have a higher quality than partial orderings since every two different elements are comparable. Well-orderings are a further improvement since they generalize the property that every finite linearly ordered set has a least element to infinite subsets of the well-ordered set.

Definition 11.1. A linear ordering $<$ of a set A is a *well-ordering* of A iff every nonempty subset $B \subseteq A$ has a least element with respect to $<$. In case that $<$ is a well-ordering of A, $(A, <)$ is called a *well-ordered* set. A set A is *well-orderable* iff there exists a well-ordering of A.

Example 11.2. *Every finite linearly ordered set is a well-ordered set. The standard linear ordering of* $\mathbb{N}$ *is a well-ordering of* $\mathbb{N}$.

Proof. Since all finite sets are order-isomorphic to subsets of $\mathbb{N}$, it is sufficient to prove the second statement. Given a nonempty $B \subseteq \mathbb{N}$, there is by the Axiom of Foundation an element $m \in B$ such that $m \cap B = \emptyset$. If $n \in B - \{m\}$ then $n \notin m$ and thus $m \in n$ by the properties of natural numbers. Thus $m < n$ and therefore m is the least element of B with respect to $<$. ▮

Example 11.3. *Recall the two linear orderings*

$$(0,0) <_1 (0,1) <_1 (0,2) <_1 (0,3) <_1 \;\; ... \;\; <_1 (1,3) <_1 (1,2) <_1 (1,1) <_1 (1,0) \text{ and}$$
$$(0,0) <_2 (0,1) <_2 (0,2) <_2 (0,3) <_2 \;\; ... \;\; <_2 (1,0) <_2 (1,1) <_2 (1,2) <_2 (1,3) <_2 ...$$

on the set $\{0,1\} \times \mathbb{N}$ *from Example 10.3. The first ordering is not a well-ordering since the subset* $\{1\} \times \mathbb{N}$ *has no least element with respect to* $<_1$. *The second ordering is a well-ordering.*

Notice that the above considered well-ordered sets $(\{0\}, <)$, $(\{0,1\}, <)$, $(\{0,1,2\}, <)$, $\ldots$, $(\mathbb{N}, <)$ and $(\{0,1\} \times \mathbb{N}, <_2)$ are mutually non-isomorphic.

Example 11.4. *The lexicographic ordering of* $\mathbb{N} \times \mathbb{N}$ *is a well-ordering of* $\mathbb{N} \times \mathbb{N}$.

Proof. Recall that $(m,n) < (i,j)$ iff $(m < i)$ or $(m = i$ and $n < j)$.

The lexicographic ordering is a linear ordering. So it is sufficient to show that it is actually a well-ordering of $\mathbb{N} \times \mathbb{N}$, that is, every nonempty subset of $\mathbb{N} \times \mathbb{N}$ has a minimal element.

Let A be a nonempty subset of $\mathbb{N} \times \mathbb{N}$. For every m, let $A_m = A \cap \{(m,n) : n \in \mathbb{N}\}$. There is a least m such that A_m is not empty. Let n be the least number in $\mathbb{N}$ with $(m,n) \in A_m$. Consider any $(i,j) \in A - \{(m,n)\}$. If $i = m$ then $j > n$ by the choice of n and $(m,n) < (i,j)$. If $i \neq m$ then $i > m$ by the choice of A_m and again $(m,n) < (i,j)$. Thus (m,n) is the minimum of A with respect to the lexicographic ordering. ▮

Example 11.5. *A further well-ordering of* $\mathbb{N} \times \mathbb{N}$ *is defined as follows:*

$$\begin{aligned}(m,n) <_{cw} (i,j) \;\Leftrightarrow\; & \max\{m,n\} < \max\{i,j\} \\ \vee\; & (\max\{m,n\} = \max\{i,j\} \wedge m < i) \\ \vee\; & (\max\{m,n\} = \max\{i,j\} \wedge m = i \wedge n < j).\end{aligned}$$

Proof. It is established that $<_{cw}$ is a well-ordering by showing that the following four conditions hold.

Antireflexiveness. $(m,n) \not<_{cw} (m,n)$ since $(m,n) <_{cw} (i,j)$ requires that either $\max\{m,n\} \neq \max\{i,j\}$ or $m \neq i$ or $n \neq j$ and none of these conditions holds if $(m,n) = (i,j)$.

Transitivity. Assume the following two conditions $(*)$:

$$(m,n) <_{cw} (i,j) \text{ and } (i,j) <_{cw} (h,k).$$

It is shown that $(*)$ implies $(m,n) <_{cw} (h,k)$.

If $\max\{m,n\} < \max\{h,k\}$ then $(m,n) <_{cw} (h,k)$. Otherwise $\max\{m,n\} = \max\{i,j\} = \max\{h,k\}$ and the relation $<_{cw}$ follows at $(*)$ the second or third case of the disjunction in its definition. If $m < h$ then again $(m,n) <_{cw} (h,k)$. Otherwise $m = i = h$ and the relation follows at $(*)$ the third case of its definition. Thus $\max\{m,n\} = \max\{h,k\}$, $m = h$ and $n < k$ by $n < j < k$. So again $<_{cw}$ holds and $<_{cw}$ is transitive.

Comparability. Assume that neither $(m,n) <_{cw} (i,j)$ nor $(i,j) <_{cw} (m,n)$. Then $\max\{m,n\} = \max\{i,j\}$, $m = i$ and $n = j$, that is, $(m,n) = (i,j)$. Thus any two different members of $\mathbb{N} \times \mathbb{N}$ are comparable.

Well-orderedness. Let $A \subseteq \mathbb{N} \times \mathbb{N}$ be nonempty. Let

$$A_k = \{(m,n) \in A \mid \max\{m,n\} = k\}.$$

Fix k as the least number such that A_k is nonempty. This set A_k is a finite set and has a least element (m,n) with respect to $<_{cw}$ since $<_{cw}$ is a linear order on $\mathbb{N} \times \mathbb{N}$ and also on its subset A_k. Now let $(i,j) \in A - \{(m,n)\}$. If $(i,j) \notin A_k$ then $\max\{i,j\} > k = \max\{m,n\}$ and $(m,n) <_{cw} (i,j)$. If $(i,j) \in A_k$ then $(m,n) < (i,j)$ by the choice of (m,n) from A_k. Thus A has a minimum and $(\mathbb{N} \times \mathbb{N}, <_{cw})$ is a well-ordered set. ∎

Remark 11.6. Notice that the ordered set $(\mathbb{N} \times \mathbb{N}, <_{cw})$ is isomorphic to $(\mathbb{N}, <)$. Indeed the function f given as

$$f(m,n) = (2\max\{m,n\} + 2)^3 + (\max\{m,n\} + m + 1)^2 + n$$

is an order-preserving one-to-one mapping into $\mathbb{N}$. Thus, $(\mathbb{N} \times \mathbb{N}, <_{cw})$ is isomorphic to an infinite subset of $(\mathbb{N}, <)$ which is then isomorphic to $(\mathbb{N}, <)$.

Example 11.7. *The following subsets of $\mathbb{Q}$ are well-ordered with respect to the natural ordering of $\mathbb{Q}$:*

$$\begin{aligned}
\{-\tfrac{1}{n+1} \quad &| \quad n \in \mathbb{N}\}, \\
\{-\tfrac{1}{m+1} - \tfrac{1}{n+1} \quad &| \quad m, n \in \mathbb{N}\}, \\
\{-\tfrac{1}{k+1} - \tfrac{1}{m+1} - \tfrac{1}{n+1} \quad &| \quad k, m, n \in \mathbb{N}\}.
\end{aligned}$$

The orderings are isomorphic to that of the lexicographic ordering on $\mathbb{N}$, $\mathbb{N} \times \mathbb{N}$, $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$, respectively; the lexicographic ordering of $\mathbb{N}$ is of course identical with the natural one.

Exercise 11.8. The set

$$\{-\tfrac{1}{m_1+1} - \tfrac{1}{m_2+1} - \ldots - \tfrac{1}{m_n+1} \mid n, m_1, m_2, \ldots, m_n \in \mathbb{N}\}$$

is not a well-ordered subset with respect to the natural ordering of $\mathbb{Q}$: show that the set is dense and is not bounded from below.

Example 11.9. *Both $\mathbb{Z}$ and $\mathbb{Q}$ are well-orderable, but the ordering differs from the standard one.*

Proof. In fact every countable set X is well-orderable. Since $|X| \leq |\mathbb{N}|$, there is a one-to-one function $f : X \to \mathbb{N}$. Now one defines on X a well-ordering $\sqsubset$ by

$$x \sqsubset y \Leftrightarrow f(x) < f(y)$$

where $x, y \in X$. This order differs from the natural order on $\mathbb{Z}$ and $\mathbb{Q}$: these sets contain the chain $-1, -2, -3, \ldots$ which is descending with respect to their natural order and which cannot be so with respect to any well-ordering of them. ▮

Definition 11.10. For a linearly ordered set $(L, <)$, an *initial segment* I of L is a proper subset of L such that $x \in I$ whenever $x \in L$ and there is an $y \in I$ with $x < y$. That is, I is an initial segment iff I is a downward closed proper subset of L:

$$I \subset L \wedge \forall x, y \in L\,(x < y \wedge y \in I \Rightarrow x \in I).$$

For $a \in L$, $L[a] = \{x \in L \mid x < a\}$. Call $L[a]$ the initial segment of L given by a.

Proposition 11.11. If $(W, <)$ is well-ordered, I is an initial segment of W, then there is an $a \in W$ such that $I = W[a]$.

Proof. Let $A = W - I$. Then A is not empty and every element of A is an upper bound of I since I is an initial segment. Let a be the least element of A with respect to $<$. Then for $x \in W$, $x < a$ if and only if $x \in I$. ▮

Notice that $\{r \in \mathbb{Q} \mid r < \sqrt{2}\}$ is an initial segment of $\mathbb{Q}$ but it is not $\mathbb{Q}[a]$ for any $a \in \mathbb{Q}$. The set of real numbers is isomorphic to every initial segment and also has a large quantity of isomorphisms onto itself. Well-ordered sets are rigid, that is, they satisfy exactly the opposite of these properties.

Theorem 11.12 (Rigidity). *Let $(A, <)$ be a well-ordered set and f be an order-preserving function from $(A, <)$ to itself. Then $a \leq f(a)$ for all $a \in A$. In particular, the range cannot be an initial segment of A and $(A, <)$ is not isomorphic to any initial segment. Furthermore, if f is an isomorphism from A to itself, then f is the identity.*

Proof. If $f(a) < a$ then $f(f(a)) < f(a)$ since f is order-preserving. Thus there is no least element a with $f(a) < a$. Since $(A, <)$ is well-ordered, there is even no element $a \in A$ with $f(a) < a$.

Given an initial segment of A, it is of the form $A[a]$ for some $a \in A$. Since $f(a) \geq a$, $f[A] \not\subseteq A[a]$ and the initial segment is not the range of f. Since the choice of f was arbitrary, there is no isomorphism from A to any initial segment.

Assume now that f is not the identity. Then there is a least element $a \in A$ with $f(a) \neq a$. As seen above, $f(a) > a$. Thus, for all b, $f(b) \neq f(a)$: If $b < a$ then $f(b) = b \neq a$ by the choice of a; if $b \geq a$ then $f(b) \geq f(a) > a$ by the fact that f is order-preserving. Thus the identity is the only isomorphism of $(A, <)$. ∎

Theorem 11.13 (Comparability Theorem). *Given two well-ordered sets, they are either isomorphic or exactly one of them is isomorphic to an initial segment of the other.*

Proof. Given $(A_1, <_1)$ and $(A_2, <_2)$, let $B = \mathcal{P}(A_1 \times A_2)$. Call $F \in B$ consistent iff the following conditions hold:

1. if $(a_1, a_2), (b_1, b_2) \in F$ then either $a_1 <_1 b_1 \wedge a_2 <_2 b_2$ or $a_1 = b_1 \wedge a_2 = b_2$ or $b_1 <_1 a_1 \wedge b_2 <_2 a_2$.

2. if $(a_1, a_2) \in F$ and $b_1 <_1 a_1$ then there is an $b_2 <_2 a_2$ such that $(b_1, b_2) \in F$.

3. if $(a_1, a_2) \in F$ and $b_2 <_2 a_2$ then there is an $b_1 <_1 a_1$ such that $(b_1, b_2) \in F$.

Now the following facts hold for all consistent F, G:

1. *If $F \not\subseteq G$ then $G \subset F$.* The elements of F are well-ordered by the ordering inherited from $(A_1, <_1)$ or $(A_2, <_2)$; both inherit the same ordering. There is a least pair $(a_1, a_2) \in F - G$. Let H be the set $\{(c_1, c_2) \in F \mid c_1 <_1 a_1\}$ of the pairs in F below (a_1, a_2). Clearly $H \subseteq F \cap G$. Assume now that $H \subset G$. Then there is a least element $(b_1, b_2) \in G - H$. Since F, G are consistent, there is for

every $c_1 <_1 a_1$ a c_2 with $(c_1, c_2) \in H$ and similarly for every $c_2 <_2 a_2$ a c_1 with $(c_1, c_2) \in H$. By consistency b_1 is the least element in A_1 different from these c_1 and b_2 the least element of A_2 different from these c_2, that is, $b_1 = a_1$ and $b_2 = a_2$ in contradiction to the choice of (a_1, a_2). Thus (b_1, b_2) does not exist and $G = H \subset F$.

2. *There is a maximal consistent set.* Every consistent set is an element of the power set $\mathcal{P}(A_1 \times A_2)$ and the property consistent is first order definable from $\mathcal{P}(A_1 \times A_2)$, A_1 and A_2 as shown above. So there is a set C of consistent sets. The consistent sets are linearly ordered by inclusion and their union is again consistent. Thus, $F = \bigcup C$ is a consistent set which is maximal.

3. *The maximal consistent set F is a partial one-to-one function with either domain A_1 or range A_2.* The property of being a bijection from the domain to the range comes from the definition, similarly the domain is a subset of A_1 and the range a subset of A_2. Assume now by way of contradiction that both subsets would be proper. Then there is a least $a_1 \in A_1$ which is outside the domain of F and a least $a_2 \in A_2$ which is outside the range of F. This would give that $F \cup \{(a_1, a_2)\}$ is also consistent in contradiction to the maximality of F. Thus only one inclusion can be proper.

If the domain of F is A_1 and the range of F is A_2 then $(A_1, <_1)$ and $(A_2, <_2)$ are isomorphic. If the range of F is a proper subset of A_2 then F is an isomorphism from $(A_1, <_1)$ is an initial segment of $(A_2, <_2)$. If the domain of F is a proper subset of A_1 then F^{-1} is an isomorphism from $(A_2, <_2)$ is an initial segment of $(A_1, <_1)$. ∎

Exercise 11.14. Define a function $f : \{0, 1, \ldots, 9\}^* \to \mathbb{N}$ which is order-preserving with respect to the length-lexicographic ordering $<_{ll}$: $v <_{ll} w \Leftrightarrow f(v) < f(w)$. Recall $0 <_{ll} 1 <_{ll} \ldots <_{ll} 9 <_{ll} 00 <_{ll} 01 <_{ll} \ldots <_{ll} 99 <_{ll} 000 <_{ll} \ldots$ and $v <_{ll} w$ if either v is shorter than w or v, w have the same length and $v <_{lex} w$.

12 Ordinals

Ordinals are a generalization of the natural numbers. While a natural number (viewed as the set which represents it) is order-isomorphic to finite well-ordered sets, ordinals are a generalization which is just taken to be order-isomorphic to any well-ordered sets. Recall from Definition 4.4 that a set A is transitive iff $\forall a \in A\, \forall b \in a\, (b \in A)$. Ordinals are now well-ordered and represented by transitive sets.

Definition 12.1. A set is an *ordinal* (or *ordinal number*) if it is transitive and well-ordered by the ordering $\in$ (restricted to its members).

Example 12.2. *Every natural number is an ordinal. The sets $\mathbb{N}$ and $\mathbb{N} \cup \{\mathbb{N}\}$ are ordinals. The set $\{2, 3, 4, 5, 6, 7, 8\}$ is well-ordered by $\in$ but not transitive. The set $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\{\emptyset\}\}\}$ is transitive but not linearly ordered by $\in$.*

Convention 12.3. Ordinals are normally written by lower case Greek letters.

Definition 12.4. ω is the first ordinal after 0 which is not the successor of any other ordinal. That is, ω is ordinal represented by $\mathbb{N}$. The ordinals strictly below ω are called finite and those beyond ω are called transfinite.

