SET THEORY I and II – Lecture notes

Department of Logic, Radek Honzik

Version: Winter 2012

Poznámka: *Skripta k výše uvedenému kurzu. Mohou se vyskytnout překlepy a chyby. Preliminary version; may contain typographical and other errors.*

Doporučená literatura:

[BS] Bohuslav Balcar, Petr Štěpánek, *Teorie množin*, Academia 2000.

[KK] Kenneth Kunen, *Set Theory: An Introduction to Independence Proofs*, Elsevier 2004.

[TJ] Thomas Jech, *Set Theory*, Springer 2003.

[AK] Akihiro Kanamori, *The Higher Infinite*, Springer 2003.

Contents

1 An axiomatic framework for set theory

1.1 Axioms of set theory

Excellent and detailed account can be found in [BS]. We just stress some points.

We define here the *Zermelo-Fraenkel* set theory (ZF), a first-order predicate theory in the language $\{=, \in\}$.

Remark 1.1 We may ask the following question again, in this more concrete context: "Why axiomatic set theory, i.e. why theory in this formal sense?". Consider the following *Russell's paradox.* Assume that every property gives rise to a set (this sounds reasonable enough). Consider the property $P(x) \equiv_{df} x \notin x$. Then existence of a set $y = \{x \mid P(x)\}$ leads to a contradiction: if $y \in y$, then y must satisfy the property $P(y)$, and so $y \notin y$; if $y \notin y$, then $P(y)$ holds, and so y qualifies to be an element of y: $y \in y$. In both cases we reached a contradiction, and so such a set y cannot exist.

We interpret this paradox in the following way: we must be more restrictive in what a set is (for instance y will not be a set – it is "too big"). We describe sets in "algebraic" fashion, listing operations which when applied to sets yield sets again. In other words, we will build our sets from bottom up: from simple sets to more complicated sets.

In the axiomatization we attempt to list all formulas which we think hold (without least doubt) about sets.

Now we formulate the principles in our chosen formal system of first order predicate logic and prove some basic properties of sets.

[ZF0] Existence of a set.

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

This is just to make sure that there is at least one set (note that because we have not constants in our language, this is necessary).

[ZF1] Extensionality (Axiom extenzionality).

$$(\forall x, y)[x = y \leftrightarrow (\forall q)(q \in x \leftrightarrow q \in y)].$$

Note that one half of ZF1 is provable from the axioms of predicate calculus (Exercise):

$$\vdash (\forall x, y)[x = y \rightarrow (\forall q)(q \in x \leftrightarrow q \in y)]. \tag{1.1}$$

We define a new binary relational symbol $\subseteq$ (a subset, podmnožina):

$$x \subseteq y \leftrightarrow (\forall q)(q \in x \rightarrow q \in y). \tag{1.2}$$

Exercise. Realize that

$$\text{ZF1} \vdash (\forall x, y)(x = y \leftrightarrow x \subseteq y \wedge y \subseteq x).$$

[ZF2] Pairing (Axiom dvojice).

$$(\forall x, y)(\exists z)(\forall q)(q \in z \leftrightarrow q = x \vee q = y).$$

Let $(\exists! x)\varphi(x)$ be a shorthand for $(\exists x)\varphi(x) \wedge [(\forall x, y)(\varphi(x) \wedge \varphi(x/y) \rightarrow x = y)]$, where x/y denotes a substitution of y for x. We read the quantifier $\exists! x$ as "there is exactly one x".

Exercise. Show

$$\text{ZF1}, \text{ZF2} \vdash (\forall x, y)(\exists! z)(\forall q)(q \in z \leftrightarrow q = x \vee q = y).$$

[Hint. ZF2 implies that there is at least one such z. ZF1 implies that there is at most one such z: if z_1 and z_2 satisfy ZF2, then they have the same elements, and by ZF1, this means that $z_1 = z_2$.]

As such z is unique we can add a new binary functional symbol which we denote as $\{\cdot,\cdot\}$ and write $\{x, y\}$ for the z in ZF2.

We define a new binary operation: $\langle x, y\rangle$ (or sometimes written as (x, y)) – an ordered pair (uspořádaná dvojice). The definition is as follows:

$$\langle x, y\rangle = \{\{x\}, \{x, y\}\}. \tag{1.3}$$

ZF2 implies that $\langle x, y\rangle$ exists. We show that this definition satisfies the property which require of an ordered pair.

Lemma 1.2 *The following holds*

$$\mathrm{ZF1}, \mathrm{ZF2} \vdash (\forall x, y, v, w)[\langle x, y\rangle = \langle v, w\rangle \leftrightarrow x = v \wedge y = w)].$$

Proof. The direction from right to left is follows from the axioms of identity. We will show the converse:

$$(\forall x, y, v, w)[\langle x, y\rangle = \langle v, w\rangle \rightarrow x = v \wedge y = w). \tag{1.4}$$

Fix arbitrary sets x, y, v, w. We will show the following equivalent reformulation of (1.4): if $x \neq v$ or $y \neq w$, then it already holds that $\langle x, y\rangle \neq \langle v, w\rangle$. By Section 1.4.1, Reasoning by cases (8), we need to show that:

(i) $x \neq v$ implies $\langle x, y\rangle \neq \langle v, w\rangle$, and
(ii) $y \neq w$ implies $\langle x, y\rangle \neq \langle v, w\rangle$.

Realize that if $\langle x, y\rangle = \langle v, w\rangle$, then it holds:

$$\big[\{x\} = \{v\} \vee \{x\} = \{v, w\}\big] \wedge \big[\{x, y\} = \{v\} \vee \{x, y\} = \{v, w\}\big]. \tag{1.5}$$

Let us shorten the expression in (1.5) as $A \wedge B$, where $A := \{x\} = \{v\} \vee \{x\} = \{v, w\}$, and $B := \{x, y\} = \{v\} \vee \{x, y\} = \{v, w\}$.

Proof of (i). Assume $x \neq v$ and assume for contradiction that $\langle x, y\rangle = \langle v, w\rangle$. Then one of the two identies in A must hold: if $\{x\} = \{v\}$, then $x = v$, and if $\{x\} = \{v, w\}$, then $x = v = w$. In both cases, this contradicts the assumption $x \neq v$. It follows A does not hold, and so $\langle x, y\rangle \neq \langle v, w\rangle$ is true.

Proof of (ii). Assume $y \neq w$ and assume for contradiction that $\langle x, y\rangle = \langle v, w\rangle$. Then A must be true, and so $x = v$. B must also be true: assume that the first part of B is true: $\{x, y\} = \{v\}$: then $x = y = v$. $x = y = v$ together with the assumption $\langle x, y\rangle = \langle v, w\rangle$ implies that $x = y = v = w$ (because $x = y$ implies $\{x\} = \{x, y\}$, and so $\{v\} = \{v, w\}$), which contradicts $y \neq w$. So assume that the second part of B is true: $\{x, y\} = \{v, w\}$: because $x = v$ is true, this can only be true if $y = w$, which again contradicts $y \neq w$. It follows that $\langle x, y\rangle = \langle v, w\rangle$ cannot be true, and so $\langle x, y\rangle \neq \langle v, w\rangle$ holds.

Note. This proof is a rather long verification of something in a sense very trivial. The apparent complexity of the proof is caused by the necessity to distinguish many cases and rule them out one by one.

Note. The expression "Fix arbitrary sets x, y, v, w" at the beginning of the proof is a correct move by Theorem on constants stated in Section 1.4.1, item (9). □

By induction[1] we can define an ordered n-tuple as follows: $\langle a_0\rangle = a_0$, and $\langle a_0, a_1, \dots, a_k\rangle = \langle\langle a_0, a_1, \dots, a_{k-1}\rangle, a_k\rangle$. The analogue of Lemma 1.2 is shown by induction.

[1] We mean an induction in the metatheory, using the natural numbers we intuitively have.

Fact 1.3

$$\mathrm{ZF1}, \mathrm{ZF2} \vdash (\forall x_0, \ldots, y_0, \ldots)[\langle x_0, \ldots, x_k\rangle = \langle y_0, \ldots, y_k\rangle \leftrightarrow x_0 = y_0 \wedge \ldots \wedge x_k = y_k].$$

Note. From now on we will not specifically say which axioms are needed to show a given claim. We just say "it is provable that"; the meaning is "it is provable from the axioms introduced so far that".

[ZF3*] Separation scheme (schema vydělení). Let $\varphi(q, p)$ be a formula with two free variables q and p.

$$(\forall x, p)(\exists z)(\forall q)[q \in z \leftrightarrow q \in x \wedge \varphi(q, p)]. \tag{1.6}$$

For each formula $\varphi(q, p)$ the axiom (1.6) is an Axiom of Separation. We view p as the parameter of the definition.

By axiom of extensionality for each x and φ the set in ZF3* is determined uniquely and we may add a new operation the value of which is written as $\{\cdot \mid \ldots\varphi\ldots\}$; for illustration, for given x and p we write

$$z = \{q \mid q \in x \wedge \varphi(q, p)\} \tag{1.7}$$

for z in (1.6).