Remark 12.5. There are in principal two options to represent the natural numbers. Both represent 0 by $\emptyset$. Having the codes for $0, 1, \ldots, n$, the first one would represent $S(n)$ as $\{Code(n)\}$ while the second one would represent $S(n)$ as $\{Code(0), \ldots, Code(n)\}$. The advantage of the second approach is that it also permits to represent transfinite ordinals (as already indicated above), which is impossible in the first approach. Thus the second approach was taken in Definition 4.3. Examples for the two representations are:

Number,	First and	Second Representation.
0	$\emptyset$	$\emptyset$
1	$\{\emptyset\}$	$\{\emptyset\}$
2	$\{\{\emptyset\}\}$	$\{\emptyset, \{\emptyset\}\}$
3	$\{\{\{\emptyset\}\}\}$	$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
4	$\{\{\{\{\emptyset\}\}\}\}$	$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}$
...	...	...
ω	—	$\{Code(\alpha) \mid \alpha \in \mathbb{N}\}$
$\omega + 1$	—	$\{Code(\alpha) \mid \alpha \in \mathbb{N} \vee \alpha = \omega\}$

So the second representation was taken in set theory since it codes natural and ordinal numbers in a uniform way and satisfies that

$$\alpha < \beta \Leftrightarrow \alpha \in \beta \Leftrightarrow \alpha \subset \beta.$$

Furthermore, the second notation also reflects the intuition from counting that n is a set of n objects, namely the representatives of the n smaller numbers. The price paid is that it takes much space to write down even small numbers, see Exercise 8.11.

Theorem 12.6. *If A is transitive and $(A, \in)$ linearly ordered then A is an ordinal.*

Proof. Let $B \subseteq A$ and B be nonempty. By the Axiom of Foundation there is $x \in B$ such that $y \notin x$ for all $y \in B$. Since $(A, \in)$ is linearly ordered, the same holds for

$(B, \in)$ and x is the least element of B with respect to the ordering given by $\in$. Thus A is an ordinal. ∎

Exercise 12.7. Verify the following properties of ordinals.

1. If α is an ordinal, then $S(\alpha)$, which is defined as $\alpha \cup \{\alpha\}$, is also an ordinal.
2. Every element of an ordinal is an ordinal.
3. An ordinal α is transfinite iff $|\alpha| = |S(\alpha)|$.
4. An ordinal α is finite iff $S(\alpha) = \{0\} \cup \{S(\beta) \mid \beta \in \alpha\}$.

Theorem 12.8. *The following basic facts hold for ordinals:*

1. *If α, β are ordinals then either $\alpha \in \beta$ or $\alpha = \beta$ or $\beta \in \alpha$.*
2. *If A is a set of ordinals then $\bigcup A$ is an ordinal.*
3. *If A is a nonempty set of ordinals then there exists an ordinal $\alpha \in A$ such that $\alpha \cap A = \emptyset$. Consequently every set of ordinals is well-ordered by $\in$.*

Proof. 1. Let α, β be ordinals such that the cases $\alpha = \beta$ and $\alpha \in \beta$ do not hold, that is, $\alpha \not\subseteq \beta$. Then there is a least element of $\alpha - \beta$, say γ. Since $\gamma = \{\delta \mid \delta < \gamma\}$ and every δ is in α by transitivity, $\gamma \subseteq \beta$. If $\gamma \subset \beta$ then $\gamma \in \beta$ in contradiction to the choice of γ. Thus $\gamma = \beta$ and $\beta \in \alpha$.

2. $\bigcup A$ is the union of transitive sets which are linearly ordered by $\in$. The union is again transitive. Furthermore, if $\alpha, \beta \in \bigcup A$ then α and β are comparable by the previous paragraph. If $\alpha \in \beta$ and $\beta \in \gamma$ then $\alpha \in \gamma$ since γ itself is an ordinal and transitive. The Axiom of Foundation gives $\alpha \notin \alpha$ for all α. So $\bigcup A$ is a transitive set which is linearly ordered by $\in$. By Theorem 12.6, $\bigcup A$ is an ordinal.

3. The set A inherits from the superset and ordinal $\bigcup A$ that $\in$ is a well-ordering. Since A is nonempty, it has a least element α with respect the ordering given by $\in$. Then $A \cap \alpha$ contains only ordinals below α and is thus empty. ∎

Exercise 12.9. Use the above results to show that there is no set containing all ordinals in V.

Definition 12.10. An ordinal α is called *successor-ordinal* if $\alpha = S(\beta) = \beta \cup \{\beta\}$ for some other ordinal β and is called *limit ordinal* otherwise. The supremum of a set A of ordinals is denoted by $\sup A$, note that $\sup A = \bigcup A$.

Example 12.11. 0 *is a limit ordinal and all the positive natural numbers are successor ordinals.* $\omega = \sup \mathbb{N}$ *is a next limit ordinal. One can combine the usage of supremums and successors to obtain every ordinal from those below. So, for given* α*, one has*

$$\alpha = \bigcup \{S(\beta) \mid \beta \in \alpha\} = \sup\{S(\beta) \mid \beta < \alpha\}$$

and this rule holds also for $\alpha = 0$ *by using the definition* $\sup \emptyset = 0$*. One should also note that above every ordinal* α *is a successor ordinal, namely* $S(\alpha)$*, and also a limit ordinal obtained as* $\sup f_\alpha[\mathbb{N}]$ *where* $f_\alpha(0) = \alpha$ *and* $f_\alpha(S(n)) = S(f_\alpha(n))$*.*

13 Transfinite Induction and Recursion

Induction and Recursion can be generalized to ordinals and the universe of sets. This says that one can prove theorems and build functions along the membership relation $\in$ from the bottom to the top.

Theorem 13.1 (Transfinite Induction on Ordinals). *Let* $p(x)$ *be a property. Assume that for every ordinal* α*,*

1. $p(0)$ *holds;*
2. *if there is an ordinal* β *with* $S(\beta) = \alpha$ *and if* $p(\beta)$ *holds then also* $p(\alpha)$ *holds;*
3. *if* α *is a limit ordinal and* $p(\beta)$ *holds for all* $\beta < \alpha$ *then* $p(\alpha)$ *holds.*

Then it can be concluded that $p(\alpha)$ *holds for all ordinals* α*.*

Remark 13.2. There are several equivalent statements of Transfinite Induction.

1. If for every ordinal α the implication $(\forall \beta < \alpha\, (p(\beta))) \Rightarrow p(\alpha)$ is true, then $p(\alpha)$ is true for all ordinals α.
2. If there is no minimal α satisfying $\neg p(\alpha)$ then $p(\alpha)$ is true for all α.
3. If for every α where $p(\alpha)$ is false there is another $\beta < \alpha$ where $p(\beta)$ is false, then $p(\alpha)$ is true for all ordinals α.

Note that due to the Axiom of Foundation one can get a counterpart to transfinite induction on $(V, \in)$.

Theorem 13.3 (Transfinite Induction in V). *Assume that for a property* p *and all* $x \in V$ *the implication*

$$(\forall z \in x\, (p(z))) \Rightarrow p(x)$$

holds. Then $p(x)$ is satisfied for all $x \in V$.

Proof. Assume that there is an $x \in V$ where $p(x)$ is false. Let $x' = \{z \in \mathcal{TC}(x) \mid p(z)\}$. The set $\mathcal{TC}(x) - x'$ is nonempty and there is by the Axiom of Foundation a $y \in \mathcal{TC}(x) - x'$ such that every $z \in y$ is in not in $\mathcal{TC}(x) - x'$. Recall that by definition, $\mathcal{TC}(x)$ is transitive and thus all members of y are in $\mathcal{TC}(x)$. Thus they are also in x' and one has that $\forall z \in y\ p(z)$ is true. So $p(y)$ holds and $y \in x'$ in contrary to its choice. Thus x' must be empty and $p(x)$ is true as well. So $p(x)$ holds for all $x \in V$. ▮

Example 13.4. *Let F be a class defining a function in one variable. If $F(x) = F[x]$ for all $x \in V$ then F is the identity.*

Proof. Assume that $F(y) = y$ for all $y \in x$. Then $F(x) = F[x] = \{F(y) \mid y \in x\}$ $= \{y \mid y \in x\} = x$. Thus the equality holds also for x. It follows from transfinite induction that F is the identity. ▮

Well-founded relations are a generalization of both, the element relation and a well-ordering.

Defintion 13.5. A relation R on a domain W (which is either a class or a set) is well-founded iff

- for every $x \in W$, $\{y \in W \mid y\, R\, x\}$ is a set;
- for every nonempty set $x \subseteq W$ there is an $y \in x$ such that no $z \in x$ satisfies $z\, R\, y$.

Note that the first condition is only important for the case that W is a proper class, that is, not a set. Furthermore, choosing for any $y \in W$ the subset $x = \{y\}$ of W, proves that (W, R) is irreflexive. Now some examples of well-founded relations are given:

- Assume that $x\, R\, y$ iff $x \in y$. Then R is well-founded relation on V by the Axiom of Foundation.
- Assume that $(W, <)$ is a well-ordered set and $x\, R\, y$ iff $x < y$. Then R is a well-founded relation on W.

One can use well-founded relations to generalize recursion from the natural numbers to many other structures like well-ordered sets, the class of ordinals and even the whole universe V along $\in$. Note that not only recursion but also transfinite induction can be carried out along any well-founded relation.

Exercise 13.6. Let A be some set and let $a_0a_1\ldots a_{n-1}\,R\,b_0b_1\ldots b_{m-1} \Leftrightarrow n < m$ and there is a function $f : n \to m$ such that $b_{f(i)} = a_i$ and $(i < j \Rightarrow f(i) < f(j))$ for all $i, j \in n$ where $a_0a_1\ldots a_{n-1}, b_0b_1\ldots b_{m-1} \in A^*$. Show that R is well-founded.

Let R be a relation on V such that $x\,R\,y$ iff there is a z with $x \in z \wedge z \in y$. Show that R is well-founded.

Let R be a relation on $V \times V$ such that $(x, y)\,R\,(v, w)$ iff either $x = v \wedge y \in w$ or $y = w \wedge x \in v$. Is R well-founded?

Is the relation R on V given as $x\,R\,y \Leftrightarrow x \cap y = x \cup y$ well-founded?

Theorem 13.7 (Transfinite Recursion). *Let R be a well-founded relation with domain W and let G be a class which is a function in $n+1$ variables. Then there is a class F which is also a function in n variables and satisfies*

$$\forall x_1, \ldots, x_n \in W\,(F(x_1, \ldots, x_n) = G(x_1, \ldots, x_n, \{(y_1, F(y_1, x_2, \ldots, x_n)) \mid y_1\,R\,x_1\})).$$

Proof The proof is similar to that of Theorem 5.2. Let W be the domain of R and R^* be the transitive closure of R, which exists by Theorem 5.2.

More formally, R^1 is the same relation as R and define inductively $v\,R^{n+1}\,w$ if $v\,R^n\,w$ or $v\,R^n\,u$ and $u\,R^n\,w$ for some $u \in W$. Furthermore, $v\,R^*\,w$ iff $v\,R^n\,w$ for some $n \in \{1, 2, 3, \ldots\}$. One can easily verify the following four facts on R^* to obtain that this relation is also well-founded: First one can show by induction that for each n and $w \in W$, the set $\{v \in W : v\,R^n w\}$ is a set. The same applies then for the inductively defined set R^*. Second, R^* is transitive. Third, a every non-empty set $A \subseteq W$ has an minimal element with respect to R^* as the set $B = \{w \in W : \exists u, v \in A : u\,R^*\,w\,R^* v\}$ has a minimal element with respect to R. Fourth, R^* is antisymmetric.

This transitive closure R^* of R will now be used to define a class C which will be used to define the function F. So let C be the class of all functions f such that

- The domain of f is a set of the form $\{(y_1, x_2, \ldots, x_n) \mid y_1 = x_1 \vee y_1\,R^*\,x_1\}$ for some $x_1, x_2, \ldots, x_n \in W$.
- If $(z_1, x_2, \ldots, x_n)$ is in the domain of the function f then $f(z_1, x_2, \ldots, x_n) = G(z_1, x_2, \ldots, x_n, \{(y_1, f(y_1, x_2, \ldots, x_n)) \mid y_1\,R\,z_1\})$.

Now the function F is defined as the union over all functions $f \in C$, that is,

$$F = \{(x_1, x_2, \ldots, x_n, f(x_1, \ldots, x_n)) \mid f \in C \wedge f(x_1, x_2, \ldots, x_n) \text{ is defined}\}.$$

It is now shown that F is actually a function. This is done by considering the following subclass D of the class of all n tuples of elements in W.

D is the class of all n-tuples $(x_1, x_2, \ldots, x_n)$ of elements in W such that there is a function $f \in C$ for which $f(x_1, x_2, \ldots, x_n)$ is defined and such that for all functions $f, \tilde{f} \in C$ where $f(x_1, \ldots, x_n), \tilde{f}(x_1, \ldots, x_n)$ are defined, these values coincide.

Now one shows by transfinite induction that $D = W^n$. Let $(x_1, x_2, \ldots, x_n)$ be any tuple of n elements in W and assume that $(y_1, x_2, \ldots, x_n) \in D$ for all $y_1 \, R^* \, x_1$. Now define $f(y_1, x_2, \ldots, x_n) = \bigcup \{\tilde{f}(y_1, x_2, \ldots, x_n) : \tilde{f} \in C \wedge \tilde{f}(y_1, x_2, \ldots, x_n)$ is defined$\}$ for all $y_1 \, R^* \, x_1$. It follows from the membership of $(y_1, x_2, \ldots, x_n)$ in D that $f(y_1, x_2, \ldots, x_n) = \tilde{f}(y_1, x_2, \ldots, x_n)$ whenever $\tilde{f} \in C$ and $\tilde{f}(y_1, x_2, \ldots, x_n)$ is defined. Consider $f \cup \{(x_1, x_2, \ldots, x_n, G(x_1, x_2, \ldots, x_n, \{(y_1, f(y_1, x_2, \ldots, x_n)) \mid y_1 \, R \, x_1\}))\}$; this function is in C. If there is a further function $\tilde{f} \in C$ for which $\tilde{f}(x_1, x_2, \ldots, x_n)$ is defined, then $\tilde{f}$ coincides with f on the domain of f and hence $\tilde{f}(x_1, x_2, \ldots, x_n)$ coincides with $G(x_1, x_2, \ldots, x_n, \{(y_1, f(y_1, x_2, \ldots, x_n)) \mid y_1 \, R \, x_1\})$. So it follows that $(x_1, \ldots, x_n) \in D$.

Hence the class $F = \bigcup C$ is actually a function mapping n-tuples in W to V and so F exists. In the case that W is a set, $F[W \times W \times \ldots \times W]$ is a set as well by the Axiom of Replacement. ∎

Informally, this means that whenever R is a well-founded relation on W and some class G says how to obtain $F(x_1, x_2, \ldots, x_n)$ from the arguments $x_1, x_2, \ldots, x_n$ and all pairs $(y_1, F(y_1, x_2, \ldots, x_n))$ with $y_1 \, R \, x_1$, then F itself exists (that is, F is a class).

Example 13.8. *The function $\mathcal{TC}$ can be defined with transfinite recursion along $\in$ via the formula*

$$\mathcal{TC}(x) = \{x\} \cup \bigcup \{\mathcal{TC}(y) \mid y \in x\};$$

and $\mathcal{TC}$ coincides with the successor S on ordinals which is an expression. They are different on sets which are not ordinals as $S(\{7,8\}) = \{7, 8, \{7,8\}\}$ and $\mathcal{TC}(\{7,8\}) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, \{7, 8\}\}$. Note that $\mathcal{TC}(\emptyset) = \{\emptyset\}$ as $\bigcup \{\mathcal{TC}(y) \mid y \in \emptyset\}$ is just $\emptyset$.

An important application of transfinite recursion is the following result.

Theorem 13.9 (Representation Theorem). Let $(W, \sqsubset)$ be a well-ordered set. Then there is an ordinal isomorphic to this set.

Proof. Using transfinite recursion one can define $F : W \to V$ by the equation

$$F(a) = \bigcup \{S(F(b)) \mid b \sqsubset a\}.$$

Note that $F(a) = 0$ if a is the least element of W with respect to $\sqsubset$ since the union over the members of the empty set gives the empty set: $\bigcup \emptyset = \emptyset$. Furthermore, 0 is the ordinal represented by $\emptyset$. It is easy to see that for all $a, b \in W$ the implication $b \sqsubset a \Rightarrow S(F(b)) \subseteq F(a) \Rightarrow F(b) \in F(a)$ holds. So F is order-preserving. Furthermore, if $\beta \in F(a)$ then there is a least $b \in W$ with $\beta \in F(b)$. All $c \sqsubset b$ satisfy $\beta \notin F(c)$. So $\beta \in \bigcup\{S(F(c)) \mid c \sqsubset b\}$ but $\beta \notin \bigcup\{F(c) \mid c \sqsubset b\}$. Thus there is a $c \sqsubset b$ with $\beta \in S(F(c)) - F(c)$. It follows that b is the successor of c with respect to $\sqsubset$ and $\beta = F(c)$. So $F[W]$ is transitive. Furthermore, by the Axiom of Replacement, $F[W]$ is a set itself. So $F[W]$ is an ordinal. ∎

Theorem 13.10. *For every set X there is an ordinal α such that $|\alpha| \not\leq |X|$.*

Proof. Given X, let $Y = \{(Z, R) \mid Z \subseteq X \wedge R$ is a well-ordering on $Z\}$. For all $(Z, R), (Z', R') \in Y$, let $(Z, R) \sqsubset (Z', R') \Leftrightarrow (Z, R)$ is isomorphic to an initial segment of (Z', R'). The relation $\sqsubset$ is well-founded: If $U \subseteq Y$ then either $U = \emptyset$ or U contains some (Z, R). In the latter case, either (Z, R) is a minimal element or there is a least $z \in Z$ such that some $(Z', R') \in U$ is isomorphic to $(Z[z], R)$ and then that (Z', R') is a minimal element of U.

So one can define by transfinite recursion a recursive function $F : Y \to V$ such that

$$F((Z, R)) = \bigcup\{S(F(Z', R')) \mid (Z', R') \in Y \wedge (Z', R') \sqsubset (Z, R)\}$$

where the well-founded relation on the domain of F is $\sqsubset$.

Assume now by way of contradiction that the range of F is not an ordinal: then one can define the nonempty set $Y' = \{(Z, R) \in Y : F((Z, R))$ is not an ordinal$\}$ and Y' has a minimal element (Z, R) with respect to $\sqsubset$. For this minimal element, one has that it is the union of all sets $S(F((Z', R')))$ with $(Z', R') \in Y \wedge (Z', R') \sqsubset (Z, R)$, hence $F((Z, R))$ is the union of ordinals and hence $F((Z, R))$ is an ordinal itself in contradiction to the assumption. Hence, the range of F is a set of ordinals and the union of these ordinals is an ordinal β. Every member of Y is isomorphic to a initial segment of $\alpha = S(\beta)$.

If there would be a one-to-one function $g : \alpha \to X$ then $g[\alpha] \in V$ and $g[\alpha] \subseteq X$. Furthermore, g induces a well-ordering R on $g[\alpha]$ and $(g[\alpha], R) \in Y$ in contradiction to the fact that no member of Y is isomorphic to α. Thus there is no such g and $|\alpha| \not\leq |X|$. ∎

Exercise 13.11. Construct by transfinite recursion a function on ordinals which tells whether an ordinal is even or odd. More formally, construct a function F such that $F(\alpha) = 0$ if α is even, $F(\alpha) = 1$ if α is odd. Limit ordinals should always be even; the successor of an even ordinal is odd and the successor of an odd ordinal is even.

Exercise 13.12. Is it possible to define a function F on all sets such that $F(X) = n$ iff n is the maximal number such that there are $Y_0, Y_1, \ldots, Y_n$ with $Y_{m+1} = S(Y_m)$ for all $m \in n$ and $X = Y_n$? If so, construct the corresponding function F by transfinite recursion.

14 The Rank of Sets

The rank is an alternative method to measure the size of a set. The cardinality asks how many elements are in the set, the rank asks how many levels are necessary to build a set. The rank is defined by transfinite recursion.

Definition 14.1. The rank ρ which is defined as $\rho(x) = \bigcup\{S(\rho(y)) \mid y \in x\}$ with $\rho(\emptyset) = \bigcup \emptyset = 0$.

Example 14.2. $\rho(1) = \rho(\{\emptyset\}) = 1$, $\rho(2) = \rho(\{\emptyset, \{\emptyset\}\}) = 2$, $\rho(\{\{\emptyset\}, \{\{\emptyset\}\}\}) = 3$, $\rho(\{A\}) = S(\rho(A))$ *and* $\rho(A \cup B) = \rho(A) \cup \rho(B)$ *for all sets* A, B.

Proposition 14.3. *The rank ρ is an ordinal-valued function with $\rho(\alpha) = \alpha$ for all ordinals α.*

Proof. By Theorem 12.6, $\rho(x)$ is an ordinal iff $\rho(x)$ is transitive and linearly ordered by $\in$. Being an ordinal is a property. So, for given $x \in V$, one can define the set

$$x' = \{y \in \mathcal{TC}(x) \mid \rho(y) \text{ is an ordinal}\}$$

by comprehension. Assume by way of contradiction that $x \notin x'$. Then $\mathcal{TC}(x) - x'$ is not empty and has an element y such that no $z \in y$ is in $\mathcal{TC}(x) - x'$, that is, $y \subseteq x'$. Then $\rho(z)$ is an ordinal for every $z \in y$ and $\rho(y) = \bigcup\{S(\rho(z)) \mid z \in y\}$ is an ordinal by Theorem 12.8. Then $y \in x'$ contradicting the choice of y; this contradiction gives $x \in x'$. In particular, $\rho(x)$ is an ordinal.

Recall that $\alpha = \bigcup\{S(\beta) \mid \beta \in \alpha\}$ for all ordinals α. Assuming that $\rho(\beta) = \beta$ for all $\beta \in \alpha$, one has that $\rho(\alpha) = \bigcup\{S(\rho(\beta)) \mid \beta \in \alpha\} = \bigcup\{S(\beta) \mid \beta \in \alpha\} = \alpha$ for α. The equality $\rho(\alpha) = \alpha$ holds for all ordinals α by transfinite induction. ∎

Exercise 14.4. For any ordinal α, consider the successor function S restricted to α, that is, consider the set

$$S \upharpoonright \alpha = \{\{\beta, \{\beta, S(\beta)\}\} \mid \beta \in \alpha\}.$$

Determine $\rho(S \upharpoonright \alpha)$ for $\alpha = 42, 1905, 2004, \omega, \omega + 1, \omega + 131501, \omega^2 + \omega \cdot 2 + 1, \omega^{17} + \omega^4$.

Theorem 14.5. *For every ordinal α let $V_\alpha = \{x \in V \mid \rho(x) < \alpha\}$. Then V_α is a set and $\rho(V_\alpha) = \alpha$.*

Proof. Define a function G by

$$G(\alpha, x) = \bigcup\{\mathcal{P}(z) \mid \exists y\, ((y, z) \in x)\}.$$

Let F be the function obtained from G by transfinite recursion on ordinals. That is, F satisfies

$$F(\alpha) = \bigcup \{\mathcal{P}(F(\beta)) \mid \beta < \alpha\}$$

for all ordinals α. Now one can show by transfinite induction that F maps ordinals to sets. $F(\emptyset) = \emptyset$ is a set. If α is a successor ordinal and $\alpha = S(\beta)$ then $F(\alpha) = \mathcal{P}(F(\beta))$ and $F(\alpha)$ is a set. If α is a limit ordinal then $F(\alpha) = \bigcup F[\alpha]$ and $F(\alpha) \in V$ by the Axiom of Replacement.

The equality $V_\alpha = F(\alpha)$ is shown by transfinite induction. That is, assuming that equality holds for all $\beta \in \alpha$, one has to show that the equality holds for α as well.

If $x \in F(\alpha)$ then $x \in \mathcal{P}(F(\beta))$ and $x \subseteq F(\beta)$ for some $\beta \in \alpha$. By induction hypothesis, $\rho(y) < \beta$ for all $y \in x$. Thus $\rho(x) < S(\beta) \leq \alpha$ and $x \in V_\alpha$.

If $x \in V_\alpha$ then $\rho(x) = \beta < \alpha$ for some β. Every $y \in x$ satisfies $\rho(y) < \beta$ and $y \in V_\beta$. By induction hypothesis, $V_\beta = F(\beta)$. Since $F(\beta)$ is a set by the Axiom of Replacement, $\mathcal{P}(F(\beta))$ exists and $x \in \mathcal{P}(F(\beta))$. It follows that $x \in F(\alpha)$.

So $F(\alpha)$ and V_α have the same elements. By the Axiom of Extensionality they are equal. Thus the mapping $\alpha \mapsto V_\alpha$ is a function and V_α is a set for every ordinal α.

On one hand, V_α consists only of elements x with $\rho(x) < \alpha$. Thus $\rho(V_\alpha) \leq \alpha$. On the other hand, every $\beta < \alpha$ satisfies $\rho(\beta) = \beta$ and $\beta \in V_\alpha$. So $\alpha \subseteq V_\alpha$ and $\rho(V_\alpha) \geq \alpha$. Thus $\rho(V_\alpha) = \alpha$. ∎

Exercise 14.6. V_ω has been defined twice. Let A be the version of V_ω as defined in Definition 7.5, that is let A consist of all hereditarily finite sets. Let $B = \bigcup \{V_n \mid n < \omega\} = \{x \in V \mid \rho(x) < \omega\}$ be the version defined here. Show that both definitions coincide, that is, show $A \subseteq B \wedge B \subseteq A$.

Show that B contains $\emptyset$, is closed under unions of two sets and is closed under the operation forming $\{v\}$ from v. Thus, by Theorem 7.9, $A \subseteq B$.

Show by induction that all members of V_n with $n < \omega$ are hereditarily finite. Thus $B \subseteq A$.

Proposition 14.7. *The definition of the function F from Theorem 14.5 can be extended to all $x \in V$ by the condition*

$$F(x) = \bigcup \{\mathcal{P}(F(y)) \mid y \in x\}.$$

For all $x \in V$, $F(x) = F(\rho(x))$.

Proof. This is proven by transfinite induction. So for any given $x \in V$, one has to show that $F(x) = F(\rho(x))$ provided that $F(y) = F(\rho(y))$ for all $y \in x$.

If $x = \emptyset$ this directly follows from $\rho(0) = 0$. So consider the case that x is nonempty. From the definition and the inductive hypothesis one has that $F(x) =$

$\bigcup\{\mathcal{P}(F(y)) \mid y \in x\} = \bigcup\{\mathcal{P}(F(\rho(y))) \mid y \in x\}$. Note that $F(\alpha) \subseteq F(\beta)$ and $\mathcal{P}(F(\alpha)) \subseteq \mathcal{P}(F(\beta))$ whenever α, β are ordinals with $\alpha \leq \beta$. Furthermore, $\alpha < \rho(x)$ iff there is $y \in x$ with $\alpha \leq \rho(y)$. So one can add $\mathcal{P}(F(\alpha))$ to the union for all $\alpha < \rho(x)$ without changing the outcome: $F(x) = \bigcup\{\mathcal{P}(F(\alpha)) \mid \alpha < \rho(x)\}$. It follows from Theorem 14.5 that $F(x) = F(\rho(x))$. ∎

15 Arithmetic on Ordinals

Addition and multiplication are defined inductively. The first parameter is fixed and the induction goes over the second one. The basic idea of addition of ordinals is that it has an easy geometric interpretation and that it can be reversed: for every ordinals α, β with $\beta > \alpha$ there is a unique ordinal γ with $\alpha + \gamma = \beta$.

Definition 15.1. For ordinals α and β, one can define the addition by transfinite induction: $\alpha + 0 = \alpha$ and, for $\beta > 0$,

$$\alpha + \beta = \sup\{S(\alpha + \gamma) \mid \gamma \in \beta\}.$$

Alternatively, one can also say that $\alpha + \beta$ is the unique ordinal which is order-isomorphic to the set $\{0\} \times \alpha \cup \{1\} \times \beta = \{(0, \gamma) \mid \gamma \in \alpha\} \cup \{(1, \delta) \mid \delta \in \beta\}$ equipped with lexicographic ordering.

Remark 15.2. *Notice that the addition of ordinals is not commutative. For example, $\omega + 1 \neq 1 + \omega = \omega$. Furthermore, if $\alpha > \beta$, one can define $\alpha - \beta$ to be the unique ordinal γ with $\beta + \gamma = \alpha$. This ordinal is the one which is isomorphic to the well-ordered set $(\{\delta \in \alpha \mid \delta \notin \beta\}, <)$. That is, arithmetic and set-theoretic difference coincide up to isomorphism for ordinals.*

Definition 15.3. Multiplication can also be defined by transfinite recursion: $\alpha \cdot 0 = 0$ and, for $\beta > 0$, $\alpha \cdot \beta = \sup\{(\alpha \cdot \gamma) + \alpha \mid \gamma \in \beta\}$.

Alternatively, one can define $\alpha \cdot \beta$ to be the unique ordinal isomorphic to the set $\beta \times \alpha$ equipped with the lexicographic ordering.

Again, the multiplication of ordinals is not commutative. For example,

$$\omega \cdot 2 = \omega + \omega \neq \omega = 2 \cdot \omega.$$

Definition 15.4. $\alpha^0 = 1$, $\alpha^1 = \alpha$, $\alpha^2 = \alpha \cdot \alpha$, $\alpha^3 = \alpha^2 \cdot \alpha$ and $\alpha^{S(n)} = \alpha^n \cdot \alpha$.

Definition 15.5. Let C_{fin} be the class of all functions F which map ordinals to natural numbers with the additional constraint that $F(\alpha) = 0$ for all but finitely many ordinals α. For F, G define that $F < G$ iff $F \neq G$ and $F(\alpha) < G(\alpha)$ for the largest ordinal with $F(\alpha) \neq G(\alpha)$. Map an ordinal α to that function $F \in C_{fin}$ for which $\{G \in C_{fin} \mid G < F\}$ is order isomorphic to α; this function is denoted by F_α from now on. Define an addition $\oplus$ on the ordinals by letting $\alpha \oplus \beta$ be that ordinal γ for which the equation

$$\forall \delta\, (F_\alpha(\delta) + F_\beta(\delta) = F_\gamma(\delta))$$

holds. So the idea is to make an isomorphism between the ordinals and the ordered free commutative semigroup over them.

Exercise 15.6. Which of the following statements are true and which are false?

1. The addition $\oplus$ is commutative.
2. There are ordinals α, β such that $\alpha + \beta$ and $\beta + \alpha$ both differ from $\alpha \oplus \beta$.
3. There are ordinals α, β such that $\alpha + \beta > \alpha \oplus \beta$.
4. There are ordinals α, β such that $\alpha < \beta$ and $\alpha \oplus \gamma \neq \beta$ for all ordinals γ.
5. There are ordinals α, β such that $\alpha < \beta$ and $\alpha + \gamma \neq \beta$ for all ordinals γ.

Remark 15.7. *Given $F \in C_{fin}$, let n be the number of ordinals which F does not map to 0 and let $\alpha_{n-1}, \alpha_{n-2}, \ldots, \alpha_1, \alpha_0$ be these ordinals in descending order. Furthermore, let G_α be the function mapping α to 1 and all other ordinals to 0. Then*

$$F = G_{\alpha_{n-1}} \cdot F(\alpha_{n-1}) + G_{\alpha_{n-2}} \cdot F(\alpha_{n-2}) + \ldots + G_{\alpha_1} \cdot F(\alpha_1) + G_{\alpha_0} \cdot F(\alpha_0).$$

Definition 15.8. Let ω^α denote the ordinal represented by G_α. Given any ordinal $\beta > 0$, consider the function $F_\beta \in C_{fin}$ which is isomorphic to β. Let n be the number of places where F_β is not 0 and let the ordinals $\alpha_{n-1}, \alpha_{n-2}, \ldots, \alpha_1, \alpha_0$ be these n places. Let $m_k = F_\beta(\alpha_k)$ for all $k \in n$. Then

$$\beta = \omega^{\alpha_{n-1}} \cdot m_{n-1} + \omega^{\alpha_{n-2}} \cdot m_{n-2} + \ldots + \omega^{\alpha_1} \cdot m_1 + \omega^{\alpha_0} \cdot m_0.$$

This unique representation is called the *Cantor Normal Form of* β. The *Cantor Normal Form of* 0 is the void sum.