Remark 1.4 Realize that for every formula $\varphi(x, p)$ we add one axiom. [ZF3*] is thus a collection of infinitely many axioms. Notice that by the syntactical rules of the first-order predicate calculus we are not allowed to quantify formulas, so there is no hope of "replacing" the infinite number of axioms in [ZF3*] by a single axiom of the type

$$\text{this is wrong: } \forall x, p \forall\varphi(x, p)(\exists z)(\forall q)[q \in z \leftrightarrow q \in x \wedge \varphi(q, p)].$$

There are good reasons to forbid the quantification over formulas: consider the following so called *Berry's paradox*: Since there are infinitely many natural numbers, there are certainly some such numbers which cannot be defined by any combination of eleven words in English or less (we do not allow numbers bigger than say 100 to appear as "words" in these sentences). Define n to be "the least number not definable with eleven words or less." Then n *is* definable using eleven words or less, which is a contradiction. Notice that if we allowed quantification such as $\forall\varphi\ldots$, then we would be in a similar situation which is paradoxical in Berry's paradox.

From the formulation of the schema with a single parameter p, it already follows that we can have more parameters (this is proved by using the ordered n-tuples: $\langle p_0, \ldots, p_n\rangle$ is just a single set):

Fact 1.5 *Let $\psi(q, p_0, \ldots, p_n)$ be a formula with the free variables shown. Then it is provable*

$$(\forall x, p_0, \ldots, p_n)(\exists z)(\forall q)[q \in z \leftrightarrow q \in x \wedge \psi(q, p_0, \ldots, p_n)]. \tag{1.8}$$

We know show that the formula $q \neq q$ from the Russell's paradox does not lead to contradiction when applied in the "safe" context of ZF3*:

Lemma 1.6 *Fix x and let*

$$z = \{q \mid q \in x \wedge q \notin q\}.$$

This set exists by ZF3. Then $z \notin z$ and $z \notin x$.*

Proof. The assumption of $z \in z$ leads to contradiction, and so $z \notin z$ must be true. To show that $z \notin x$, assume for contradiction that $z \in x$. Then we can show both $z \notin z$ and $z \in z$ which is a contradiction, and hence $z \notin x$. (Note that in the original Russell's paradox, we did not have the extra assumption that $z \in x$, and so all we could say was that the whole system was contradictory, not just the assumption that $z \in x$.)

Note that assuming Axiom of Fundation (see below), we actually have $z = x$. □

Corollary 1.7 *There is no set containing all sets.*

Proof. Assume V is the set containing all sets. Then we do obtain contradiction from the existence of set $z = \{q \mid q \in V \wedge q \notin q\}$ because $z \in V$ is true in this case. □

The Separation scheme enables us to define a lot of other operations common in set theory (we can do that since by Axiom of extensionality these operations are correctly defined). Let x, y, z be sets.

- Intersection (průnik) $x \cap y = \{q \mid q \in x \wedge q \in y\}$, difference (rozdíl) of two sets $x \setminus y = \{q \mid q \in x \wedge q \notin y\}$.[2]
- Emptyset: $\emptyset = \{q \mid q \in x \wedge q \neq q\}$, where x is an arbitrary set.

  Exercise. More precisely, consider the property $\varphi(y)$ given by "$(\forall q) q \notin y$. It can be shown: (i) that there is at least one set which satisfies $\varphi(y)$ – to show that such a set exists we use ZF3*: for instance the set $z = \{q \mid q \in x \wedge q \neq q\}$ above satisfies $\varphi(z)$; (ii) it can be shown that there is at most one such set: if there were two such sets, they would need to differ by an element (because of the Axiom of extensionality), but this is impossible. We can therefore add to our language a new symbol, $\emptyset$, to denote this set.
- Definition: x and y are *disjoint* (disjunktní) if $x \cap y = \emptyset$.
- Intersection (generalization of intersection):

  $$\bigcap x = \{q \mid (\forall q')(q' \in x \to q \in q')\}. \tag{1.9}$$

  If x is nonempty, then x is a set because if $y \in x$ is some set, then

  $$\bigcap x = \{q \mid q \in y \wedge (\forall q')(q \in x \to q \in q')\}. \tag{1.10}$$

  If $x = \emptyset$, then $\bigcap \emptyset$ is not a set; in fact every set at all is the element of $\bigcap \emptyset$ (that is $\bigcap \emptyset$ is the whole universe of sets, denoted as V).

[ZF4] Axiom of union (sjednocení).

$$(\forall x)(\exists z)(\forall q)[q \in z \leftrightarrow (\exists y)(y \in x \wedge q \in y)]. \tag{1.11}$$

We introduce the following abbreviations:

$$(\exists y \in x)\varphi \text{ for } (\exists y)(y \in x \wedge \varphi)$$

and

$$(\forall y \in x)\varphi \text{ for } (\forall y)(y \in x \to \varphi).$$

By Axiom of extensionality, we can define a new operation

$$\bigcup x = \{q \mid (\exists y \in x) q \in y\}.$$

[2] Some authors write just $x - y$ for $x \setminus y$.

Define $\bigcup\{x,y\} = x \cup y$. ($\bigcup$ is an infinite version of $\cup$).

Exercise. Note that $\{x\} \cup \{y\} = \{x,y\}$ but ZF4 does not imply ZF2. [Hint. To show that $\{x\}$ is a set still requires ZF2.]

[ZF5] Power set (potence).

$$(\forall x)(\exists z)(\forall q)(q \in z \leftrightarrow q \subseteq x). \tag{1.12}$$

Definition. We say that q is a proper subset (vlastní podmnožina) of x if $q \subseteq x$ but $q \neq x$.

We can form a new unary operation:

$$\mathscr{P}(x) = \{q \,|\, q \subseteq x\}.$$

Lemma 1.8 *There is no set x such that $\mathscr{P}(x) \subseteq x$. As a corollary, this again shows that V (the universe of all sets) is not a set (because if V were a set, then $\mathscr{P}(V) \subseteq V$ must be true).*

Proof. Assume for contradiction that there is x such that

$$\mathscr{P}(x) \subseteq x \tag{1.13}$$

Fix such an x. Consider the set z defined in Lemma 1.6. z is clearly a subset of x. By our assumption (1.13) it must hold that $z \in x$. But this is contradictory by Lemma 1.6. □

The **powerset axiom** allows us to define the following operations:

- **Product** $x \times y$, where:

  $$x \times y = \{q \,|\, q \in z \wedge (\exists q_0, q_1)(q = \langle q_0, q_1 \rangle \wedge q_0 \in x \wedge q_1 \in y)\}. \tag{1.14}$$

  The product is a set because $z = \mathscr{P}(\mathscr{P}(x \cup y))$ is a set and $x \times y \subseteq z$, then apply Schema of Comprehension.

  By induction on $n \in \mathbb{N}$ we define $(x_1 \times \ldots x_n \times x_{n+1}) = (x_1 \times \ldots \times x_n) \times x_{n+1}$. We write x^n to denote the set $x \times x \ldots$, where x occurs n-times.

- **An n-ary relation (relace)** on sets $x_1, \ldots, x_n$ is a subset of $x_1 \times \ldots \times x_n$. An n-ary relation r is a relation on x if $r \subseteq x^n$.

  For a binary relation $r \subseteq x \times y$ we define **domain**(definiční obor) of r as

  $$\mathrm{dom}(r) = \{q \,|\, (\exists q' \in y)\langle q, q' \rangle \in r\}. \tag{1.15}$$

  and similarly we define **range** (obor hodnot) of r as

  $$\mathrm{rng}(r) = \{q \,|\, (\exists q' \in x)\langle q, q' \rangle \in r\}. \tag{1.16}$$

  Exercise. Verify that dom(r) and rng(r) are sets. [Hint. Both are subsets of $\bigcup\bigcup r$.]

  We also define the **inverse** (inverz relace)

  $$r^{-1} = \{\langle q, q' \rangle \,|\, \langle q', q \rangle \in r\}. \tag{1.17}$$

  and if $a \subseteq x$ we define the **image of r on a** (obraz r přes a):

  $$r''a = \{q \,|\, (\exists q' \in a)\langle q', q \rangle \in r\}. \tag{1.18}$$

  If $a \subseteq x$ then we say that $r \restriction a$ is the **restriction** (zúžení) of r to a, where

  $$r \restriction a = \{\langle q, q' \rangle \,|\, \langle q, q' \rangle) \in r \wedge q \in a\}. \tag{1.19}$$

  Verify that r^{-1}, $r''a$ and $r \restriction a$ are sets.

  Exercise.

1. Consider the relation $\leq$ defined on natural numbers (we write $x \leq y$ for $\langle x, y\rangle \in\leq$). In set theory, the set of all natural numbers $\mathbb{N}$ is customarily denoted as ω (and by convention includes 0).[3] It follows that $\leq \subseteq \omega \times \omega$, and $\leq^{-1} = \geq$. What is $\leq'' \{2\}$?
2. $\in$ is a binary relation with domain the universe of all sets: by Pairing axiom, x is an element of $\{x\}$ for every set x. What is the range of $\in$? Let x be a set; what is $\mathrm{dom}(\in\restriction x)$?
3. Check the following for a binary relations x, x' and sets y, z:
   (a) $x \cup x', x \cap x', x \setminus x'$ are binary relations,
   (b) $x''(y \cup z) = x''y \cup x''z$,
   (c)

   $$x''(y \cap z) \subseteq x''y \cap x''z \text{ and } x''y \setminus x''z \subseteq x''(y \setminus z). \tag{1.20}$$

   Give an example where the converse inclusion $\supseteq$ does not hold in (1.20). Compare with (1.24).