Proposition 15.9. *$\omega^0 = 1$ and $\omega^\alpha = \sup\{\omega^\beta \cdot m \mid \beta < \alpha \wedge m \in \mathbb{N}\}$ for ordinals $\alpha > 0$. In particular, Definitions 15.4 and 15.8 coincide for ω^n with $n \in \mathbb{N}$.*

Proof. This proposition uses Definition 15.8 and the equivalence to Definition 15.4 is established in the last paragraph for the case $n \in \mathbb{N}$. Since 0 is represented by the void sum, ω^0 is greater than 0 and takes the next value 1.

Clearly $\omega^\alpha > \omega^\beta \cdot m$ for all $\beta \in \alpha$ and $m \in \mathbb{N}$. On the other hand, let γ be an ordinal with $1 \leq \gamma < \omega^\alpha$. γ can be represented as

$$\gamma = \omega^{\alpha_{n-1}} \cdot m_{n-1} + \omega^{\alpha_{n-2}} \cdot m_{n-2} + \ldots + \omega^{\alpha_1} \cdot m_1 + \omega^{\alpha_0} \cdot m_0.$$

where $n \in \mathbb{N} - \{0\}$ and $m_k > 0$ for all $k \in n$. Recall that F_γ is the function in C_{fin} representing γ and G_α represents ω^α. Let δ be the largest ordinal with $F_\gamma(\delta) \neq G_\alpha(\delta)$. Since $F_\gamma < G_\alpha$, $F_\gamma(\delta) < G_\alpha(\delta) = 1$ and $\delta = \alpha$. Thus $F_\gamma(\delta) = 0$ and $\delta > \alpha_{n-1} > \alpha_{n-2} > \ldots > \alpha_1 > \alpha_0$. Now let $\beta = \alpha_{n-1}$ and $m = \sum_{k \in n} m_k$. Then $F_\gamma \leq G_\beta \cdot m < G_\alpha$ and $\gamma \leq \omega^\beta \cdot m < \omega^\alpha$. Thus ω^α is indeed the supremum of all $\omega^\beta \cdot m$ with $\beta \in \alpha$ and $m \in \mathbb{N}$.

Note that the equality $\omega^0 = 1$ coincides with Definition 15.4. Assume now that the equivalence is established for some $n \in \mathbb{N}$. Then, using the definition of $\omega^{S(n)}$ and of the multiplication, one has that $\omega^{S(n)} = \sup_{m \in \mathbb{N}} \omega^n \cdot m$. Since the sequence $\omega^0, \omega^1, \ldots, \omega^n$ is increasing, $\omega^{S(n)}$ is also by Definition 15.4 the supremum of all $\omega^k \cdot m$ with $k \in S(n)$ and $m \in \mathbb{N}$, thus the equivalence of both definitions transfers to $S(n)$ and the last statement of the proposition follows by induction. ∎

Example 15.10. One can view the Cantor Normal Form as a finite sum of powers of ω in descending order as in this example:

$$\omega^5 + \omega^2 + \omega^2 + \omega^2 + \omega^1 + \omega^0 + \omega^0 \; = \; \omega^5 + \omega^2 \cdot 3 + \omega + 2.$$

Instead of repeating same ordinals, one can also multiply them with the corresponding natural number, instead of ω^1, one can write just ω, instead of ω^0 just 1. The void sum is represented by the symbol 0. This all is done on the right hand side of the equation above. Also transfinite ordinals can be in the power:

$$\omega^{\omega^2} + \omega^{\omega \cdot 5+8} \cdot 7 + \omega^{\omega \cdot 5+7} \cdot 12345 + \omega^{22222} \cdot 33333 + \omega^4 + \omega^3 + \omega^2 + \omega + 1.$$

If one adds ordinals $\omega^\alpha \cdot a + \omega^\beta \cdot b$ with $\alpha < \beta$, then $\omega^\alpha \cdot a$ can be omitted; if $\alpha = \beta$ the coefficients can be added giving $\omega^\alpha \cdot (a + b)$; if $\alpha > \beta$, no simplification is possible:

$$\begin{aligned} \omega^3 + \omega^5 &= \omega^5; \\ \omega^3 \cdot 5 + \omega^4 \cdot 8 + \omega^5 \cdot 0 &= \omega^4 \cdot 8; \\ \omega^3 \cdot 234 + \omega^3 \cdot 111 &= \omega^3 \cdot 345; \\ \omega^5 + \omega^3 + \omega^2 + \omega^3 + \omega^1 &= \omega^5 + \omega^3 \cdot 2 + \omega. \end{aligned}$$

The last line has the application of two rules: first ω^2 is omitted as it is in front of a higher ω-power; second the two ω^3-terms are unified to one; no further simplification is possible.

The Cantor Normal Form can also be used in order to express the rank of sets. Recall the following rules:

- The rank of an ordinal α is α. So $\rho(0) = 0, \rho(1) = 1, \rho(\omega) = \omega$.
- The rank of sets is determined by the rank of their elements. For example, $\rho(\{x, y, z\}) = \max\{\rho(x)+1, \rho(y)+1, \rho(z)+1\}$ and $\rho(\{\omega, \omega\cdot 2, \omega+5\}) = \omega\cdot 2+1$.
- In general, $\rho(X) = \sup\{\rho(Y) + 1 \mid Y \in X\}$.

Then, one can get the following ranks expressed in Cantor Normal Form:

$$\begin{array}{rcl}
\rho(\{\{\emptyset\}\}) & = & 2; \\
\rho(\{\emptyset, \{\emptyset\}\}) & = & 2; \\
\rho(\{\{\{\{\omega\}\}\}\}) & = & \omega + 4; \\
\rho(\{\omega^4 + 2, \omega^3 \cdot 8\}) & = & \omega^4 + 3; \\
\rho(\{\omega^\alpha \mid \alpha < \omega^2\}) & = & \omega^{\omega^2}; \\
\rho(\{\omega^\alpha + \omega^\beta \mid \alpha, \beta < \omega + 8\}) & = & \omega^{\omega+7} \cdot 2 + 1.
\end{array}$$

The Cantor Normal Form is in particular useful to denote ordinals formed by finite sums over small powers of ω.

Exercise 15.11. Determine the Cantor Normal Form of the following ordinals.

1. $\omega + \omega^2 + \omega^3 + \omega^4 + 2$,
2. $(\omega + 3)^5 + (\omega^2 + 17) \cdot (\omega + 8) + \omega^{12}$,
3. $\omega^2 + \omega + 1 + \omega^2 + \omega + 1 + \omega^2 + \omega + 1$,
4. $1 \oplus \omega \oplus \omega^2 \oplus \omega^3$,
5. $\omega^{\omega+5} + \omega^{\omega+2} \cdot \omega + \omega^2$,
6. $256^{256} + \omega \cdot 42$.

Exercise 15.12. Assume that $\alpha = \omega^{\gamma_1} + \omega^{\gamma_2}$ and $\beta = \omega^{\delta_1} + \omega^{\delta_2}$ with $\gamma_1 > \gamma_2$ and $\delta_1 > \delta_2$. What condition on $\gamma_1, \gamma_2, \delta_1, \delta_2$ is equivalent to the equation $\alpha + \beta = \alpha \oplus \beta$.

Example 15.13. An ϵ-number is an ordinal ϵ satisfying $\epsilon = \omega^\epsilon$. In particular, for any ordinal α, ϵ_α is the first ordinal such that the set

$$\{\epsilon : \epsilon \leq \epsilon_\alpha \wedge \omega^\epsilon = \epsilon\}$$

is isomorphic to $\{\beta : \beta \leq \alpha\}$. The ordinal ϵ_0 is the limit of the sequence

$$\omega, \omega^\omega, \omega^{(\omega^\omega)}, \omega^{(\omega^{(\omega^\omega)})}, \ldots$$

of iterated powers of ω; the brackets (and) are normally omitted. In particular, $\alpha < \epsilon_0$ iff α can be expressed by a formula consisting of the constants $0, 1$, the power ω^β for a subformula β and addition. For example,

$$\omega^{\omega^\omega} + \omega^2 + \omega^2 + \omega^2 + 1 + 1$$

is such an expression. It is of course convenient to write $\omega^{\omega^\omega} + \omega^2 \cdot 3 + 2$ instead.

Example 15.14. An ordinal α is called *constructive* or *recursive* iff $\alpha < \omega$ or there is a relation $\sqsubset$ on $\mathbb{N} \times \mathbb{N}$ which can be computed by a computer programme such that $(\mathbb{N}, \sqsubset)$ is isomorphic to α.

The ordinal ω^ω is constructive. There is a one-to-one enumeration of all polynomials $p_0, p_1, \ldots$ in ω where the coefficients are natural numbers. Now let $n \sqsubset m$ iff the polynomials p_n, p_m are different and satisfy $a_n < a_m$ for the coefficients a_n in p_n and a_m in p_m for the largest power ω^k where these coefficients are different. So $\omega^5{+}\omega^3{\cdot}2 \sqsupset \omega^5{+}\omega^3 \sqsupset \omega^5{+}\omega^2 \sqsupset \omega^4{\cdot}17{+}\omega^3 \sqsupset \omega^4{\cdot}17 \sqsupset \omega^2{+}1 \sqsupset \omega \sqsupset 1243134123412342$. An equivalent definition is that $p_n \sqsubset p_m$ iff $p_n(x) < p_m(x)$ for almost all natural numbers x when x replaces ω viewed upon as a formal variable in the polynomials p_n, p_m.

Other examples of constructive ordinals are $\epsilon_0, \epsilon_1, \ldots, \epsilon_\omega, \epsilon_{\omega+1}$.

The first non-constructive ordinal is ω_{CK} named after the mathematicians Church and Kleene who studied this ordinal.

Example 15.15. There is a first ordinal ω_1 such that the set $\{\alpha : \alpha < \omega_1\}$ representing ω_1 is not countable. It is larger than all previously considered ordinals. For example, $121234312 < \omega < \omega^\omega + 2341234431124123 < \epsilon_0 < \epsilon_1 < \epsilon_2 < \epsilon_5 + \omega^7 < \omega_{CK} < \omega_{CK} + \omega \cdot 17 + 4 < \omega_1$.

16 Cardinals

There are two different usages of the natural numbers: First to denote the quantity of something, say "three" represents the set {apple, banana, pear} of three fruits. Second to enduce an order, so the third word "pear" comes after the second word "banana". The English language reflects these two ways to use numbers by having the different words "three" and "third". When dealing with infinite objects, it is even more necessary to distinguish cardinals (representing the quantity) and ordinals (representing an order): object number $\omega+8$ should come after object $\omega+7$ and not before it. But the cardinality of the sets represented by these two cardinals is the same, since f given as

$$f(\alpha) = \begin{cases} S(\alpha) & \text{if } \alpha \in \omega; \\ 0 & \text{if } \alpha = \omega + 7; \\ \alpha & \text{if } \omega \subseteq \alpha \wedge \alpha \in \omega + 7; \end{cases}$$

is a bijective function from $\omega+8$ to $\omega+7$. So one would want to assign to $\omega+8$ and $\omega+7$ the same cardinal. This is done by defining that the cardinal α of a set A is the least ordinal such that there is a bijective mapping from A into α; note that the Axiom of Choice defined below is required to guarantee that every set has a cardinal.

Definition 16.1. An ordinal α is a *cardinal* (or cardinal number) if $|\beta| < |\alpha|$ for all $\beta \in \alpha$. A cardinal α is called *the cardinal number* (or sometimes *the cardinality*) of A, denoted by $\alpha = |A|$, if and only if $|\alpha| = |A|$.

Example 16.2. *Let α be an ordinal.*

1. *If α is a cardinal then $\alpha = |\alpha|$.*
2. *If $\alpha \leq \omega$ then α is a cardinal. ω is the least infinite cardinal.*
3. *$\omega+1, \omega+17, \omega^2, \omega^{51}$ are not cardinals. In particular, if α is countable and $\alpha > \omega$ then α is not a cardinal.*

Theorem 16.3. *For every ordinal β there is a unique cardinal α such that $\alpha = |\beta|$ and $\alpha \leq \beta$. Furthermore, if A is well-orderable then there is a unique cardinal α such that $\alpha = |A|$.*

Proof. The set $\{\gamma \in S(\beta) \mid |\gamma| = |\beta|\}$ has a minimum α. Then $|\alpha| = |\beta|$ but $|\gamma| < |\beta|$ for all $\gamma < \alpha$. It follows that α is a cardinal, that is, $\alpha = |\beta|$.

Given a well-ordered set $(A, \sqsubset)$, there is by Theorem 13.9 a unique ordinal β representing $(A, \sqsubset)$ in the sense that $(A, \sqsubset)$ is isomorphic to $(\beta, <)$. So there is a cardinal α such that $\alpha = |\beta|$ and therefore also $\alpha = |A|$. ∎

Given an ordinal β, there is by Theorem 13.10 an ordinal γ with $|\gamma| \not\leq |\beta|$. Now let α be the minimum of all $\delta \in S(\gamma)$ such that $|\delta| \not\leq |\beta|$. Then α is the first cardinal larger than β.

Property 16.4. *For every ordinal β there is a first cardinal α such that $|\beta| < \alpha$. α is denoted as β^+.*

Definition 16.5. For an ordinal β, the least cardinal α satisfying that $\beta \in \alpha$ is denoted by β^+. A cardinal α is called a *successor cardinal* if $\alpha = \beta^+$ for some ordinal β. A cardinal is called a *limit cardinal* if it is not a successor cardinal. Furthermore, cardinals are denoted by *alephs*: $\aleph_0 = \omega$ and

$$\aleph_\alpha = \sup\{\aleph_\beta^+ \mid \beta \in \alpha\}$$

for ordinals $\alpha > 0$. Although cardinals are identified with the ordinals representing them, there is still the traditional name ω_α for the least ordinal β satisfying $|\beta| = \aleph_\alpha$.

Example 16.6. *So $\omega_1 = \omega_0^+$, $\omega_2 = \omega_1^+$, $\omega_3 = \omega_2^+$ and $\omega_\omega \neq \alpha^+$ for all ordinals α.*

Due to the identification of sets and cardinals with ordinals, the natural numbers and their cardinality can all be denoted by the following symbols: $\mathbb{N}$, ω, ω_0, $\aleph_0$. Similarly, ω_1, ω^+, ω_0^+, $\aleph_1$, $\aleph_0^+$ are all names for the first uncountable ordinal which is identified with the set representing it and its cardinal.

All ω_α are limit ordinals, but $\aleph_\alpha$ is a limit cardinal only if α is a limit ordinal; $\aleph_\alpha$ is a successor cardinal otherwise.

The addition and multiplication of cardinals is different from that of ordinals since one enforces that the result is a cardinal. So $\aleph_\alpha + 1$ will be different from both $\omega_\alpha + 1$ (obtained by looking on $\aleph_\alpha$ as an ordinal) and $\aleph_{\alpha+1} = \aleph_\alpha^+$ (obtained by adding the indices).

Definition 16.7 (Arithmetic for Cardinals). For cardinals κ and λ, define $\kappa+\lambda = |\kappa + \lambda|$, $\kappa \cdot \lambda = |\kappa \times \lambda|$ and $2^\kappa = |\mathcal{P}(\kappa)|$.

In the following it will be proven that the addition and the multiplication of infinite cardinals are really trivial and coincide with forming the maximum.

Proposition 16.8 (Hessenberg). *If κ is an infinite cardinal and λ a cardinal with $\lambda \leq \kappa$ then $\kappa + \lambda = \lambda + \kappa = \kappa$.*

Proof. For every infinite ordinal α, one has $|\alpha| = |S(\alpha)|$ witnessed by the bijection f defined as

$$f(\beta) = \begin{cases} 0 & \text{if } \beta = \alpha; \\ \beta & \text{if } \omega \leq \beta < \alpha; \\ \beta + 1 & \text{if } \beta < \omega. \end{cases}$$

So $S(\alpha)$ cannot be a cardinal and κ is a limit ordinal. Note that $\kappa \leq |\kappa + \lambda| \leq |\kappa \times \{0,1\}|$. So it is sufficient to show that $|\kappa| = |\kappa \times \{0,1\}|$ and using the fact that κ is an infinite limit ordinal. This is witnessed by the following function g:

$$g(\omega \cdot \gamma + n, a) = \omega \cdot \gamma + 2n + a$$

for all ordinals γ and $n \in \mathbb{N}$ such that $\omega \cdot \gamma + n \in \kappa$. Since every $(\beta, a) \in \kappa \times \{0,1\}$ can be uniquely represented as $(\omega \cdot \gamma + n, a)$ with γ being an ordinal, $n \in \mathbb{N}$ and $a \in \{0,1\}$, the function g is well-defined. Furthermore, it is easy to see that g is a bijection. ∎

Exercise 16.9. Construct a one-to-one function h which maps $\alpha \times \omega$ to α for any infinite limit ordinal α. This function can without loss of generality assume that the input is of the form $(\omega \cdot \gamma + n, m)$ where $m, n \in \mathbb{N}$ and γ is an ordinal with $\omega \cdot S(\gamma) \leq \alpha$; the image should be of the form $\omega \cdot \gamma + h(n,m)$.

In order to see that $\kappa \times \kappa$ is the same as κ, recall the canonical well-ordering from Example 11.5.

Definition 16.10. For an infinite ordinal κ, the *canonical well-ordering* of $\kappa \times \kappa$, denoted by $<_{cw}$, is defined as follows for $(\alpha, \beta), (\gamma, \delta) \in \kappa \times \kappa$:

$$\begin{aligned}(\alpha, \beta) <_{cw} (\gamma, \delta) \;\Leftrightarrow\;& \max\{\alpha, \beta\} < \max\{\gamma, \delta\} \\ \vee\;& (\max\{\alpha, \beta\} = \max\{\gamma, \delta\} \wedge \alpha < \gamma) \\ \vee\;& (\max\{\alpha, \beta\} = \max\{\gamma, \delta\} \wedge \alpha = \gamma \wedge \beta < \delta).\end{aligned}$$

Theorem 16.11. *For all infinite cardinals κ, $(\kappa \times \kappa, <_{cw}) \cong (\kappa, \in)$.*

Proof. The mapping $(\alpha, \beta) \mapsto \omega^{\max\{\alpha,\beta\}\cdot 2+2} + \omega^{\max\{\alpha,\beta\}+\alpha+1} + \omega^{\beta}$ is an isomorphism from $(\kappa \times \kappa, <_{cw})$ to some subset of $(\omega^{\kappa \cdot 2+3}, \in)$. Thus $(\kappa \times \kappa, <_{cw})$ is a well-ordered set.

So it remains to show that the two well-ordered sets are isomorphic. Remark 11.6 states that it is true for $\kappa = \omega$. Assume that there is a counterexample, say κ.

Let $A = \{\lambda \in S(\kappa) \mid \lambda \geq \omega \text{ is a cardinal and } (\lambda \times \lambda, <_{cw}) \not\cong (\lambda, \in)\}$. Then $\kappa \in A$. Let $\mu \in A$ be the least element of A. Note that $\mu > \omega$. Furthermore, for all λ, if $\omega \leq \lambda < \mu$ and λ is a cardinal, then $(\lambda \times \lambda, <_{cw}) \cong (\lambda, \in)$.

By the Comparability Theorem, $(\mu, \in)$ is isomorphic to an initial segment of $(\mu \times \mu, <_{cw})$ since $|\mu \times \mu| \geq \mu$ and μ is a cardinal. Let $(\alpha, \beta) \in \mu \times \mu$ be such that $(\mu, \in)$ is isomorphic to the initial segment of $(\mu \times \mu, <_{cw})$ given by (α, β). Let h be the isomorphism. Let $\eta = \max\{\alpha, \beta\}$. Then $(\alpha, \beta) \leq_{cw} (\eta, \eta)$. Hence, $h : \mu \to \eta \times \eta$ is injective. Let $\lambda = |\eta|$. Then $|\lambda \times \lambda| = |\eta \times \eta|$. But $|\lambda \times \lambda| = \lambda < \mu$. This is a contradiction. Hence, there is no counterexample to the theorem. ∎

Theorem 16.12 (Hessenberg). *If κ, λ are cardinals with $\aleph_0 \leq \kappa$ and $1 \leq \lambda \leq \kappa$ then $\kappa \cdot \lambda = \lambda \cdot \kappa = \kappa$.*

Remark 16.13. Let κ be a cardinal. Recall that $2^\kappa = |\mathcal{P}(\kappa)|$ by Definition 16.7 and $2^\kappa > \kappa$ by Theorem 6.12. Thus $2^\kappa \geq \kappa^+$. Note that $2^0 = 1 = 0^+$, $2^1 = 2 = 1^+$ and $2^4 = 16 > 5 = 4^+$.

17 The Axiom of Choice

If a set is not empty, one can find an element in it. Somehow, it is not guaranteed that one can find the element in a systematic way, that is, by a function. This is formalized by the Axiom of Choice.

Definition 17.1. Let X be a set of sets. A function $\mathcal{C}$ defined on all nonempty members of X is called *a choice function* of X if $\mathcal{C}(x) \in x$ for every nonempty $x \in X$. This permits to state the Axiom of Choice and its countable counterpart as follows.

Definition 17.2 (Axiom of Choice). *Let X be a set of sets. Then X has a choice function.*

Example 17.3. *A choice function $\mathcal{C}$ on $\mathbb{N}$ can be defined as $\mathcal{C}(S(n)) = n$ for all $n \in \mathbb{N}$.*

Example 17.4. *If $(W, \sqsubset)$ is a well-orderable set and let $X = \mathcal{P}(W)$. Then the function which assigns to every nonempty subset of W its minimum with respect to $\sqsubset$ is a choice function.*

Theorem 17.5. *Assuming all axioms except the Axiom of Choice, the following conditions are equivalent:*

1. *The Axiom of Choice.*
2. *Every set can be one-to-one mapped into a set of ordinals.*
3. *Every set is well-orderable.*
4. *For all sets X, Y, either $|X| < |Y|$ or $|X| = |Y|$ or $|Y| < |X|$.*

Proof. *First Statement $\Rightarrow$ Second Statement.* Let X be any given set and $u \notin X$ a target which will be used to guarantee that the below mapping is invertible on X. By Theorem 13.10 there is an ordinal α such that $|\alpha| \not\leq |X|$. Now one constructs by transfinite induction the following $f : \alpha \to X \cup \{u\}$ where $\mathcal{C}$ is a choice function which is defined at least on all subsets of X. For every $\gamma \in \alpha$ one defines

$$f(\gamma) = \begin{cases} \mathcal{C}(X - f[\gamma]) & \text{if } f[\gamma] \subset X; \\ u & \text{otherwise.} \end{cases}$$

Note that whenever $f(\gamma) \in X$ then $f(\gamma) \notin f[\gamma]$ and thus f does not take any elements of X twice. Since $|\alpha| \not\leq |X|$ there must be some ordinal in α which is mapped to u. Let β be the least such ordinal. Then $f : \beta \to X$ is a bijection and has an inverse one-to-one function g which maps X into a set of ordinals.

Second Statement $\Rightarrow$ *Third Statement.* If $g : X \to Y$ is a one-to-one function and Y is a set of ordinals, then g induces a well-ordering of X: for all $x, y \in X$, $x \sqsubset y \Leftrightarrow g(x) \in g(y)$.

Third Statement $\Rightarrow$ *First Statement.* Let a set X of sets be given. There is a well-ordering $\sqsubset$ on $\bigcup X$. Now one can define a choice function $\mathcal{C}$ which maps every nonempty subset Y of $\bigcup X$ to its minimum with respect to $\sqsubset$. Hence $\mathcal{C}$ also maps every nonempty $Y \in X$ to its minimum with respect to $\sqsubset$. Hence X has a choice function.

Third Statement $\Rightarrow$ *Fourth Statement.* Let X, Y be sets and assume that $|X| \not\leq |Y|$. There are well-orderings on X, Y and by Theorem 11.13 these sets are either order-isomorphic or one is order-isomorphic to some initial segment of the other one. Since $|X| \not\leq |Y|$, Y is order-isomorphic to an initial segment of X and the corresponding mapping is one-to-one. Thus $|Y| < |X|$.

Fourth Statement $\Rightarrow$ *Second Statement.* Given a set X, there is by Theorem 13.10 there is an ordinal α such that $|\alpha| \not\leq |X|$. Since α, X are comparable, $|X| < |\alpha|$. Thus there is a one-to-one mapping from X into α. ▮

The next two results are applications of the Axiom of Choice. They are based on the fact that every set X there is an ordinal α and a bijection $f : \alpha \to X$. Then there is a cardinal $\kappa \leq \alpha$ with $\kappa = |\alpha|$. Furthermore, for every cardinal $\lambda \leq \kappa$, $f[\lambda]$ is a subset of X of cardinality λ.

Theorem 17.6. *For every set X, there is a unique cardinal κ such that $\kappa = |X|$. Furthermore, if X is infinite, then X has a countable subset.*

Exercise 17.7. Let A, B, C be any sets and, as in Example 3.16,

$$D = \{f \in C^A \mid \exists g \in B^A \exists h \in C^B \, (f = h \circ g)\}.$$

Show that $D = C^A$ iff $|B| \geq \min\{|A|, |C|\}$.

Theorem 17.8. *If f is a function defined on A, then $|f[A]| \leq |A|$.*

Proof. Let $\mathcal{C}$ be a choice function on all nonempty subsets of A. Now define for all $b \in f[A]$ the mapping g by $g(b) = \mathcal{C}(\{a \in A \mid f(a) = b\})$. The function g is one-to-one and witnesses $|f[A]| \leq |A|$. ∎

Using the Axiom of Choice, one can prove the following result.

Theorem 17.9. *The union of a countable set of countable sets is countable.*

Proof. Let A be a countable set of countable sets, that is, every $B \in A$ is countable. There is surjective function $F : \mathbb{N} \to A$, $F(n)$ is the $n+1$-st set contained in A. For each n, let $E(n) = \{f : \mathbb{N} \to F(n) \mid f \text{ is surjective}\}$. Note that E is a function from $\mathbb{N}$ to $(\bigcup A)^{\mathbb{N}}$, each set $E(n)$ has cardinality $2^{\aleph_0}$. By the Axiom of Choice, there is a function g which selects from every $E(n)$ an element g_n of this set. Now let $G(n,m) = g_n(m)$. G is a surjective mapping from $\mathbb{N} \times \mathbb{N}$ to $\bigcup A$, thus $\bigcup A$ is at most countable. Since A is not empty, there is a countable and thus infinite $B \in A$ and by $B \subseteq \bigcup A$, the set $\bigcup A$ is infinite. So $\bigcup A$ is countable. ∎

Corollary 17.10. *The first uncountable ordinal ω_1 is not the union of a countable set of countable ordinals.*

The Axiom of Choice can be used to construct an example of a set of cardinality $\aleph_1$.

Example 17.11. *Define for $A, B \subseteq \mathbb{N}$ the following relations:*

$$
\begin{aligned}
A \leq_{lin} B &\iff \exists m, n \in \mathbb{N}\, \forall a \in \mathbb{N}\, (a \in A \Leftrightarrow a \cdot m + n \in B);\\
A <_{lin} B &\iff A \leq_{lin} B \wedge B \not\leq_{lin} A.
\end{aligned}
$$

Let $L \subseteq \mathcal{P}(\mathbb{N})$ be such that L is not empty and $(L, <_{lin})$ is a linearly ordered set. If L is bounded by some $A \subseteq \mathbb{N}$ in the sense that $\forall B \in L\, (B \leq_{lin} A)$ then $|L| \leq \aleph_0$ else $|L| = \aleph_1$. Furthermore L can indeed be chosen such that $(L, <_{lin})$ is linearly ordered and $|L| = \aleph_1$; so the else-case is not only a theoretical case.

Proof. Note that $\leq_{lin}$ is transitive: If $A \leq_{lin} B$ and $B \leq_{lin} C$ then there are $m, n, i, j \in \mathbb{N}$ such that for all $a, b \in \mathbb{N}$, $a \in A \Leftrightarrow a \cdot m + n \in B$ and $b \in B \Leftrightarrow b \cdot i + j \in C$. Thus for all $a \in \mathbb{N}$, $a \in A \Leftrightarrow a \cdot (m \cdot i) + (n \cdot i + j) \in C$. Note that $\leq_{lin}$ is not antisymmetric: if A is the set of even and B of odd numbers then $a \in A \Leftrightarrow a + 1 \in B$ and $b \in B \Leftrightarrow b + 1 \in A$. But by Remark 9.11, $<_{lin}$ is defined from the transitive relation $\leq_{lin}$ such that it is automatically transitive and antireflexive, so $(\mathcal{P}(\mathbb{N}), <_{lin})$ is a partially ordered set.

If L is bounded by A then one has for every $B \in L$ a pair (m_B, n_B) such that $\forall b \in B \Leftrightarrow b \cdot m_B + n_B \in A$. If $C \neq B$ and $C \leq_{lin} A$ then $(m_C, n_C) \neq (m_B, n_B)$. Thus

L is at most countable.

If L is at most countable then there is a surjective function F from $\mathbb{N}$ to L, that is, $L = \{F(0), F(1), \ldots\}$. Now let

$$A = \{(2b+1)2^a \mid b \in F(a)\}.$$

It is easy to see that $F(a) \leq_{lin} A$ by $\forall b \in \mathbb{N}\,(b \in F(a) \Leftrightarrow b \cdot 2^{a+1} + 2^a \in A)$.

So let L be unbounded. L is uncountable by the previous paragraphs. Now functions $f : \omega_1 \to \mathcal{P}(\mathbb{N})$ and $g : \omega_1 \times \omega \times \omega \to \mathcal{P}(\mathbb{N})$ with $f[\omega_1] \subseteq L \subseteq g[\omega_1, \omega, \omega]$ in order to witness that $|L| = \aleph_1$. The construction uses transfinite recursion and a choice function $\mathcal{C}$ defined on all nonempty subsets of $\mathcal{P}(\mathbb{N})$.

So, for given α, assume that $f(\beta)$ and $g(\beta, i, j)$ are defined for all $\beta \in \alpha$ and $i, j \in \mathbb{N}$. Now let

$$\begin{aligned} f(\alpha) &= \mathcal{C}(L - g[\alpha, \omega, \omega]); \\ g(\alpha, i, j) &= \{a \mid a \cdot i + j \in f(\alpha)\}. \end{aligned}$$

The resulting functions f, g are then defined on the whole sets ω_1 and $\omega_1 \times \omega \times \omega$, respectively. Now the following properties hold.

- For all $\alpha \in \omega_1$ is the set $g[\alpha, \omega, \omega] = \{g(\beta, i, j) \mid \beta \in \alpha, i, j \in \omega\}$ at most countable. Hence $L - g[\alpha, \omega, \omega]$ is not empty and $f(\alpha) = C(L - g[\alpha, \omega, \omega])$ is an element of L outside the set $g[\alpha, \omega, \omega]$.

- For all $\alpha \in \omega_1$ and all $A \subseteq \mathbb{N}$, $A <_{lin} f(\alpha) \Leftrightarrow \exists i, j \in \mathbb{N}\,(A = g(\alpha, i, j))$. In particular, $f(\alpha) = g(\alpha, 1, 0)$.

- If $\alpha, \beta \in \omega_1$ with $\beta < \alpha$ then $f(\beta) <_{lin} f(\alpha)$. The reason is that all sets $A \leq_{lin} f(\beta)$ are in $g[\alpha, \omega, \omega]$ and thus $f(\alpha) = C(L - g[\alpha, \omega, \omega]) \not\leq_{lin} f(\beta)$. As $(L, <_{lin})$ is linearly ordered, $f(\beta) <_{lin} f(\alpha)$.

It follows that $(f[\omega_1], <_{lin})$ is a linearly ordered set isomorphic to $(\omega_1, <_{lin})$ and $f[\omega_1] \subseteq L$. The set $f[\omega_1]$ has cardinality $\aleph_1$. Let $A \in L$. Then there is some $B \in f[\omega_1]$ with $B \not\leq_{lin} A$. Hence $A <_{lin} B$. Furthermore, there is an $\alpha \in \omega_1$ with $B = f(\alpha)$. It follows that $A = f(\alpha, i, j)$ for some $i, j \in \mathbb{N}$ and $B \in g[\omega_1, \omega, \omega]$. So $f[\omega_1] \subseteq L \subseteq g[\omega_1, \omega, \omega]$. It follows that $|L| = \aleph_1$.