- **Composition of relations** (skládání relací). If r, s are binary relations then the composition $r \circ s$ is defined as

  $$r \circ s = \{\langle x, z\rangle \mid (\exists y)\langle x, y\rangle \in r \wedge \langle y, z\rangle \in s\}. \tag{1.21}$$

  Exercise. Check the following for all binary relations r, s, t:
  1. $\mathrm{dom}(r^{-1}) = \mathrm{rng}(r)$, $\mathrm{rng}(r^{-1}) = \mathrm{dom}(r)$, $(r^{-1})^{-1} = r$, $\mathrm{dom}(r \circ s) \subseteq \mathrm{dom}(r)$, $\mathrm{rng}(r \circ s) \subseteq \mathrm{rng}(s)$. When the identity holds in the last two formulas?
  2. $(r \circ s)^{-1} = s^{-1} \circ r^{-1}$.
  3. $r \circ (s \circ t) = (r \circ s) \circ t$.
  4. Let Id be the identity relation (it is a class, see below for some notes on classes), $\mathrm{Id} = \{\langle x, x\rangle \mid x \in V\}$. Then $r \circ \mathrm{Id} = \mathrm{Id} \circ r = r$.

- A binary relation r is called a **function** if it satisfies the following:

  $$(\forall x, y, \bar{y})(\langle x, y\rangle \in r \wedge \langle x, \bar{y}\rangle \in r \rightarrow y = \bar{y}). \tag{1.22}$$

  Since every function f is a relation, we can use for f the notation defined above for relations: Let f be a function.

  – If $x \in \mathrm{dom}(f)$ we write $f(x)$ for y such that $\langle x, y\rangle \in f$.

  – If $a \subseteq \mathrm{dom}(f)$ we write $f[a]$ for $\{y \mid (\exists x \in a)\langle x, y\rangle \in f\}$.
  Note that as f is a relation, it holds that $f[a] = f''a$ by the notation for relations; when dealing with functions however, we often (not always) prefer to use the notation $f[a]$.

  – If $b \subseteq \mathrm{rng}(f)$ we write $f^{-1}[b]$ for $\{x \mid (\exists y \in b)\langle x, y\rangle \in f\}$. Note again that this can be written as $f^{-1''}b$. Which notation is used depends on the context.

  – **Notation.** Let $f : x \rightarrow y$ and $g : y \rightarrow z$ be two functions. This notation means that $\mathrm{dom}(f) = x$, $\mathrm{dom}(g) = y$ and $\mathrm{rng}(f) \subseteq y$ and $\mathrm{rng}(g) \subseteq z$. We denote by $f \circ g$ the composition of the functions f and g viewed as relations. It follows that $f \circ g$ is a function $f \circ g : x \rightarrow z$ such that for each $q \in x$, $f \circ g\ (q) = g(f(q))$.

  – f is called 1-1 (injective) (prostá) if it satisfies:

  $$(\forall x_0, x_1 \in \mathrm{dom}(f))\ x_0 \neq x_1 \rightarrow f(x_0) \neq f(x_1).$$

  Exercise. Verify that f is 1-1 iff f^{-1} is a function.

  – $f : x \rightarrow y$ is *onto* (na) y if $\mathrm{rng}(f) = y$.

  Exercises.

[3] We have not yet shown how to construct ω in set theory, but we will do it later.

1. Let $f : x \to y$ and $g : y \to z$ be 1-1 functions, then:
   (a) $f \circ g$ is 1-1.
   (b) $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.
2. Let x, y be any sets and f a function (and so f^{-1} is a relation):

   $$(f^{-1})''(x \cap y) = (f^{-1})''x \cap (f^{-1})''y \text{ and } (f^{-1})''(x \setminus y) = (f^{-1})''x \setminus (f^{-1})''y. \tag{1.23}$$

   If f is morever 1-1, then it also holds:

   $$f''(x \cap y) = f''x \cap f''x \text{ and } f''(x \setminus y) = f''x \setminus f''y. \tag{1.24}$$

   [Notice that this is a strengthening of the results in (1.20); you can use the results in (1.20) here.]

- **Index sets** (indexové množiny). Let i and a be sets and $\bar{a}$ a 1-1 function such that $\mathrm{dom}(\bar{a}) = i$ and $\mathrm{rng}(\bar{a}) = a$. Then

  $$a = \{\bar{a}(j) \,|\, j \in i\} \tag{1.25}$$

  and we say that a is indexed by i. In practice it is customary to write I instead of i and a_j instead of $\bar{a}(j)$ so that

  $$a = \{a_i \,|\, i \in I\}. \tag{1.26}$$

  Compare with (1.37).

[ZF6] Axiom of infinity (axiom nekonečna).

$$(\exists x)[\emptyset \in x \wedge (\forall q)(q \in x \to q \cup \{q\} \in x)]. \tag{1.27}$$

Under all reasonably definitions of *finiteness*, a set in the axiom ZF6 is infinite. We will show later that this is indeed true once we define in detail when a set is finite.

Note that a set x in (1.27) is note determined uniquely, there are more sets satisfying (1.27). If x satisfies (1.27), we say that x is *inductive*. We will define the set of natural numbers $\mathbb{N}$, or ω, as follows:[4]

$$\omega = \mathbb{N} = \bigcap \{x \,|\, x \text{ is inductive}\}. \tag{1.28}$$

Remark 1.9 We are not entitled to use the operation $\bigcap$ here according to (1.9) unless we show first that $S = \{x \,|\, x \text{ is inductive}\}$ is a set. But S is *not* a set. We still find useful to refer to objects such as S and we call them *classes*. We say that a class is any system of **sets** which is defined by a formula with parameters.[5] Every set is a class because if x is a set then $x = \{q \,|\, q \in x\}$ and so is defined by the formula $q \in x$ with x as a parameter. Some classes however are not sets, and these are called *proper classes* (vlastní třídy). S is a proper class. Another example of a proper class is the *universe of all sets* (univerzum všech množin), which we denote as $V = \{x \,|\, x = x\}$. Classes are usually written in capital letters: A, B etc. When dealing with (proper) classes we must remember that these are not sets (so for instance the expression $A \in B$ is meaningless because $\in$ is a binary relation between *sets*), and most importantly we *must not* quantify them. More about classes is in Subsection 1.2.

[4] There are more equivalent definitions of $\mathbb{N}$. We will state another one, using *ordinals*, later in the text. See Remark 4.28. Also note that we will verify the properties of $\mathbb{N}$ here; we will rely on our intuition. Formal development of properties of natural numbers is deferred to Section 4.4 where we deal with ordinal numbers. The symbol ω is used because the set $\mathbb{N}$ is itself an ordinal number, and ordinal numbers are denoted with Greek letters.

[5] So if $\varphi(x, p_0, \dots p_n)$ is a formula and $p_0, \dots, p_n$ are sets then $\{q \,|\, \varphi(q, p_0, \dots, p_n)\}$ is a class. So for instance $\{x \,|\, x \notin x\}$ is a class.

However we can argue that the operation of intersection $\bigcap$ can be generalized so that it can be used to classes, and moreover when applied to a class, it will yield a *set*. Indeed if A is a nonempty class defined by a formula $\varphi_A(x, p_0, \dots, p_n)$, that is $q \in A \leftrightarrow \varphi_A(q, p_0, \dots, p_n)$, and $y \in A$ is arbitrary then

$$\bigcap A = \{q \,|\, q \in y \wedge (\forall q')(\varphi_A(q', p_0, \dots, p_n) \rightarrow q \in q'\} \tag{1.29}$$

and so $\bigcap A$ is a set by the axiom of separation.

It follows that $\mathbb{N}$ is a set.

Lemma 1.10 *$\mathbb{N}$ is an inductive set and it is the least such.*

Proof. $\emptyset$ is clearly in every inductive set, and so in $\mathbb{N}$. Also, if q is in every inductive set, so must be by definition $q \cup \{q\}$. Hence $\mathbb{N}$ is inductive. $\mathbb{N}$ is the least such (in the ordering $\subseteq$) because clearly $\mathbb{N} = \bigcap\{x \,|\, x \text{ is inductive}\} \subseteq y$ for every inductive set y. □

The following is a key definition of notation for natural numbers:

Definition 1.11 *We define by induction the following notation for natural numbers in ω: $\emptyset = 0$, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, etc.*

Note that a natural number n is thus defined to be the set of all smaller natural numbers $\{0, \dots, n-1\}$.

[ZF7*] Replacement scheme (schema nahrazení). We say that a formula $\varphi(u, v, p)$ determines a function (compare with (1.22)) if

$$(\forall p, u, v_0, v_1)[\varphi(u, v_0, p) \wedge \varphi(u, v_1, p) \rightarrow v_0 = v_1]. \tag{1.30}$$

If $\varphi(u, v, p)$ determines a function, we view u as an argument of the function and v the value of the function. In keeping with the notation for classes in Subsection 1.2, we can write F to denote the class $\{\langle u, v\rangle \,|\, \varphi(u, v, p)\}$; since $\varphi(u, v, p)$ determines a function, F is a function and we can write $F(u) = v$ instead of $\langle u, v\rangle \in F$.

Let $\varphi(u, v, p)$ be a formula determining a function, then the following statement is an axiom of replacement (axiom nahrazení) for the formula $\varphi(u, v, p)$:

$$(\forall p)(\forall x)(\exists z)(\forall q)[q \in z \leftrightarrow (\exists q' \in x)\varphi(q', q, p)]. \tag{1.31}$$

For each formula $\varphi(u, v, p)$ which determines a function, the formula in (1.31) is an axiom of ZF. Since there are infinitely many of such formulas, the Replacement scheme contains infinitely many axioms.

If we denote as F the function determined by $\varphi(u, v, p)$, we can reformulate the axiom as follows:

$$\text{For every set } x, F[x] \text{ is a set.} \tag{1.32}$$

Note the following properties:

- As in the scheme of Separation, we can show that more parameters as in $\varphi(u, v, p_0, \dots, p_n)$ are allowed (see Fact 1.5).
- Replacement scheme implies Separation scheme. This means that we can "cancel" the axioms ZF3* from our system while retaining its strength. Hint: Fix x. Given $\varphi(u, p)$, let $\psi(u, v, p) = \varphi(u, p) \wedge u = v$. We can show that Axiom of Replacement applied to $\psi(u, v, p)$ proves the existence of a set $a = \{q \in x \,|\, \varphi(q, p)\}$.

- Replacement scheme plus Powerset Axiom imply the Pairing Axiom: Assuming the existence of $\emptyset$, Powerset axiom implies that $\mathscr{P}(\mathscr{P}(\emptyset)) = \{\emptyset, \{\emptyset\}\}$ is a set. Let a, b be sets. We want to show that there is a set $c = \{a, b\}$. Apply Axiom of replacement with the formula $(u = \emptyset \wedge v = a) \vee (u = \{\emptyset\} \wedge v = b)$ to $\{\emptyset, \{\emptyset\}\}$; it will yield the set c as required.
- If φ is not a function, then we may not obtain a set. Consider a class relation determined by the formula $u \neq v$ (we can view this as some $\varphi(u, v, p)$ with p missing). Then for every u, the class $\{v \,|\, u \neq v\}$ is equal to $V - \{u\}$ and thus is a proper class [If $V - \{u\}$ were a set, say a, then by union axiom $a \cup \{u\} = V$ is also a set, and this is a contradiction. In general if $A = V - b$, where b is a set, then A is a proper class.]

[ZF8] Axiom of foundation.

$$(\forall x)[x \neq \emptyset \rightarrow (\exists q)(q \in x \wedge x \cap q = \emptyset)]. \tag{1.33}$$

Little reflection shows that [ZF8] says that every non-empty x has a minimal element with respect to the relation $\in$: that is, there is some $y \in x$ such that there is no $z \in x$ which satisfies $z \in y$. Note that a minimal element may not be unique – x may have more minimal elements, for instance the set $x = \{a, b\}$, where $a \neq b$, $a \notin b$, and $b \notin a$, has exactly two minimal elements: a and b.

Axiom of foundation is a structural axiom which prohibits the existence of "bad" sets, i.e. sets which are unpleasant to deal with and which, importantly, are not required for mathematical arguments. We show later that if our axiomatic system is consistent without ZF8, then it stays consistent with ZF8. This means that we are not running the risk of introducing inconsistency by using ZF8.

Axiom of foundation implies (Exercise):

- There is no set x such that $x = \{x\}$.
- There is no set y such that $y \in y$ (if $y \in y$ then existence of $\{y\}$ violates foundation because $y \cap \{y\} = \{y\}$ and is thus non-empty).
- There are no cycles $y_0 \in y_1 \in y_0$ because Axiom of Foundation would fail for $\{y_0, y_1\}$; in general there are no finite cycles $y_0 \in y_1 \in \ldots \in y_n \in y_0$, for the same reason.
- There can be no infinite $\in$-chain: $y_0 \ni y_1 \ni y_2 \ni \ldots$. (If there were such, then ZF8 would fail for $x = \{y_i \,|\, i \in \omega\}$).

Definition 1.12 *Axioms* [ZF0] – [ZF8] *are called the* Axioms of Zermelo-Fraenkel set theory*, and are denoted as* ZF*.*

Note that we have shown in the discussion concerning the Axioms of Replacement ZF7* that ZF3* and ZF2 follow from the remaining axioms. So we may define ZF to contain just the axioms ZF1,ZF4-8. However, do not forget that in any case there are *infinitely* many axioms in ZF (because of ZF7*).

[ZF9] Axiom of choice (AC).

$$(\forall x)(\exists f)[(f \text{ is a function with } \mathrm{dom}(f) = x - \{\emptyset\}) \wedge (\forall q)[(q \in x \wedge q \neq \emptyset) \rightarrow f(q) \in q)]]. \tag{1.34}$$

Such f is called a *choice function* (for x). It can be shown that ZF $\nvdash$ AC and ZF $\nvdash \neg$AC. We say that AC is independent on the axioms of ZF (or equivalently ZF does not decide whether AC holds or not). ZF together with AC is written as ZFC and is called Zermelo-Fraenkel with Choice.

1.2 Classes

Recall the brief discussion of classes in Remark 1.9. A class is a collection of sets satisfying some formula φ with parameters $p_0, \ldots, p_n$; if $\varphi(u, p_0 \ldots p_n)$ is a formula we denote the collection of all sets q such that $\varphi(q, p_0 \dots p_n)$ by a capital letter, for instance A. We then write $q \in A$ as a shorthand for $\varphi(q, p_0 \dots p_n)$.

If A, B are classes then we my still reasonably define some set-theoretical operations:

- $A = B$ if for all q, $q \in A \leftrightarrow q \in B$.
- $A \cap B$, $A \cup B$, $A - B$.
- Universal class defined by the formula $u = u$ is written as V.
- $\bigcup A$, $\bigcap A$ (if $A = \emptyset$ then by definition $\bigcap \emptyset = V$; if $A \neq \emptyset$, then $\bigcap A$ is a *set*).
- If $a \in A$, then $\bigcap A \subseteq a \subseteq \bigcup A$.
- If A is a class and a a set, then $A \cap a$ is always a *set*.
- $A \times B = \{\langle a, b\rangle \mid a \in A \wedge b \in B\}$.

Note that Scheme of Comprehension states that for every class P and a set x, the class $P \cap x$ is a set. Similarly, the Scheme of Replacement states that for every class function F and a set x, the class $F[x]$ is a set.

Remark 1.13 It can be shown by induction that classes can be eliminated from the language of set theory (replace them by their defining formulas).

1.3 Basic properties of sets: Boolean algebra of sets

Let x be a non-empty set. Consider the set $\mathscr{P}(x)$ together with the operations $\cap, \cup, -$; i.e. for $a, b \in \mathscr{P}(x)$, $a \cap b$ is the intersection, $a \cup b$ the union, and $-a = x \setminus a$ the complement of a.

Lemma 1.14 *The set $\mathscr{P}(x)$ together with operations $\cup, \cap, -$ satisfies the following formulas, where a, b, c are arbitrary elements of $\mathscr{P}(x)$:*

(i) Associativity. $a \cap (b \cap c) = (a \cap b) \cap c$, $a \cup (b \cup c) = (a \cup b) \cup c$.
(ii) Commutativity. $a \cap b = b \cap a$, $a \cup b = b \cup a$.
(iii) Absorption. $a \cap (a \cup b) = a$, $a \cup (a \cap b) = a$.
(iv) Distributivity. $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$, $a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$.
(v) Complement. $a \cup -a = x$, $a \cap -a = \emptyset$.

Note that properties (i)–(iv) hold for all sets, we do not have to restrict ourselves to $\mathscr{P}(x)$. $\mathscr{P}(x)$ is used to define the complement of a: $-a = x \setminus a$.

Proof. By the definition of operations $\cap, \cup, -$ (they use the propositional connectives $\wedge, \vee, \neg$), we first show that the above formulas (i)–(v) hold for the propositional connectives $\wedge, \vee, \neg$, and constants $1 =$ truth, and $0 =$ falsity in place of x and $\emptyset$, respectively.

Recall the definition of connectives: they are functions with domain $\{0, 1\}$ and range $\{0, 1\}$. We write $p \wedge q$ for cunjunction (konjunkce) and $p \vee q$ for disjunction (disjunkce). $p \wedge q$ is 1 only if both p and q are 1. $p \vee q$ is 0 only if both p and q are 0. $\neg p$ is 1 if $p = 0$, and 0 if $p = 1$.

Exercise. Show that $\wedge, \vee, \neg, 0, 1$ satisfy the formulas (i)–(v) above. [Hint. These are just propositional tautologies (výrokové tautologie).]

As soon as we know that $\wedge, \vee, \neg, 0, 1$ satisfy (i)–(v), the proof of lemma is easy. For instance to argue that $a \cap b = b \cap a$ we need to show that for every q: $q \in a \wedge q \in b$ is equivalent to $q \in b \wedge q \in a$; however this is true because the conjunction $\wedge$ is commutative. Similarly for other formulas in (i)–(v). □

Remark 1.15 A structure B of the form $B = \langle B, \wedge, \vee, -, 0, 1\rangle$ is called a *Boolean algebra* if it satisfies the formulas (i)–(v) (when $\cup$ is replaced by $\vee$, $\cap$ by $\wedge$, $\emptyset$ by 0 and x by 1). Thus we have shown above that propositional connectives are a Boolean algebra over the domain $\{0, 1\}$, and operations $\cap, \cup, -$ are a Boolean algebra over a domain $\mathscr{P}(x)$ for any x. On every Boolean algebra $B = \langle B, \wedge, \vee, -, 0, 1\rangle$, where the operations $\wedge, \vee, -$ are arbitrary operations on B which satisfy (i)–(v), one can define the so called *canonical* partial ordering $\leq_B$ on B for all $x, y \in B$:

$$x \leq_B y \leftrightarrow x \wedge y = x \leftrightarrow x \vee y = y. \tag{1.35}$$

The ordering $\leq_B$ is usually not linear. 0 is the least element and 1 the greatest element in $\leq_B$.