At the end it is shown that there is indeed such an uncountable and unbounded linearly ordered subset of $(\mathcal{P}(\mathbb{N}), <_{lin})$. This is done by constructing an order-preserving mapping h from $(\omega_1, \in)$ into $(\mathcal{P}(\mathbb{N}), <_{lin})$ via transfinite recursion and using the choice function C on $\mathcal{P}(\mathcal{P}(\mathbb{N}))$:

$$h(\alpha) = C(\{A \subseteq \mathbb{N}\, \forall \beta \in \alpha\,(h(\beta) <_{lin} A)\})$$

for all $\alpha \in \omega_1$. It is clear from the construction that the mapping is order preserving, one only has to verify that the set

$$\{A \subseteq \mathbb{N} \,\forall \beta \in \alpha \,(h(\beta) <_{lin} A)\}$$

is not empty for any $\alpha \in \omega_1$. To see this, note that $h[\alpha]$ is at most countable and that there is therefore a set $B \subseteq \mathbb{N}$ such that $B \not\leq_{lin} h(\beta)$ for every $\beta \in \alpha$. It follows that $\{B\} \cup h[\alpha]$ is at most countable. As argued above, there is an $A \subseteq \mathbb{N}$ bounding every set in $\{B\} \cup h[\alpha]$. Then $h(\beta) <_{lin} A$ as $h(\beta) \leq_{lin} A$ and $B \not\leq_{lin} h(\beta)$ but $B \leq_{lin} A$. So there is a proper upper bound of $h[\alpha]$ and the value $h(\alpha)$ is such an upper bound selected by the choice function. The linearly ordered set $(h[\omega_1], <_{lin})$ is order-isomorphic to $(\omega_1, \in)$, has cardinality $\aleph_1$ and is a subset of the partially ordered set $(\mathcal{P}(\mathbb{N}), <_{lin})$. This completes the proof. ∎

Exercise 17.12. Consider the following partial ordering given on the set $\mathbb{N}^{\mathbb{N}}$ of all functions from $\mathbb{N}$ to $\mathbb{N}$:

$$f \sqsubset g \Leftrightarrow \exists n \,\forall m > n \,(f(m) < g(m)).$$

This partial ordering only shares some but not all of the properties of the ordering $<_{lin}$ considered above. In order to see this, show the following two properties:

- For countably many functions $f_0, f_1, \ldots$ there is a function g such that $\forall n \in \mathbb{N}\,(f_n \sqsubset g)$;
- There are uncountably many f below the exponential function $n \mapsto 2^n$. Namely for every $A \subseteq \mathbb{N}$ the function $c_A : n \mapsto \sum_{m \in n} 2^{n-m-1} \cdot A(m)$ is below the exponential function.

Note that $c_A \sqsubset c_B \Leftrightarrow A <_{lex} B$. Thus there is an uncountable linearly ordered set of functions below the exponential function.

Exercise 17.13. Use the Axiom of Choice to prove the following: If $|A| = \aleph_1$ and every $B \in A$ satisfies $|B| \leq \aleph_1$ then $|\bigcup A| \leq \aleph_1$.

18 The Set of Real Numbers

The real numbers are one of the most important topics of mathematics. This section deals with some basic properties of this set. In particular, several ways to represent the set of real numbers are proposed. Other than in the case of the natural numbers, there is no standard convention how to do it. The given representations are build in the standard way using already defined objects like sequences of digits or subsets of the rational numbers. It is convenient to introduce representations of the integers numbers first.

Example 18.1. The set of integers can be represented as that of ordered pairs of natural numbers where one of the parts of the pair is 0:

$$\mathbb{Z} = \{(m, n) \in \mathbb{N} \times \mathbb{N} \mid m = 0 \vee n = 0\}.$$

The pair (m, n) represents the integer normally denoted by $m - n$, so $(10, 0)$ is 10 and $(0, 4)$ is -4. The addition of two integers (i, j) and (k, l) can be defined as follows.

$$(i, j) + (k, l) = (m, n) \Leftrightarrow (m, n) \in \mathbb{Z} \wedge \exists h \in \mathbb{N}\, (i + k = m + h \wedge j + l = n + h)$$

Furthermore, $(i, j) < (k, l)$ if and only if $i + l < j + k$ as natural numbers.

Note that this representation has the disadvantage that it recodes the natural numbers in a nonstandard way, replacing n by $\{n, \{n, 0\}\}$. An alternative approach would be to let the natural numbers unchanged, to represent -1 by $\{\{\emptyset\}\}$ and, for all $n > 0$, $-n - 1$ by $S(-n) = -n \cup \{-n\}$. So -2 would be $\{\{\emptyset\}, \{\{\emptyset\}\}\}$ and -3 would be $\{\{\emptyset\}, \{\{\emptyset\}\}, \{\{\emptyset\}, \{\{\emptyset\}\}\}\}$. The disadvantage of this representation is that the addition and other operations are a bit more difficult to define.

Theorem 18.2. $|\mathbb{R}| = |\mathbb{N}^{\mathbb{N}}|$.

Proof. To see that $|\mathbb{N}^{\mathbb{N}}| \leq |\mathbb{R}|$, define the function $F : \mathbb{N}^{\mathbb{N}} \to \mathbb{R}$ as follows:

$$F(f) = \sum_{n \in \mathbb{N}} 10^{\sum_{m=0,\ldots,n} -S(f(n))}$$

That is, the decimal representation of the number $F(f)$ is $0.0^{f(0)}10^{f(1)}10^{f(2)}10^{f(3)}1\ldots$ and the injectiveness follows from the fact that one can reconstruct f from the representation of its image. For example, $F(f) = 0.1100000001000101\ldots$ iff $f(0) = 0$, $f(1) = 7$, $f(2) = 3$ and $f(3) = 1$. Since one deals with decimal and not with binary representation, the numbers $0.100000\ldots = \frac{1}{10}$ and $0.0111111\ldots = \frac{1}{90}$ are different, so there is no messing up caused by the images of functions which are almost everywhere 0.

For the converse direction, one takes a one-to-one enumeration $q_0, q_1, \ldots$ of $\mathbb{Q}$ and constructs the following one-to-one mapping from $\mathbb{R}$ to $\mathbb{N}^{\mathbb{N}}$:

$$G(r)(n) = \begin{cases} 0 & \text{if } r < q_n; \\ 1 & \text{if } r = q_n; \\ 2 & \text{if } r > q_n. \end{cases}$$

The function G is one-to-one. Let r_1, r_2 be two different real numbers, say $r_1 < r_2$. There is a number n such that $r_1 < q_n < r_2$ since $\mathbb{Q}$ is a dense subset of $\mathbb{R}$. It follows that $G(r_1)(n) = 0$ and $G(r_2)(n) = 2$. Thus $G(r_1)$ and $G(r_2)$ are different members of $\mathbb{N}^{\mathbb{N}}$.

By the Cantor-Bernstein Theorem, it follows from $|\mathbb{R}| \leq |\mathbb{N}^{\mathbb{N}}|$ and $|\mathbb{N}^{\mathbb{N}}| \leq |\mathbb{R}|$ that these two sets have the same cardinality. ▮

In the following, explicit representations are given and the addition and ordering defined on them. Set theorists do not much care how to represent real numbers. If there are two representations, one can go from one to the other with a bijective function f and then carry over the operations: If addition is defined on the image of f, then one can inherit the definition to the domain of f by

$$x + y = f^{-1}(f(x) + f(y))$$

and so on. Here some examples based on the idea to represent real numbers by digits.

Exercise 18.3. Show that the standard representation can be defined in set-theory: First define a representation for the set $A = \mathbb{Z} \cup \{sign\}$. Then look at the class of all functions $r : A \to \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -\}$, call it B; the decimal point could be placed between $r(0)$ and $r(-1)$ and need not to be represented explicitly.

Define which elements of B represent real numbers and get $\mathbb{R}$ by comprehension, state the property explicitly. For this and further definitions, integer constants, integer addition and $<$ on the integers can be used to in order to deal with positions of digits. The selection should be made such that r represents $\sum_{z \in \mathbb{Z}} r(z) \cdot 10^z$ in the case that $r(sign)$ is $+$ and $-\sum_{z \in \mathbb{Z}} r(z) \cdot 10^z$ in the case that $r(sign)$ is $-$. Make sure that every real occurs in the representation exactly once. For example, fix the sign of 0 to either $+$ or $-$.

This representation has the disadvantage that $\mathbb{N} \not\subseteq \mathbb{R}$. So one distinguishes as in many programming languages like FORTRAN between the natural number 2 and the real number 2.0. Nevertheless, there is a one-to-one mapping $f : \mathbb{N} \to \mathbb{R}$ which maps every natural number to its representative in $\mathbb{R}$. f can be defined inductively using a $g : \mathbb{R} \to \mathbb{R}$ such that $f(S(n)) = g(f(n))$ for all natural numbers n. Give two properties $nat, succ$ such that $nat(r)$ is true iff r is in the range of f and $succ(r, q)$ is true if $nat(r) \wedge nat(q) \wedge q = g(r)$.

In the early days of computing, integers were represented by bytes, more precisely, they were limited to the numbers -128 up to 127. The negative numbers started with a 1 and the positive (including 0) with a 0. So one had that $-128 = 10000000$, $-127 = 10000001$, ..., $-2 = 11111110$, $-1 = 11111111$, $0 = 00000000$, $1 = 00000001$,

$2 = 00000010$, …, $127 = 01111111$. The next exercise shows how to transfer this idea to the representation of the reals.

Exercise 18.4. Consider the following set WS representing the reals Without Sign:

$$\begin{aligned} WS = \{r : \mathbb{Z} \to \{0,1\} \quad | \quad & \forall n \in \mathbb{Z}\, \exists m < n\, (r(m) = 0) \\ \wedge \quad & \exists n \in \mathbb{Z}\, \forall m > n\, (r(m) = r(m-1))\} \end{aligned}$$

One can define on WS an addition $+$. Let $(r+q)(n) = 1$ if one of the following three conditions holds:

1. $r(n) = q(n)$ and $r(k) \neq q(k)$ for all $k < n$;
2. $r(n) = q(n)$ and there is $m < n$ such that $r(m) = 1$ and $q(m) = 1$ and $r(k) \neq q(k)$ for all k with $m < k < n$;
3. $r(n) \neq q(n)$ and there is an $m < n$ such that $r(m) = 0$ and $q(m) = 0$ and $r(k) \neq q(k)$ for all k with $m < k < n$.

Let $(r+q)(n) = 0$ otherwise. From $+$ one can define an ordering $<$ on WS by

$$r < q \Leftrightarrow \exists s \in WS\, (q = r + s \wedge \exists n \in \mathbb{Z}\, \forall m > n\, (s(m) = 0)).$$

Verify that $(WS, +)$ is a commutative group: show that the function *null* mapping $\mathbb{Z}$ to 0 is the neutral element, that $q + r = r + q$ for all $r, q \in WS$, that $r + (q + s) = (r + q) + s$ for all $r, q, s \in WS$ and that for every $r \in WS$ there is a $q \in WS$ with $r + q = null$. Show that $<$ is an ordering of WS.

Definition 18.5. $A \subseteq \mathbb{R}$ is *open* iff for every $a \in A$ there is a positive $\epsilon > 0$ such that $\{b \in \mathbb{R} \mid a - \epsilon < b < a + \epsilon\}$ is a subset of A. A set $A \subseteq \mathbb{R}$ is *closed* iff $\mathbb{R} - A$ is open. A point $a \in A$ is isolated iff there is an open set B with $A \cap B = \{a\}$. A set is perfect iff it is closed but does not have isolated points. Say that $a \in A$ is approximable from above in A iff $a = \inf\{b \in A \mid b > a\}$. A set is compact iff it is closed and if it is bounded in the sense that there are $b, c \in \mathbb{R}$ with $b < a < c$ for all $a \in A$.

Example 18.6. *Let $a, b \in \mathbb{R}$ with $a < b$. The closed interval $\{r \in \mathbb{R} \mid a \leq r \leq b\}$ is a closed set, perfect and compact. The open interval $\{r \in \mathbb{R} \mid a < r < b\}$ is an open set. Every open set is the union of open intervals.*

Remark 18.7. A pair (X, Y) is called a Hausdorff space iff the following four axioms hold.

1. $Y \subseteq \mathcal{P}(X)$ and $\emptyset, X \in Y$.

2. If $A, B \in Y$ then $A \cap B \in Y$.

3. If $W \subseteq Y$ then $\bigcup W \in Y$.

4. If $a, b \in X$ and $a \neq b$ then there are $A, B \in Y$ such that $a \in A$, $b \in B$ and $A \cap B = \emptyset$.

Hausdorff observed that these axioms are true for $X = \mathbb{R}$ and Y being the open subsets of X. He discovered that the structure Y of open subsets of a set X has many characteristic properties of a space. In general, one calls any structure (X, Y) satisfying the first three axioms a *topological space* and Y is called the *topology on* X. By the way, Hausdorff introduced his axioms in his book "Mengenlehre" which is the German translation of the word "set theory".

An example of a property defined by topology is the dimension n of $\mathbb{R}^n$. Lebesgue discovered that the space $\mathbb{R}^n$ has the property that for every set A consisting of open subsets of $\mathbb{R}^n$ with $\bigcup A = \mathbb{R}^n$ there is a set B consisting of open subsets of $\mathbb{R}^n$ such that the following three conditions hold:

- $\bigcup B = \mathbb{R}^n$;
- $\forall b \in B \exists a \in A[b \subseteq a]$;
- $\forall r \in \mathbb{R}^n$ there are at most $n+1$ sets $b \in B$ with $r \in b$.

But one can choose A such that it is possible to replace $n+1$ by n in the third condition of the refinement B chosen for A; so the $n+1$ in the third condition is an invariant describing the dimension of $\mathbb{R}^n$.

Example 18.8. *Call a set A upward-open iff for every $a \in A$ there is $r \in \mathbb{R}$ with $r > a$ and $\{s \in \mathbb{R} \mid a \leq s < r\} \subseteq A$. Then the class of all upward-open sets on $\mathbb{R}$ satisfies Hausdorff's axioms and differs from that class of the open sets in $\mathbb{R}$ defined in Definition 18.5.*

Proof. It is easy to see that the set $\{a \in \mathbb{R} \mid a \geq 0\}$ is upward-open. But this set is not open in the usual sense since it contains 0 without containing any number less than 0. Now Hausdorff's axioms are verified.

1. The empty set is upward-open. Also $\mathbb{R}$ itself is upward-open since for every $a \in \mathbb{R}$ the set $\{s \in \mathbb{R} \mid a \leq s < a+1\}$ is a subset of $\mathbb{R}$.

2. If A, B are upward-open and $a \in A \cap B$. There is $r > a$ such that $\{s \in \mathbb{R} \mid a \leq s < r\} \subseteq A$. Since $a \in B$ one can also find a q such that $a < q < r$ and $\{s \in \mathbb{R} \mid a \leq s < q\} \subseteq B$. The latter set is then also in $A \cap B$. Since such a set exists for all elements of $A \cap B$, $A \cap B$ is upward-open.

3. Let W consist of upward-open subsets of $\mathbb{R}$ and let $a \in \bigcup W$. There is an $A \in W$ with $a \in A$. Since A is upward-open there is an $r > a$ with $\{s \in \mathbb{R} \mid a \leq s < r\} \subseteq A$. This set is also contained in $\bigcup W$ and $\bigcup W$ is upward-open.

4. Assume that $a, b \in \mathbb{R}$ and $a \neq b$. One of them is smaller, say $a < b$. Then $A = \{s \in \mathbb{R} \mid a \leq s < b\}$ and $B = \{s \in \mathbb{R} \mid b \leq s\}$ are two disjoint upward-open sets with $a \in A$ and $b \in B$.

So all four axioms of Hausdorff are satisfied. ∎

Exercise 18.9. Verify that Hausdorff's axioms are true for the set $\mathbb{R}$. That is, verify that $(\mathbb{R}, \{A \subseteq \mathbb{R} \mid A \text{ is open}\})$ is a Hausdorff space.

Exercise 18.10. Let α be any ordinal different from 0 and 1. Define a topology on α by saying that a set $\beta \subseteq \alpha$ is open iff β is an ordinal. Verify that the first three axioms of Hausdorff are satisfied, but not the last fourth one.

Exercise 18.11. Find a topology on the set of ordinals up to a given ordinal α which satisfies the Axioms of Hausdorff and in which an ordinal $\beta \in \alpha$ is isolated iff it is either a successor ordinal or 0.

19 The Continuum Hypothesis

Cantor showed that the cardinality of $\mathbb{R}$ is the same as the cardinality of $\mathcal{P}(\mathbb{N})$. Furthermore he showed that the cardinality of $\mathbb{N}$ is smaller than the of $\mathbb{R}$. But he did not find any set of intermediate cardinality.