Exercise. Verify that the inclusion relation $\subseteq$ is the canonical ordering $\leq_B$ on the powerset algebra $B = \langle \mathscr{P}(x), \cap, \cup, -, \emptyset, x\rangle$.

Exercise. Show that $\mathscr{P}(x)$ also satisfies the following formulas (where a, b, c are arbitrary elements of $\mathscr{P}(x)$):

(1.36)

(1) $--a = a$
(2) $-a = -b \to a = b$
(3) (de Morgan laws) $-(a \cup b) = -a \cap -b$, $-(a \cap b) = -a \cup -b$
(4) If $a \subseteq b$ then $a \cup c \subseteq b \cup c$, $a \cap c \subseteq b \cap c$, and $-b \subseteq -a$
(5) $a \subseteq c \wedge b \subseteq c \leftrightarrow a \cup b \subseteq c$, and
$a \subseteq b \wedge a \subseteq c \leftrightarrow a \subseteq b \cap c$

[Hint. Again show first that propositional connectives satisfy these formulas.] *Note:* The formulas above are true in any Boolean algebra.

The Boolean algebra of sets $\mathscr{P}(x)$ satisfies also the so called *infinite* versions of de Morgan's laws and distributivity. We introduce some notation first.

Let I is an *index set* of a set a, in the sense of (1.25).

It follows we can write

$$a = \{q \,|\, q \in a\} = \{a_i \,|\, i \in I\} \tag{1.37}$$

If $a = \{a_i \,|\, i \in I\}$, then of course

$$\begin{aligned} \bigcup\{a_i \,|\, i \in I\} &= \bigcup a \\ \bigcap\{a_i \,|\, i \in I\} &= \bigcap a \end{aligned} \tag{1.38}$$

The lefthand side of (1.38) is sometimes written as $\bigcup_{i\in I} a_i$, and $\bigcap_{i \in I} a_i$.

Lemma 1.16 *Let $\{a_i \,|\, i \in\}$ be a family of subsets of x, i.e. for every $i \in I$, $a_i \in \mathscr{P}(x)$. The Boolean algebra of sets $\mathscr{P}(x)$ satisfies the following infinite laws:*

(i) (infinite de Morgan laws)
$-\bigcap_{i\in I} a_i = \bigcup_{i\in I} -a_i$, $-\bigcup_{i\in I} a_i = \bigcap_{i\in I} -a_i$
(ii) (infinite distributive laws)
$b \cap \bigcup_{i\in I} a_i = \bigcup_{i\in I}(b \cap a_i)$, $b \cup \bigcap_{i\in I} a_i = \bigcap_{i\in I}(b \cup a_i)$

Proof. Exercise. [Hint. For de Morgan's laws, use the fact that $\neg(\forall x)\varphi$ is logically equivalent to $(\exists x)\neg\varphi$, and $\neg(\exists x)\varphi$ is equivalent to $(\forall x)\neg\varphi$, for arbitrary φ. The proof of infinite distributive laws uses the fact that $\psi \wedge [(\exists x)\varphi(x)]$ is logically equivalent to $(\exists x)[\psi \wedge \varphi(x)]$ providing that x is not free in ψ, and $\psi \vee [(\forall x)\varphi(x)]$ is logically equivalent to $(\forall x)[\psi \vee \varphi(x)]$ providing that x is not free in ψ.] □

1.4 Optional topics

1.4.1 Practical guide: what we can use in proofs

In practice, we can use the following, for more see either [AS] or [VS]: Let us denote as $\perp$ the contradiction, i.e. a formula of the form $\varphi \wedge \neg\varphi$ for some φ.

(1) All propositional tautologies. Very often one uses the contraposition (obměna):

$$\vdash (\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi), \tag{1.39}$$

other examples include for instance $\vdash \neg(\varphi \wedge \psi) \leftrightarrow \neg\varphi \vee \neg\psi$, $\vdash \neg(\varphi \vee \psi) \leftrightarrow \neg\varphi \wedge \neg\psi$, $\vdash \varphi \to \psi \leftrightarrow \neg\varphi \vee \psi$, $\vdash \neg(\varphi \to \psi) \leftrightarrow \varphi \wedge \neg\psi$, etc.

(2) We define $(\exists x)\varphi$ iff $\neg(\forall x)\neg\varphi$.

   (a) $\vdash \neg(\forall x)\varphi \leftrightarrow (\exists x)\neg\varphi$

   (b) $\vdash \neg(\exists x)\varphi \leftrightarrow (\forall x)\neg\varphi$. etc

(3) (Modus Ponens.) If we derive $\varphi \to \psi$ and also φ, we can derive ψ.

(4) (Generalization.) If we derive $\varphi(x)$ without assuming anything special about x, we can derive $(\forall x)\varphi(x)$.

(5) (Deduction theorem.) If σ and ψ do not contain free variables, then to argue $T \vdash \sigma \to \psi$, it is sufficient to argue $T, \sigma \vdash \psi$. Or using (1.39), it is sufficient to argue $T, \neg\psi \vdash \neg\sigma$.

(6) To argue for $T \vdash \varphi \wedge \psi$, argue separately for φ, ψ.

(7) (Reasoning by contradiction.) To argue for $T \vdash \varphi$, argue $T, \neg\varphi \vdash \perp$.

   Very often it occurs in the following form: assume we want to prove $T \vdash \varphi \to \psi$ for some sentences φ, ψ, then it suffices to show that $T, \varphi \wedge \neg\psi \vdash \perp$.

(8) (Reasoning by cases) To argue that $T, \varphi \vee \psi \vdash \theta$, argue that $T, \varphi \vdash \theta$ and $T, \psi \vdash \theta$.

   Note that it is generally not enough to argue either $T, \varphi \vdash \theta$ or $T, \psi \vdash \theta$. Consider the example of a tautology $\varphi \vee \neg\varphi$. Clearly, $T, \varphi \vee \neg\varphi \vdash \theta$ if and only if $T \vdash \theta$. It follows that $T, \varphi \vdash \theta$ cannot suffice to argue that $T, \varphi \vee \neg\varphi \vdash \theta$.

(9) (Theorem on constants – a technical device which makes it possible to apply Deduction theorem even in case when the formula is not a sentence) Let T be a theory. Assume $\varphi(x_0, \ldots)$ is a formula and $c_{x_0}, c_{x_1}, \ldots$ are symbols for constants which do not occur in the language of T (they are completely new). Then

$$T \vdash (\forall x_0, \ldots)\varphi(x_0, \ldots) \text{ iff } T \vdash \varphi[x_0/c_{x_0}, \ldots],$$

   where x/c_x means substitution of the symbol c_x for x. This is very important in the context of Deduction theorem above: the Deduction theorem requires that σ has no free variables. If it does, we can still use Deduction theorem providing we replace the variables by new constants (which means that we can no longer quantify them!).

2 Comparing sizes

2.1 Relation "to be bigger than" for infinite objects

Recall that a binary relation R on a class A, i.e. $R \subseteq A \times A$, is called a *partial order* or shortly an *ordering* (česky: uspořádání) (on A) if it satisfies the following properties for all $x, y, z \in A$:

(i) $\langle x, x\rangle \in R$ (reflexivity)
(ii) $\langle x, y\rangle \in R \wedge \langle y, z\rangle \in R \rightarrow \langle x, z\rangle \in R$ (transitivity)
(iii) $\langle x, y\rangle \in R \wedge \langle y, x\rangle \in R \rightarrow x = y$ (weak antisymmetry)

We say that a relation R on A is *linear* if for all $x, y \in A$, either $\langle x, y\rangle \in R$ or $\langle y, x\rangle \in R$.

It is customary to write the symbol $\leq$ (and its variants such as $\preceq$) to denote a partial order; we also write $x \leq y$ instead of $\langle x, y\rangle \in \leq$. From now on we will use this convention.

The partial order $\subseteq$ is too strong for comparing sizes of sets – if $x \subseteq y$ is true than x might be really considered "smaller" than y; however if $a \neq b$ are two sets then $\{a\} \not\subseteq \{b\}$ and $\{b\} \not\subseteq \{a\}$ (we say that $\{a\}$ and $\{b\}$ are *incomparable* in $\subseteq$), but as they both have just one element, they should have the same "size".

Goal. We want to define a partial order $\preceq$ on the universal class V that could be interpreted as correctly capturing the intuitive notion of one set being smaller in size than another. We argued above that the inclusion relation $\subseteq$ is not suitable.

Definition 2.1 *Let x, y be two sets.*

(i) We say that x, y have the same size, and denote this as $x \approx y$, if there is a bijection from x onto y.
(ii) We say that x has size smaller or equal to y, and denote it as $x \preceq y$, if there is a 1-1 function from x into y.
(iii) If $x \preceq y$ is true but there is not bijection from x onto y (i.e. $x \not\approx y$) then we say that x is strictly smaller than y and denote it as $x \prec y$.

Examples.

(i) If $x \subseteq y$ then $x \preceq y$ [Hint. Use the identity function.]. So our definition of $\preceq$ includes the inclusion relation.
(ii) $\{x\} \approx \{y\}$ for every x, y. So our definition corrects the drawback of $\subseteq$ mentioned above.
(iii) If $x \subseteq y$ but $x \neq y$ then $x \approx y$ is still possible: Let y be the set of natural numbers and x the set of all even numbers $E = \{2, 4, \ldots\}$. Then $i : n \mapsto 2n$ is a bijection from $\mathbb{N}$ onto E, and so $\mathbb{N} \approx E$. This means that the property of x being a proper subset y (i.e. $x \subseteq y \wedge x \neq y$) does not imply that x has size strictly smaller than y.
(iv) The idea of comparing sizes using some 1-1 functions works correctly with finite objects:[6] if x has n elements and y has m elements then

$$n < m \text{ iff } x \prec y \tag{2.40}$$

2.2 Basic properties

Lemma 2.2 *(i) The relation $\approx$ is an equivalance relation on V.*
(ii) The relation $\preceq$ is reflexive and transitive on V, but not weakly antisymmetric.

Proof. Ad (i). Let x, y be sets. Then $x \approx x$ because the identity function id_x on x, where $\mathrm{id}_x = \{\langle a, a\rangle \mid a \in x\}$, is a bijection between x and x. If $x \approx y$ via some bijection $f : x \rightarrow y$, then

[6] Here we work intuitively, not within our formal ZFC; we have not yet defined what a finite object is.

$f^{-1} : y \to x$ shows $y \approx x$. If $x \approx y$ and $y \approx z$ via $f : x \to y$ and $g : y \to z$, then the composition $g \circ f : x \to z$ shows $x \approx z$.

Ad (ii). Similarly as in (i). To show that $\preceq$ is not weakly antisymmetric note that for two sets $a \neq b$, it clearly holds $\{a\} \preceq \{b\}$ (as witnesses by the bijection $\{\langle a, b\rangle\}$) and $\{b\} \preceq \{a\}$, but $\{a\} \neq \{b\}$.

□

Note that we write $\preceq$ but $\preceq$ is not an ordering by (ii). But it is "almost" an ordering: we say that $\preceq$ is a *pre-ordering* (quasi-uspořádání). The important Theorem 2.6 connects the relations $\approx$ and $\preceq$ in the natural way and shows that although $x \preceq y \wedge y \preceq x$ does not imply $x = y$, it does imply that x and y have the some size: $x \approx y$. Using the fact that $\approx$ is an equivalance, this means that x and y are in the same equivalance class.

Before we show Theorem 2.6, we first show Theorem 2.3, which we will use in the proof of Cantor-Bernstein theorem, but which is interesting by itself.

Theorem 2.3 (Fixed point theorem, věta o pevném bodě.) *Let x be a set and let H be a monotonic map from $\mathscr{P}(x)$ to $\mathscr{P}(x)$, i.e. for all $a, b \in \mathscr{P}(x)$, if $a \subseteq b$ then $H(a) \subseteq H(b)$. Then there exists a* fixed point *$c \subseteq x$ of H, i.e. a set $c \subseteq x$ such that $H(c) = c$.*

Proof. Consider the following set

$$C = \{u \subseteq x \mid u \subseteq H(u)\} \tag{2.41}$$

and denote $c = \bigcup C$. Note that C is non-empty because it contains at least the set $\emptyset$: $\emptyset \subseteq H(\emptyset)$ ($\emptyset$ is a subset of any set). We want to show that c is a fixed point. First we prove

$$c \subseteq H(c). \tag{2.42}$$

First notice that if u is C then $u \subseteq \bigcup C = c$. Now: if q is in c, there is some $u \in C$ such that $q \in u$. Because $u \in C$, it follows $u \subseteq H(u)$, and also $H(u) \subseteq H(c)$ because H is monotonic and $u \subseteq c$. Thus $u \subseteq H(u) \subseteq H(c)$, and so q is in $H(c)$ as required.

We now need to show the converse, i.e.

$$H(c) \subseteq c. \tag{2.43}$$

We will apply monotonicity of H to (2.42), obtaining $H(c) \subseteq H(H(c))$. This means that $H(c)$ is an element of C, and so in particular

$$H(c) \subseteq c. \tag{2.44}$$

(2.42) and (2.44) together imply $c = H(c)$ as required. □

Notice the role of the set $c = \bigcup C$ in the above proof. c is the *supremum* of the set C with respect to the ordering $\subseteq$. This fact is important for the idea behind the proof of the fixed point theorem. Recall the definition of the supremum and the infimum:

Definition 2.4 *If $\langle A, \leq\rangle$ is a partial order, and $x \subseteq A$ is a nonempty set, then we say that $r_1 \in A$ is the* supremum *of x (with respect to $\leq$) if:*

(i) For all $a \in x$, $a \leq r_1$ (r_1 is an upper bound *(horní závora) of x),*
(ii) r_1 is the least upper bound, i.e. if s is in A and for all $a \in x$, $a \leq s$, then $r_1 \leq s$.

Similarly, we say that $r_2 \in A$ is the infimum *of x (with respect to $\leq$) if:*

(i) For all $a \in x$, $a \geq r_2$ (r_2 is a lower bound *(dolní závora) of x),*
(ii) r_2 is the greatest lower bound, i.e. if s is in A and for all $a \in x$, $a \geq s$, then $r_2 \geq s$.

The ordering $\subseteq$, and the operations $\bigcup$ and $\bigcap$ satisfy the following general lemma:

Lemma 2.5 *Let a be a non-empty set. Then $\langle \mathscr{P}(a), \subseteq \rangle$ is a partial order. If $\emptyset \neq x \subseteq \mathscr{P}(a)$, then $\bigcup x$ is the supremum of x and $\bigcap x$ the infimum of x in $\langle \mathscr{P}(a), \subseteq \rangle$.*

Proof. We need to show that $\bigcup x$ is the lowest upper bound of x in $\langle \mathscr{P}(a), \subseteq \rangle$. Clearly $\bigcup x$ is an upper bound because if $c \in x$, then $c \subseteq \bigcup x$. If y is an upper bound of x, then $\bigcup x \subseteq y$ (because if $q \in \bigcup x$, then there is some $z \in x$ such that $q \in z$; since y is an upper bound of x, $z \subseteq y$, and therefore $q \in y$ as required).

Similarly argue for $\bigcap x$. □

We now return to the formulation and the proof of the Cantor-Bernstein theorem.

Theorem 2.6 (Cantor, Bernstein) *For every x, y:*

$$x \approx y \leftrightarrow (x \preceq y \wedge y \preceq x). \tag{2.45}$$

Proof. The direction from left to right in (2.45) is obvious, so we need to prove the converse: $(x \preceq y \wedge y \preceq x) \rightarrow x \approx y$.

Let $f : x \rightarrow y$ a 1-1 function from x to y and $g : y \rightarrow x$ a 1-1 function from y to x. If $a \subseteq x$, recall the notation $f[a] = \{b \in y \,|\, (\exists a' \in a) f(a') = b\}$; clearly, if $a \subseteq b \subseteq x$, then $f[a] \subseteq f[b]$. We can view $f[\cdot]$ as a new function, determined by f; $f[\cdot]$ is a function from $\mathscr{P}(x)$ into $\mathscr{P}(x)$ which is monotonic with respect to $\subseteq$. Since f is 1-1, the function $f[\cdot]$ is also 1-1. The same applies to g. We now define a monotonic map from $\mathscr{P}(x)$ to $\mathscr{P}(x)$ as follows:

$$H(u) = x \setminus g[y \setminus f[u]], \text{ for every } u \subseteq x. \tag{2.46}$$

Claim: H is monotonic with respect to $\subseteq$: let $a \subseteq b$ be subsets of x; we need to show that $H(a) \subseteq H(b)$. If $a \subseteq b$, $y \setminus f[a] \supseteq y \setminus f[b]$ and also $g[y \setminus f[a]] \supseteq g[y \setminus f[b]$ (the operation $\setminus$ reverses the inclusion relation; see Exercises (1.36), item (4)). Finally after applying $\setminus$ again, we get $x \setminus g[y \setminus f[a]] \subseteq x \setminus g[y \setminus f[g]]$ as required.

Let c be a fixed point of the map H ensured by Theorem 2.3:

$$c = H(c) = x \setminus g[y \setminus f[c]]. \tag{2.47}$$

It implies that

$$x \setminus c = g[y \setminus f[c]]. \tag{2.48}$$

Thus we can define a bijection h from x onto y as follows:

$$h(a) = \begin{cases} f(a) & \text{for } a \in c, \\ g^{-1}(a) & \text{for } a \in x \setminus c. \end{cases}$$

Note that h is equal to the union of f restricted to c with g^{-1} restricted to $x \setminus c$; in symbols $h = f \restriction c \cup g^{-1} \restriction (x \setminus c)$.