Recall that the class of infinite cardinals can be identified with the following class of ordinals:

$$\{\alpha \mid \alpha \geq \omega \wedge \forall \beta < \alpha \, (|\beta| < |\alpha|)\}$$

This subclass is well-ordered and there is an order-preserving isomorphism from every ordinal α to the infinite cardinal ω_α. So ω_0 is just ω. ω_1 is the first uncountable ordinal. Recall that the cardinalities of these ordinals are just called "$\aleph$" with the same index: $\aleph_\alpha = |\omega_\alpha|$ and that 2^κ denotes the cardinality of the power set of any set of cardinality κ.

In the following, it is shown for many natural types of subsets of $\mathbb{R}$ that they have either the cardinality $2^{\aleph_0}$ or are finite or have the cardinality $\aleph_0$. But an intermediate cardinality does not show up. Therefore, Cantor conjectured that there is no intermediate cardinality. That is, he stated the following continuum hypothesis (CH) where "continuum" refers to the set of real numbers.

Conjecture 19.1 (Continuum Hypothesis). $2^{\aleph_0} = \aleph_1$.

This result cannot be proven. But this section deals with partial results obtained by attempts to prove the Continuum Hypothesis. These results show that certain types of sets of real numbers satisfy this hypothesis in the sense that there is no set of intermediate cardinality of this type. That is, sets of this type are either at most countable or have the cardinality of the continuum.

But before dealing with these results, a general theorem is given. This theorem shows how to prove one of the directions in Theorem 20.10 below since its proof easily generalizes to one for the fact that $2^{\aleph_0}$ is not the limit of an ascending sequence of countably many other cardinals.

Theorem 19.2 (König). $2^{\aleph_0} \neq \aleph_\omega$.

Proof. The following implication is proven: $2^{\aleph_0} \geq \aleph_\omega \Rightarrow 2^{\aleph_0} > \aleph_\omega$. Thus $2^{\aleph_0} \neq \aleph_\omega$.

So assume that $2^{\aleph_0} \geq \aleph_\omega$. Let α be the least ordinal having the cardinality $2^{\aleph_0}$, that is, let α be the ordinal representing the cardinal $2^{\aleph_0}$. Then $\alpha \geq \omega_\omega$ where ω_ω is the ordinal representing $\aleph_\omega$ and $\omega_\omega = \bigcup\{\omega_n \mid n \in \mathbb{N}\}$. Since $|\mathcal{P}(\mathbb{N})| = 2^{\aleph_0}$, there is a bijection $F : \alpha \to \mathcal{P}(\mathbb{N})$.

Recall the definition of $\leq_{lin}$ from Example 17.11 which is a relation on subsets of $\mathbb{N}$. This relation is transitive and has two important properties.

- For every A there are at most countably many $B \subseteq \mathbb{N}$ with $B \leq_{lin} A$.
- For every countable set $\{B_1, B_2, \ldots\}$ of subsets of $\mathbb{N}$ there is an A with $B_k \leq_{lin} A$ for all k.

Now extend F to $G : \alpha \times \omega \times \omega \to \mathcal{P}(\mathbb{N})$ with $G(\alpha, i, j) = \{a \mid a \cdot i + j \in F(\alpha)\}$. $G[\alpha, \omega, \omega]$ is the closure downwards under $\leq_{lin}$ of $F[\alpha]$. Each set $F[\omega_n]$ has cardinality $\aleph_n$. Furthermore, using Hessenberg's Theorem, $|G[\omega_n, \omega, \omega]| \leq \aleph_n \cdot \aleph_0 \cdot \aleph_0 = \aleph_n$. Since $|\alpha| > \aleph_n$ there is for each n a set B_n with $B_n \in \mathcal{P}(\mathbb{N}) - G[\omega_n, \omega, \omega]$. Thus there is a set $A \subseteq \mathbb{N}$ with $B_n \leq_{lin} A$ for all n. Since all sets $G[\omega_n, \omega, \omega]$ are closed under $\leq_{lin}$ and each of them does not contain B_n, none of them contains A. Since $\omega_\omega = \bigcup\{\omega_n \mid n \in \mathbb{N}\}$, $A \in \mathcal{P}(\mathbb{N}) - G[\omega_\omega, \omega, \omega]$ and $A \in \mathcal{P}(\mathbb{N}) - F[\omega_\omega]$. By assumption, $A = F(\beta)$ for some $\beta \in \alpha$. Since $\beta \geq \omega_n$ for all $n \in \mathbb{N}$, $\beta \geq \omega_\omega$ and $\alpha > \omega_\omega$. Since α is a cardinal, it is larger than $|\omega_\omega|$ and $2^{\aleph_0} > \aleph_\omega$. This completes the proof. ∎

The next results are the first step on the way to prove Theorem 19.8 which says that every closed set is either at most countable or has the cardinality of the continuum. Cantor's Discontinuum in Exercise 19.7 is one of the first examples of a set of cardinality $2^{\aleph_0}$ which has Lebesgue measure 0.

Theorem 19.3. *Every nonempty open set has cardinality* $2^{\aleph_0}$.

Proof. Let A be an open nonempty subset of $\mathbb{R}$. Open sets are unions of basic open sets of the type $\mathbb{R}_{r,\epsilon} = \{q \in \mathbb{N} \mid r - \epsilon < q < r + \epsilon\}$ where ϵ is a positive real number. So let $r, \epsilon \in \mathbb{R}$ be such that $\epsilon > 0$ and $\mathbb{R}_{r,\epsilon} \subseteq \mathbb{R}$. The following f is a one-to-one mapping from $\mathbb{R}$ into A, in fact $f[\mathbb{R}] = \mathbb{R}_{r,\epsilon}$:

$$f(x) = \begin{cases} r + \epsilon \cdot \frac{x}{x+1} & \text{if } x > 0; \\ r & \text{if } x = 0; \\ r - \epsilon \cdot \frac{x}{x-1} & \text{if } x < 0. \end{cases}$$

Then $|\mathbb{R}| = |A|$ by Proposition 6.5. ∎

Exercise 19.4. If $A \subseteq \mathbb{R}$ is at most countable, then $\mathbb{R} - A$ has cardinality $2^{\aleph_0}$.

Proposition 19.5. *Let $A \subseteq \mathbb{R}$ be perfect. Then there is a countable subset $B \subseteq A$ such that $(B, <)$ is a dense linearly ordered set without end points where $<$ is the natural ordering inherited from $\mathbb{R}$.*

Proof. Let $a, b \in A$ such that there is a third element $c \in A$ with $a < c < b$. Since A is perfect, there is for every $\epsilon > 0$ a further element $c' \in A - \{c\}$ with such that the distance between c and c' is less than ϵ. Starting with $\epsilon = \frac{1}{2} \cdot \min\{c - a, b - c\}$, one obtains that $a < c' < b$. By iterating these argument with an ϵ also smaller than the distance between c and c', one can establish that there are three numbers $c_0, c_1, c_2 \in A$ with $a < c_0 < c_1 < c_2 < b$. Thus, if there is $c \in A$ between a, b then one can find a new element $c_1 \in A$ such that A has elements between a, c_1 and c_1, b. Note that it might be impossible to take $c_1 = c$.

Let $X_0 = \{l, h\}$ with $l, h \in A$ such that there is a further $c \in A$ with $l < c < h$. Using the Axiom of choice, one can define a function f giving $X_{S(n)}$ from X_n such that

1. $X_{S(n)}$ is finite and $X_{S(n)} \subseteq A$;
2. for all $a, b \in X_n$ with $a < b$ there is a $c \in X_{S(n)}$ with $a < c < b$;
3. for all $a, b \in X_{S(n)}$ there is a $c \in A$ with $a < c < b$.

That is, given X_n, f does the following: for all pairs $(a, b) \in X_n$ satisfying $a < b \wedge \forall c \in X_n\, (c \leq a \vee b \leq c)$, f searches a $c_0, c_1, c_2 \in A$ such that $a < c_0 < c_1 < c_2 < b$ and puts a, b, c_1 into $X_{S(n)}$.

The set $C = \cup\{X_0, X_1, \ldots\}$ is in V and also the set $B = \{c \in C \mid l < c < h\}$. Clearly $B \subseteq A$. If $a, b \in B$, then $a, b \in X_n$ for some n and there is a $c \in X_{S(n)}$ such that $a < c < b$. Thus B is dense. Furthermore, B has no end points, since l is the infimum and h the supremum of B with respect to C but $l, h \notin B$; note that they might not be the infimum and supremum with respect to $\mathbb{R}$. Furthermore, B is the union of countably many finite sets and thus countable. ▮

Theorem 19.6. *Every perfect set $A \subseteq \mathbb{R}$ has cardinality $2^{\aleph_0}$.*

Proof. Since $(B, <)$ is dense, there is an isomorphism f from $\mathbb{Q}$ to B. Now define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(r) = \sup\{f(q) \mid q \in \mathbb{Q} \wedge q < r\}.$$

If $r, r' \in \mathbb{R}$ then there are $q, q' \in \mathbb{Q}$ such that $r < q < q' < r'$. It follows that $g(r) \leq f(q) < f(q') \leq f(r')$. Thus g is one-to-one.

Since A is closed, the complement of A is an open set. If there would be an $r \in \mathbb{R}$ with $g(r) \notin A$, then there would also be an $\epsilon > 0$ such that $\{s \in \mathbb{R} \mid g(r) - \epsilon < s < g(r) + \epsilon\}$ is disjoint to A. But then $g(r) - \frac{\epsilon}{2}$ is an upper bound for the subset $\{f(q) \mid q \in \mathbb{Q} \wedge q < r\}$ of A in contradiction to $g(r)$ being the supremum of this subset. Thus $g(r) \in A$ and $|\mathbb{R}| \leq |A|$. Since $A \subseteq \mathbb{R}$, $|A| = 2^{\aleph_0}$. ▮

Exercise 19.7. Cantor's Discontinuum is given as $F(\{0,2\}^{\mathbb{N}})$ where F maps every $f \in \{0,2\}^{\mathbb{N}}$ to the real number having the digits $f(0)f(1)f(2)\ldots$ in the ternary digital representation after the point:

$$F(f) = \sum_{n \in \mathbb{N}} f(n) \cdot 3^{-1-n}.$$

For example, if $f(0) = 2$ and $f(n) = 0$ for all $n \geq 1$ then $F(f)$ represents the ternary number $0.02222\ldots$, that is, $F(f) = \frac{1}{3}$. If E is the set of even numbers and $f(n) = 0$ for $n \in E$ and $f(n) = 2$ for $n \notin E$ then $F(f)$ is the ternary number $0.20202020\ldots$, that is, $F(f) = \sum_{n \in E} 2 \cdot 3^{-1-n} = \frac{3}{4}$. Show that

1. F restricted to $\{0,2\}^{\mathbb{N}}$ is one-to-one;
2. $F(\{0,2\}^{\mathbb{N}})$ does not have any nonempty open subset;
3. $F(\{0,2\}^{\mathbb{N}})$ is perfect.

Furthermore, show that $F(\{0,2\}^{\mathbb{N}})$ is given as

$$\begin{aligned} F(\{0,2\}^{\mathbb{N}}) &= \{r \in \mathbb{R} \mid 0 \leq r \leq 1\} - T \text{ where} \\ T &= \{r \in \mathbb{R} \mid \exists m, n \in \mathbb{N}\,(m \cdot 3^{-n} + 3^{-1-n} < r < m \cdot 3^{-n} + 2 \cdot 3^{-1-n})\}, \end{aligned}$$

that is, T is the set of all positive real numbers for which the digit 1 appears in every ternary representation after the point; $\frac{1}{3} \notin T$ since it has besides $0.1000\ldots$ also the representation $0.02222\ldots$ where no 1 occurs after the point.

Theorem 19.8. *Every uncountable closed subset of reals has a perfect subset. Hence every closed subset of $\mathbb{R}$ is either at most countable or of the cardinality $2^{\aleph_0}$.*

Proof. Let A be a closed subset of $\mathbb{R}$ and let

$$B = \{r \in A \mid \forall \epsilon > 0 \, (\, |A \cap \mathbb{R}_{r,\epsilon}| \geq \aleph_1)\}$$

Then B satisfies the following properties.

1. *$A - B$ is countable.* If $r \in A - B$ there is an $\epsilon > 0$ such that $\mathbb{R}_{r,\epsilon} \cap A$ is countable. Thus there are $q, \delta \in \mathbb{Q}$ such that $\delta > 0$ and $\{r\} \subseteq \mathbb{R}_{q,\delta} \subseteq \mathbb{R}_{r,\epsilon}$ since $\mathbb{Q}$ is dense in $\mathbb{R}$. It follows that

   $$A - B = \bigcup \{A \cap \mathbb{R}_{q,\delta} \mid q, \delta \in \mathbb{Q} \wedge \delta > 0 \wedge |A \cap \mathbb{R}_{q,\delta}| \leq \aleph_0\}$$

   which is countable since it is the union of a countable set of countable sets.

2. *B is closed.* The sets $\mathbb{R} - A$ and

   $$C = \bigcup \{\mathbb{R}_{q,\delta} \mid q, \delta \in \mathbb{Q} \wedge \delta > 0 \wedge |A \cap \mathbb{R}_{q,\delta}| \leq \aleph_0\}$$

   are open, thus $B = A - C = \mathbb{R} - ((\mathbb{R} - A) \cup C)$ is a closed set.

3. *B has no isolated points.* Assume that $r \in B$. Then, for every $\epsilon > 0$, $A \cap \mathbb{R}_{r,\epsilon}$ is uncountable and thus $A \cap \mathbb{R}_{r,\epsilon} - (A - B) - \{r\}$ is also uncountable. Thus r is not an isolated point of B.

So either A is at most countable or B is not empty. In the latter case, B is nonempty and satisfies the last two properties. That is, the cardinality of B is $2^{\aleph_0}$ and since $B \subseteq A \subseteq \mathbb{R}$, A has the same cardinality. ▮

20 The Axioms of Zermelo and Fraenkel

First-order logic permits to state axioms which quantify over elements of V but not over subclasses of V. Furthermore, one can use expressions to define subclasses of V. Consider now a subclass G which is a function, that is, there is a domain (either class or set) such that there is for all $x_1, x_2, \ldots, x_n, W$ exactly one y with $(x_1, x_2, \ldots, x_n, y) \in G$ and all elements of F are of there type. Then one can use Recursion (if $W = \mathbb{N}$) or transfinite recursion (if $W \neq \mathbb{N}$ using some suitable well-founded relation R on W) to construct a new class as done in Theorems 5.2 and 13.7. Furthermore, one can do with classes the usual operations like concatenation. For example, given three classes which are functions $G_1, G_2, G_3 : V^2 \to V$ then the function $x, y, z \mapsto G_1(G_2(x, y), G_3(x, z))$ is also a class and can be used in the below Axioms of Replacement and Comprehension. Nevertheless, it is understood that all the classes considered can be build in finitely many steps from sets and the expressions and properties in Definition 3.7 with these methods.

Axioms 20.1 (Zermelo and Fraenkel). Let V be the class of all sets. There are classes coding functions $x, y \mapsto \{x, y\}$, $x \mapsto \bigcup x$, $x \mapsto \mathcal{P}(x)$ and special sets $\emptyset, \mathbb{N}$ such that the following holds:

Foundation: $\forall x \in V\, (x \neq \emptyset \Rightarrow \exists y \in x\, \forall z \in x\, (z \notin y))$;

Extensionality: $\forall x, y \in V\, (x = y \Leftrightarrow \forall z \in V\, (z \in x \Leftrightarrow z \in y))$;

Existence (of empty set): $\forall x\, (x \notin \emptyset)$;

Pairing: $\forall x, y \in V\, \forall z \in V\, (z \in \{x, y\} \Leftrightarrow (z = x \vee z = y))$;

Schema of Comprehension: For all classes which are unary functions F and for all sets $x \in V$, $\{y \in x \mid F(y) \neq \emptyset\} \in V$;

Union: $\forall x, y \in V\, (y \in \bigcup x \Leftrightarrow \exists z \in x\, (y \in z))$;

Power Set: $\forall x, y \in V\, (y \in \mathcal{P}(x) \Leftrightarrow \forall z \in y\, (z \in x))$;

Infinity: $\emptyset \in \mathbb{N}$ and $\forall y \in V\, (y \in \mathbb{N} \Rightarrow y \cup \{y\} \in \mathbb{N})$,
$\forall x \in V\, (\emptyset \in x \wedge \forall y \in V\, (y \in x \Rightarrow y \cup \{y\} \in x) \Rightarrow \forall z \in \mathbb{N}\, (z \in x))$;

Schema of Replacement: For every n and every class coding an n-ary function F and every sets $x_1, \ldots, x_n$ the set $F[x_1, \ldots, x_n]$ is in V.