Let us verify that h is really a bijection from x onto y. (2.48) implies that g^{-1} is defined for all elements of $x \setminus c$ and so $\mathrm{dom}(h) = x$. h is clearly 1-1 on the set c because h is the same as f on c and f is 1-1. g^{-1} is clearly 1-1 on $x \setminus c$; to check that h is 1-1 it suffices to show that $f[c] \cap g^{-1}[x \setminus c] = \emptyset$, and to show that h is onto it suffice to show that that $f[c] \cup g^{-1}[x \setminus c] = y$. But both these identities are true because $g^{-1}[x \setminus c]$ is the complement of $f[c]$ in y, i.e. $g^{-1}[x \setminus c] = y \setminus f[c]$. This ends the proof. □

Another basic, but important, theorem states that the powerset operation strictly increases the size of the original set. The technique of the proof utilizes the so called *diagonalization method* (diagonalizace).

Theorem 2.7 (Cantor.) *For every set x,*

$$x \prec \mathscr{P}(x).$$

Proof. Define for $a \in x$: $f(a) = \{a\}$. f is a 1-1 function from x to $\mathscr{P}(x)$, which shows $x \preceq \mathscr{P}(x)$.

To show $x \prec \mathscr{P}(x)$ assume for contradiction that there is a bijection $g : x \to \mathscr{P}(x)$. Define

$$a = \{y \in x \,|\, y \notin g(y)\} \tag{2.49}$$

The set a is a subset of x, and hence an element of $\mathscr{P}(x)$. Since g is onto, there must be some $z \in x$ such that $g(z) = a$. We reach contradiction by showing that $z \in a$ and also $z \notin a$. Assume first $z \in a$; then by definition of a, $z \notin g(z) = a$. Conversely, if $z \notin a$, then $z \in a$ by the definition of a. □

Corollary 2.8 *The set $\mathscr{P}(\omega)$ is strictly larger than ω.*

How big is $\mathscr{P}(\omega)$? This is an important question because as we will see in Theorem 2.17 below, the size of $\mathscr{P}(\omega)$ is exactly the size of $\mathbb{R}$.

But first we verify how the relation $\approx$ interacts with the operations we already have in set theory:

But let us first define:

Definition 2.9 *Let x, y be arbitraly sets, then we write ${}^{x}y$ to denote the following set:*

$$^{x}y = \{f \,|\, f : x \to y\}, \tag{2.50}$$

where we write $f : x \to y$ to denote a function f with $\mathrm{dom}(f) = x$ *and* $\mathrm{rng}(f) \subseteq y$.

Exercise. Show that if y is any set, then ${}^{\emptyset}y = \{\emptyset\}$ and if $y \neq \emptyset$, then ${}^{y}\emptyset = \emptyset$.

Remark 2.10 The x, y in the definition of ${}^{x}y$ can be finite. We can take Definition 2.9 as a definition of the usual exponentiation (mocnění) on the natural numbers: n^m is defined as the number of all functions in ${}^{m}n$. The following Lemmas then show that this definition satisfies all the intuitive properties: for instance that $n^{mk} = (n^m)^k$, etc. It is also easy to check that this definition of n^m is equivalent to the definition by recursion: $n^0 = 1$ and $n^{m+1} = n^m n$.

In the following Lemmas we show some basic properties of the relations $\preceq$ and $\approx$ with respect to operations $\times$ and ${}^{x}y$. The proofs below involve finding a 1-1 or 1-1 and onto (bijection) function f from some set v to some other set w. Note the following easy observation: Let f_1 and f_2 be functions from v to w, then:

$$f_1 \neq f_2 \Leftrightarrow \text{ there is some } q \in v \text{ such that } f_1(q) \neq f_2(q). \tag{2.51}$$

In words, two functions with the same domain are different iff there is an argument on which they are different.

Notational note. If $f : v \to w$ is a function then $f(q)$ for $q \in v$ can also be a function, for instance when $w = {}^{x}y$ for some sets x, y. If $x' \in x$ we write $f(q)(x')$ to denote the value which the function $f(q)$ takes at x'.

Lemma 2.11 *For all x, y, x_1, y_1:*

(i) $x \times y \approx y \times x$,
(ii) $x \times (y \times z) \approx (x \times y) \times z$,

(iii) $(x \approx x_1 \wedge y \approx y_1) \to (x \times y) \approx (x_1 \times y_1)$,
(iv) $x \approx y \to \mathscr{P}(x) \approx \mathscr{P}(y)$,
(v) $\mathscr{P}(x) \approx {}^x2$, *where* $2 = \{\emptyset, \{\emptyset\}\}$.
(vi) $x^2 \approx {}^2x$. *The exponentiation* ${}^yx$ *can thus be viewed as a generalization of the Cartesian product.*

Proof. We will just define the relevant functions f. As an Exercise show that the functions defined are really bijections between the respective sets.

Ad (i). Define $f : x \times y \to y \times x$ as the function which to a pair $\langle a, b\rangle$ assigns the pair $\langle b, a\rangle$.

Ad (ii). Define $f : x \times (y \times z) \to (x \times y) \times z$ as the function which to a pair $\langle a, \langle b, c\rangle\rangle$ assigns $\langle\langle a, b\rangle, c\rangle$.

Ad (iii). Let $g_1 : x \to x_1$ and $g_2 : y \to y_1$ be bijections. Define $f : (x \times y) \to (x_1 \times y_1)$ as the function which to a pair $\langle a, b\rangle$ assigns the pair $\langle g_1(a), g_2(b)\rangle$.

Ad (iv). Let $g : x \to y$ be a bijection. Define $f : \mathscr{P}(x) \to \mathscr{P}(y)$ as the function which to $a \subseteq x$ assigns $g[a] = \{b \in y \,|\, (\exists q \in a) g(q) = b\} \subseteq y$.

Ad (v). If $y \subseteq x$ is a subset, then we define the *characteristic function of* y (charakteristická funkce množiny y) $\chi_y : x \to 2$ by defining for each $q \in x$:

$$\chi_y(q) = \begin{cases} 1 & \text{if } q \in y, \\ 0 & \text{if } q \notin y. \end{cases}$$

Intuitively, χ_y says about each element of x whether it belongs to y (value 1), or does not belong to y (value 0). We define the bijection $f : \mathscr{P}(x) \to {}^x2$ as the function which to each $y \in \mathscr{P}(x)$ assigns the characteristic function χ_y.

Ad (vi). Recall that $x^2 = x \times x$. Define $f : x^2 \to {}^2x$ by assigning to each $\langle a, b\rangle$ the function $\{\langle 0, a\rangle, \langle 1, b\rangle\}$. □

Lemma 2.12 *Let x, y, u, v be sets:*

(i) $\emptyset \neq x \preceq y \to {}^xu \preceq {}^yu$,
(ii) $u \preceq v \to {}^yu \preceq {}^yv$,
(iii) ${}^{(x\times y)}u \approx {}^x({}^yu) \approx {}^y({}^xu)$.

Proof. We will just define the relevant function f. As an Exercise show that they are 1-1 or bijections, depending on the context.

Ad (i) By the assumption $\emptyset \neq x$, both x and y are non-empty. We can also assume that u is non-empty because if $u = \emptyset$, we get ${}^xu \approx {}^yu$. Let $g : x \to y$ be 1-1. Define $f : {}^xu \to {}^yu$ as the function which to $h : x \to u$ assigns the function $h' : y \to u$ defined by

$$h'(q) = \begin{cases} h(g^{-1}(q)) & \text{for } q \in g[x], \\ a & \text{otherwise,} \end{cases}$$

where a is some fixed element of u. Show that f is 1-1.

Ad (ii). Let $g : u \to v$ be 1-1. Define $f : {}^yu \to {}^yv$ by assigning to a function $h : y \to u$ the function $h' : y \to v$ defined by $h'(q) = g(h(q))$ for each $q \in y$. Show that f is 1-1.

Ad (iii). We will define a bijection $f : {}^{x\times y}u \to {}^x({}^yu)$. The bijection between ${}^x({}^yu)$ and ${}^y({}^xu)$ is left to the reader as a (simple) exercise. Given a function $h : (x \times y) \to u$ and an element $a \in x$ let us define a unary function $h_a : y \to u$, where $h_a(b) = h(\langle a, b\rangle)$ for every $b \in y$ (view the function h_a as the function h with the argument a fixed: $h(a, \cdot) = h_a(\cdot)$ where $\cdot$ denotes the argument of the function). Define f as the function which to a $h : (x \times y) \to y$ assigns to the function $h' : x \to {}^yu$ defined by $h'(a) = h_a$. Show that f is a bijection. [Hint. To show that f is 1-1 note that if $h_1 \neq h_2$ are different functions in ${}^{x\times y}u$, then there is some argument $\langle a,b\rangle$ on which they are different: $h_1(\langle a,b\rangle) \neq h_2(\langle a,b\rangle)$. It follows that $(h_1)_a(b) \neq (h_2)_a(b)$ and so $f(h_1) \neq f(h_2)$. f is onto because if $h' : x \to {}^y u$ is given, then $h' = f(h)$ for h defined by $h(\langle a,b\rangle) = h'(a)(b)$.] □

Corollary 2.13 *For all x, y, u, v:*

(i) $(x \approx y \wedge v \approx u) \to {}^x u \approx {}^y v$,
(ii) $(\emptyset \neq x \preceq y \wedge u \preceq v) \to {}^x u \preceq {}^y v$.