Choice: For all sets $x \in V$ there is a function C_x such that for all nonempty $y \in x$, $C_x(y) \in y$.

These axioms are called the *Zermelo-Fraenkel Axioms with Choice* or just *ZFC*. The axiom system *ZF* is obtained by taking all above axioms except the Axiom of Choice.

Definition 20.2. A model of ZF consists of the class V and the relations $\in$ such that the above axioms are satisfied. Similar for models of ZFC.

Remark 20.3. There are several models of set theory, that is, the models are not uniquely defined. A hypothesis H is called independent under ZF if there are two models of ZF such that one satisfies H and the other satisfies $\neg H$.

One method to build models is to start in a large model $(V, \in)$ and then to build inside $(V, \in)$ a smaller model (W, R) with $W, R \in V$ such that (W, R) satisfies the a certain desired combination of axioms. If R is the restriction of $\in$ to W, then one writes $(W, \in)$ instead of (W, R).

Definition 20.4. A structure (W, R) is called an inner model of $(V, \in)$ iff $W, R \in V$ and (W, R) satisfy all set-theoretic axioms with R being a relation standing for the element relation $\in$.

Exercise 20.5. Given any model $(V, \in)$, show that $(V_{\omega_1}, \in)$ is not an inner model of ZFC. Take a well-ordering of $\mathcal{P}(\mathbb{N})$ and show that ω_1 is contained in its range. Therefore, ω_1 is the candidate for the inner model and this can be used to show that $(V_{\omega_1}, \in)$ cannot be an inner model.

Inaccessible cardinals are one example of large cardinals. Intuitively a cardinal is called large if it exists in some but not all models; that is, its existence cannot be proven from the existence of lower cardinals. Note that the cardinal $\aleph_0$ only exists because of the Axiom of Infinity or an equivalent one. Similar, a large cardinal would only be guaranteed to exist if an additional axiom is added and there are models of ZFC where no large cardinals exist. While the notion of a large cardinal is not precisely defined and is just used to denote anything which is not guaranteed to exist due to being large, the notion of an inaccessible cardinal is much more precise and defined as follows.

Definition 20.6. A cardinal $\kappa > \aleph_0$ is *inaccessible* iff

- for all cardinals $\lambda < \kappa$, $2^\lambda < \kappa$ and
- for all sets $L \subseteq \kappa$ of cardinals, either $\sup(L) < \kappa$ or $|L| = \kappa$.

Inaccessible cardinals are interesting since every inaccessible cardinal permits to build a submodel for ZFC from a given model for ZFC. There the condition $\kappa > \aleph_0$ is important because otherwise the Axiom of Infinity will go lost. Recall that a cardinal is identified with the least ordinal of the same cardinality, thus V_κ is defined for every cardinal. The following proposition is given without a proof.

Proposition 20.7. Given a model $(V, \in)$ of ZFC, the following conditions are equivalent for every cardinal $\kappa > \aleph_0$:

1. *κ is inaccessible;*
2. *for every $x \in V_\kappa$, $\sup\{2^{|y|} \mid y \in x\} < \kappa$;*
3. *$|V_\kappa| = \kappa$;*
4. *for every class F being a function in one argument which maps V_κ to V_κ and every $\alpha \in \kappa$ there is $\beta \in \kappa$ with $F[V_\alpha] \in V_\beta$.*

Theorem 20.8. *Let κ be an inaccessible cardinal. Then $(V_\kappa, \in)$ is an inner model of ZFC.*

The proof of this result is omitted. The central idea of the proof would be to show that every set $X \subseteq V_\kappa$ satisfies $\rho(X) = \kappa \Leftrightarrow |X| = \kappa$. Then if f is a class which is a function, either $f(X) \not\subseteq V_\kappa$ and f does not need to be considered or $f(X) \in V_\kappa$ and there is no problem. Using this key idea, one can verify the other axioms.

Theorem 20.9. *The existence of inaccessible cardinals cannot be proven in ZFC.*

Proof. It has to be shown that there is a model of ZFC not containing an inaccessible cardinal. Then it follows that one cannot prove the existence of these cardinals.

Given a model $(V, \in)$, assume that it contains inaccessible cardinals – otherwise there is nothing to prove. So let λ be a cardinal which bounds some inaccessible cardinal. Then $\{\kappa \in \lambda \mid \aleph_0 < |V_\kappa| = |\kappa|\}$ is a set of ordinals in V and thus well-ordered. This set contains all inaccessible cardinals below λ. Thus it has a least element κ. Now $(V_\kappa, \in)$ is a model of ZFC.

It remains to show that this model does not contain "new inaccessible ordinals": So assume that α is an ordinal in $(V_\kappa, \in)$ with $|V_\alpha| = \alpha$. Ordinals are transitive sets such that any two members are comparable with respect to $\in$, so α is also an ordinal in $(V, \in)$. Furthermore, there is a bijection $f : V_\alpha \to \alpha$ and this f is in V_κ. Since $V_\kappa \subseteq V$, $f \in V$ and $|V_\alpha| = \alpha$ also with respect to the model $(V, \in)$. It follows that $\alpha \leq \omega$ in $(V, \in)$ since κ is the least inaccessible cardinal in $(V, \in)$. Since ω is the same in $(V, \in)$ and $(V_\kappa, \in)$, $\alpha \leq \omega$ also in $(V_\kappa, \in)$ and $(V_\kappa, \in)$ has no inaccessible cardinals. Thus the existence of such cardinals in unprovable from ZFC. ∎

The next theorems are given without proof. The first one shows that one cannot decide the Continuum Hypothesis from ZFC because there are models of ZFC where this hypothesis is true and others where it is false. Every statement ϕ which is decidable from ZFC is either true in all models of ZFC or false in all models of ZFC.

Theorem 20.10 (Cohen, Easton, Gödel, König). *Let α be an at most countable ordinal. Then there is a model with $2^{\aleph_0} = \aleph_\alpha$ iff α is a successor ordinal.*

Cantor proved 1878 that $2^\kappa > \kappa$ for all cardinals κ and α cannot be the limit ordinal 0. König proved 1905 that $2^{\aleph_0} \neq \aleph_\alpha$ for all countable limit ordinals, see Theorem 19.2.

Gödel proved 1938 the above Theorem with the parameter $\alpha = 1$. So he obtained that the Continuum Hypothesis is consistent with ZFC and constructed a model with $2^{\aleph_0} = \aleph_1$. Starting with a model $(V, \in)$ of ZFC, Gödel defined a new model $(L, \in)$ by transfinite recursion. He first defined a suitable class $\mathcal{AF}$ of absolute functions and defined then the following classes L_α inductively for all ordinals α:

$$L_\alpha = \{x \in V_\alpha \mid \exists \beta \in \alpha\, \exists y \in L_\beta \cup \{L_\beta\}\, \exists F \text{ in } \mathcal{AF}\, (x = F[y] \wedge x \subseteq L_\beta)\}$$

Note that $L_\beta = V_\beta$ for $\beta \leq \omega$ but it might be that $L_{\omega+1} \subset V_{\omega+1}$. For every $x \in L$, if $y = \mathcal{P}(x)$ in L and $z = \mathcal{P}(x)$ in V then $y = z \cap L$ but it can happen that $y \subset z$. Indeed there is such an $x \in L$ whenever $L \neq V$. Furthermore, functions which exist in V and witness that $|x| \leq |y|$ might fail to exist in L and thus it might happen that $|x| < |y|$ in L. Gödel's model has the following properties.

- $(L, \in)$ satisfies ZFC.
- For all ordinals α, $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

The second property is called the Generalized Continuum Hypothesis (GCH).

Cohen constructed 1963 for every countable ordinal α which is not a limit ordinal a model such that $2^{\aleph_0} = \aleph_\alpha$; actually the method works also for some larger ordinals. Easton investigated 1970 the question what possible outcomes exist for the function f satisfying $2^{\aleph_\alpha} = \aleph_{f(\alpha)}$. In this terminology, Cantor showed $\forall \alpha\, (f(\alpha) > \alpha)$; Gödel showed that it is consistent with ZFC to assume $\forall \alpha\, (f(\alpha) = \alpha + 1)$ (which is GCH), König's result is that $f(0) \neq \alpha$ whenever α is a limit ordinal which is the union of countably many smaller ordinals, Cohen showed that $f(0)$ can be any countable successor ordinal. Easton showed that many functions are possible, for example if $\forall n \in \mathbb{N}\, (\max\{f(n), n+1\} \leq f(n+1) < \omega)$ then there is a model of ZFC with $2^{\aleph_n} = \aleph_{f(n)}$ for all $n \in \mathbb{N}$. Easton's result was indeed a bit more general and showed that one can prescribe the cardinality of the power set for all successor cardinals and some limit ones. But the problem was not yet completely solved, for example Silver showed 1974 that if $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all $\alpha < \omega_1$ then $2^{\aleph_{\omega_1}} = \aleph_{\omega_1+1}$ and cannot take any other value. So in some cases $2^{\aleph_\alpha}$ might be determined by the values of $2^{\aleph_\beta}$ with $\beta < \alpha$.

The next theorem shows that there are nonstandard models of ZFC. A nonstandard model is a model in which the natural numbers do not exist as a set. Instead of the set $\{0, 1, 2, \ldots\}$, there is a set containing some additional elements, here denoted as $\mathbb{N}$ or ω. This set contains also nonstandard numbers beyond the usual natural numbers which behave like natural numbers but are not such numbers. Then the collection $\{\beta \in \omega \mid \beta$ is a nonstandard number$\}$ does not have a least element and therefore is neither a set nor a class. This nonstandard model is not what is intended but one can show that every first-order axiomatization of set theory has a nonstandard model. Only infinite axioms like

$$\forall x\ (x \in \mathbb{N} \Leftrightarrow x = 0 \vee x = 1 \vee x = 2 \vee x = 3 \vee \ldots)$$

can rule out nonstandard models, but such axioms are normally not considered in set theory as infinite formulas are much more difficult to handle than finite formulas. Nevertheless, one often considers axioms which are infinite sets of finite formulas.

Theorem 20.11. *There is a model $(V, \in)$ of ZFC where $\mathbb{N}$ contains an element α such that $0 \neq \alpha$, $1 \neq \alpha$, $2 \neq \alpha$, ...; more informally, $\{0, 1, 2, \ldots\} \subset \mathbb{N}$ and the collection $\{0, 1, 2, \ldots\}$ of all natural numbers is neither a set or a class. α is called a "non-standard number".*

A further pathology is that one can have a countable model of ZFC. That is, one constructs the model of ZFC within a model $(V, \in)$ as an inner model (W, R) and has that $|W| = \aleph_0$ with respect to the model $(V, \in)$. The members of W have of course cardinalities higher than $\aleph_0$ with respect to (W, R); so the notion of cardinality is depending on the view point which one has.

21 References

Klaus Gloede from the University of Heidelberg gave the permission to include the literature from his lecture notes which are in German language. Thus many of the cited books are also in German. Although only English language books are relevant for most of the students of this module, German and French language titles are kept for those who know these languages.

Introductory texts.

1. Cameron, P. J. *Sets, Logic and Categories.* Springer 1999.

2. Deiser, O. *Einführung in die Mengenlehre.* Springer 2004 (German).

3. Devlin, K. *The Joy of Sets.* Springer 1993.

4. Ebbinghaus, D. *Einführung in die Mengenlehre.* Spectrum 2003 (German).

5. Friedrichsdorf, U., Prestel, A. *Mengenlehre für den Mathematiker.* Vieweg 1985 (German).

6. Hrbacek, K., Jech, T.H. *Introduction to Set Theory.* Second edition. Marcel Dekker, New York, 1984.

7. Klaua, D. *Mengenlehre* de Gruyter 1979 (German).

8. Krivine, J.L. *Théorie axiomatique des ensembles.* Presse de l'Université Paris 1969 (French).

9. Moschovakis, Y. *Notes on Set Theory.* Springer 1994.

10. Oberschelp, A. *Allgemeine Mengenlehre.* BI 1994 (German).

11. Lévy, A. *Basic Set Theory.* Springer 1979.

12. Rubin, J.E. *Set Theory for the Mathematician.* Holden-Day 1967.

More comprehensive literature and related fields.

1. Jech, Th. *Set Theory.* Springer 2003.

2. Mendelson, E. *Introduction to Mathematical Logic.* Chapman & Hall 1997.

3. Rautenberg, W. *Einführung in die Mathematische Logik.* Vieweg 2002 (German).

4. Shoenfield, J.R. *Mathematical Logic.* Addison-Wesley 1967.

Special aspects of set theory.

1. Devlin, K. *Aspects of Constructibility.* Springer 1984.

2. Drake, F.R. *Set Theory. An Introduction to Large Cardinals.* North Holland 1974.

3. Felgner, U. *Models of ZF-set theory.* Springer 1971.

4. Jech, Th. *The Axiom of Choice.* Springer 1993.

5. Kanamori, A. *The Higher Infinite. Large Cardinals in Set Theory from Their Beginnings.* Springer 2003.

6. Kechris, Alexander S. *Classical descriptive set theory.* Springer 1995.

7. Kunen, K. *Set Theory. An Introduction to Independence Proofs.* North Holland 1983.

8. Moschovakis, Yiannis N. *Descriptive set theory.* North Holland 1980.

Original texts of the founders of set theory.

1. Bolzano, B. *Paradoxien des Unendlichen.* Leipzig 1851; Meiner 1955 (German).

2. Cantor, G. *Gesammelte Abhandlungen.* Springer 1980 (German).

3. Dauben, J.W. *Georg Cantor. His Mathematics and Philosophy of the Infinite.* Princeton 1990.

4. Felgner, U. (editor) *Mengenlehre.* Wissenschaftliche Buchgesellschaft 1979 (German).

5. Fraenkel, A. *Einleitung in die Mengenlehre.* Springer 1919.

6. Fraenkel, A., Bar Hillel, Y., Lévy, A. *Foundations of Set Theory.* North Holland 1973.

7. Gödel, K. *Collected Works.* Oxford 1986-2003.

8. Hallett, M. *Cantorian set theory and limitation of size.* Oxford University Press 1984.

9. Hausdorff, F. *Gesammelte Werke Band II. Grundzüge der Mengenlehre.* Springer 2002.

10. Lavine, S. *Understanding the Infinite.* Harvard 1998.

11. Meschkowski, H. *Probleme des Unendlichen. Werk und Leben Georg Cantors* Vieweg 1967.

12. Meschkowski, H. *Hundert Jahre Mengenlehre* Dtv 1973.

13. Moore, G.H. *Zermelo's Axiom of Choice.* Springer 1982.

14. Purkert, E., Ilgauds, H.J. *Georg Cantor 1845-1918.* Birckhäuser 1987.

15. Quine, Willard V. *Set Theory and its Logic.* Harvard 1969.

16. SPALT, D. (editor) *Rechnen mit dem Unendlichen* Birckhäuser 1990 (German).

Set theory using alternative systems of axioms.

1. BERNAYS, P. *Axiomatic set theory.* North Holland 1968.
2. CHUAQUI, R. *Axiomatic Set theory. Impredicative Theory of Classes.* North Holland 1981.
3. FORSTER, T. E. *Set theory with a universal set.* Oxford University Press 1992.
4. POTTER, M.D. *Sets. An Introduction.* Oxford University Press 1990.
5. QUINE, WILLARD V. *Mengenlehre und ihre Logik.* Vieweg 1973.
6. SCHMIDT, J. *Mengenlehre I.* BI 1966.
7. VOPENKA, P., HAJEK, P. *The Theory of Semisets.* North Holland 1972.

Bibliography and Links.

1. MÜLLER, G.H., LENSKI, V. (editors) Ω-*Bibliography of Mathematical Logic.* Vol. I-VI. Springer 1987. See also: http://www-logic.uni-kl.de/BIBL/index.html.
2. Klaus Gloede's lecture notes on set theory: http://www.math.uni-heidelberg.de/logic/skripten.html (German)
3. Homepage for Mathematical Logic: http://www.uni-bonn.de/logic/world.html
4. Biographies of mathematicians: http://www-groups.dcs.st-and.ac.uk/~history/BiogIndex.html