2.3 The size of $\mathbb{R}$

Recall that the important property which distinguishes $\mathbb{R}$ from $\mathbb{Q}$ is its *order completeness.* We review the relevant concepts in the appropriately general framework.

Let $\langle X, <\rangle$ be a linearly ordered set. Recall the definition of supremum and infimum in Definition 2.4.

Exercise. Verify that if supremum (infimum) exits for a set $A \subseteq X$, then it is unique. This is true even when $<$ is not linear.

*Exercise.** Show that if a non-empty A is finite, then it has both the supremum and the infimum. [Hint. Use induction on the number of elements.]

We say that A is *bounded below* (zdola omezená) if it has a lower bound, and is *bounded above* (shora omezená) if it has an upper bound. We say that A is bounded (omezená) if it is bounded below and above.

Definition 2.14 *We say that a linearly ordered set $\langle X, <\rangle$ is* order-complete *(úplně uspořádaná) if every non-empty bounded set $A \subseteq X$ has the supremum and the infimum.*

The order-completeness can be formulated in an apparently weaker form, which however turns out to be equivalent.

Lemma 2.15 *Let $\langle X, <\rangle$ be an infinite linearly ordered set without end points.*[7] *The following are equivalent.*

(i) $\langle X, <\rangle$ is order-complete.
(ii) Every non-empty A bounded below has the infimum.
(iii) Every non-empty A bounded above has the supremum.

Proof. (i)→(ii). Choose any $x \in A$, such that $A_x = A \cap \{y \in A \,|\, y < x\}$ is non-empty (such x always exists if A has more than one element; it is has just one element, then this element is both the supremume and the infimum of A). Then A_x is bounded and so has the infimum, which is also the infimum of A, which can be easily verified.

(ii)→(iii). Define B to be the set of all upper bounds of A. Since A is bounded above, B is non-empty and bounded below. By (ii), it has the infimum b. We will show that in fact b is the supremum of A. To show that b is the supremum, we need to check two things:

(1) b is the upper bound of A.
(2) b is the least upper bound of A.

(1). Given $a \in A$, we want to show $a \leq b$: notice that a is a lower bound of B and because b is the greatest lower bound of B, this implies $a \leq b$ as required.

[7] More general formulations are possible. We will use this one because it is of the main interest.

(2). Let b' be another upper bound of A, we want to show $b \leq b'$. Since b' is an upper bound of A, $b' \in B$. Since b is a lower bound of B, it satisfies $b \leq b'$ as required.

Since also (iii)→(ii) by an analogous argument, we can conclude that (iii) implies (i). □

Recall:

Fact 2.16 *$\mathbb{R}$ is the unique (up to isomorphism) order-complete extension of $\mathbb{Q}$ which contains $\mathbb{Q}$ as a dense subset (that is for all $r < r'$ real numbers there is a rational number q such that $r < q < r'$).*

In preperation of Theorem 2.17, we show the following lemma which concerns geometrical progressions (geometrické řady). We call a sequence (a_n) which is of the form $a_0, a_0r, a_0r^2, \dots$ for some a_0 in $\mathbb{R}$ and $r \in \mathbb{R}$ a gemetrical progression. We will only be interested in the case when $0 < r < 1$. If (a_n) is a gemetrical progression, we denote by s_n the sum of its first n elements:

$$s_n = \textstyle\sum_{i=0}^{n-1} a_i = a_0 + a_0r + a_0r^2 + \cdots + a_0r^{n-1}. \tag{2.52}$$

Assuming $0 < r < 1$ one can show

$$s_n = a_0 \cdot \frac{1 - r^n}{1 - r} = \frac{a_0}{1 - r} - \frac{a_0}{1 - r} r^n. \tag{2.53}$$

To see this, argue as follows: Clearly $(1 - r)s_n = s_n - rs_n = (a_0 + a_0r + a_0r^2 + \cdots + a_0r^{n-1}) - (a_0r + a_0r^2 + \cdots + a_0r^n) = a_0 - a_0r^n = a_0(1 - r^n)$.

Recall the following fact about convergence of sequences: If (a_n) and (b_n) are convergent sequences in $\mathbb{R}$ and $q \in \mathbb{R}$ then $(a_n + b_n)$ and (qa_n) are convergent and

$$\lim(a_n + b_n) = \lim a_n + \lim b_n, \text{ and } \lim(qa_n) = q \lim a_n. \tag{2.54}$$

Using this, we can see that:

$$\text{the limit of } (s_n) \text{ exists and } \lim(s_n) = \frac{a_0}{1 - r}. \tag{2.55}$$

Argue as follows: by (2.54), the limit of (s_n), if it exists, is using the expression in (2.52), equal to $\frac{a_0}{1-r} - \frac{a_0}{1-r}\lim(r^n)$. It is obvious that for $0 < r < 1$ the $\lim(r^n)$ exists and is equal to 0. Thus $\lim(s_n) = \frac{a_0}{1-r}$.

Theorem 2.17 *The size of real numbers is the same as the size of the powerset of ω, i.e.*

$$\mathbb{R} \approx \mathscr{P}(\omega).$$

Proof. First recall that $\mathscr{P}(\omega) \approx {}^{\omega}2$, and so it suffices to show that $\mathbb{R} \approx {}^{\omega}2$.

Furthermore, using Cantor-Bernstein theorem 2.6, it suffices to find a 1-1 function $f : \mathbb{R} \to {}^{\omega}2$ and a 1-1 $g : {}^{\omega}2 \to \mathbb{R}$.

Construction of f. Let $\{q_n \mid i \in \omega\}$ be some enumeration of all rational numbers $\mathbb{Q}$ (recall that $\mathbb{Q}$ is countable: there is some bijection $h : \omega \to \mathbb{Q}$, if we set $q_n = h(n)$ we get one such enumeration). Define f so that it assigns to $x \in \mathbb{R}$ a function $f(x) \in {}^{\omega}2$ defined for each $n \in \omega$:

$$f(x)(n) = 1 \text{ if } q_n < x, \text{ or } f(x)(n) = 0 \text{ if } x \leq q_n. \tag{2.56}$$

We need to show that that f is 1-1: if $x \neq y$ are two real numbers then either $x < y$ or $y < x$. Assume without loss of generality that $x < y$. Then by density of $\mathbb{Q}$ in $\mathbb{R}$ there is some $n \in \omega$ such that $x < q_n < y$. This implies that $f(x)(n) = 1$ while $f(y)(n) = 0$; this implies that $f(x) \neq f(y)$ as required.

Construction of g. Let us define an auxiliary function F which to each finite sequence σ of 0s and 1s assigns the value $F(\sigma) = \sum\{1/3^i \,|\, \sigma(i) = 1\}$. Clearly, every $F(\sigma)$ is a rational number. Given a function $x \in {}^{\omega}2$, the set $\{F(x \restriction n) \,|\, n \in \omega\}$ is increasing, that is $F(x \restriction n) \leq F(x \restriction (n+1))$ for every $n \in \omega$. We now claim that for any $x \in {}^{\omega}2$, the set $\{F(x \restriction n) \,|\, n \in \omega\}$ of rational numbers is bounded above (shora omezená) and has therefore a supremum. This follows from the claim in (2.55): the sum of the geometric progression $\frac{1}{3^0} + \frac{1}{3^1} + \frac{1}{3^2} +$ exists and is equal to $\frac{3}{2}$ when we substitute $a_0 = 1$ and $r = \frac{1}{3}$ in (2.55). For any $x \in {}^{\omega}2$ it is clearly true that

$$F(x \restriction n) < \sum\{1/3^i \,|\, i \in \omega\}, \text{ for every } n \in \omega$$

and so $\{F(x \restriction n) \,|\, n \in \omega\}$ is bounded above and by order-completeness of $\mathbb{R}$ it has a supremum. We can now define our function g, for every $x \in {}^{\omega}2$:

$$g(x) = \sup\{F(x \restriction n) \,|\, n \in \omega\}. \tag{2.57}$$

It remain to check that g is 1-1. It is here that we make use of the fact that we have defined the value of $g(x)$ by using a geometrical progression with the factor $\frac{1}{3}$ (one might ask why we have not used factor $\frac{1}{2}$; it will be apparent that it would not work). So assume $x \neq y$ are two sequences in ${}^{\omega}2$. Let n be least such that $x(n) \neq y(n)$; then either $0 = x(n) < y(n) = 1$ or $0 = y(n) < x(n) = 1$. Without loss of generality let the first case be true. Because n is the least where $x(n) \neq y(n)$, $F(x \restriction n) = F(y \restriction n)$. Let us denote this number as a, so that $a = F(x \restriction n) = F(y \restriction n)$. Since $y(n) = 1$, the value of $g(y)$ is at least as big as $a + \frac{1}{3^n}$. To argue that $g(x) < g(y)$, and so $g(x) \neq g(y)$, it suffices to show that $g(x) < a + \frac{1}{3^n}$. Clearly, $g(x) \leq a + \sum\{1/3^{i+1} \,|\, n \leq i, i \in \omega\}$, so it remains to see that

$$\sum\{1/3^{i+1} \,|\, n \leq i, i \in \omega\} < \frac{1}{3^n}. \tag{2.58}$$

Applying again (2.55) with $a_0 = \frac{1}{3^{n+1}}$ and $r = \frac{1}{3}$, the sum $\sum\{1/3^{i+1} \,|\, n \leq i, i \in \omega\}$ is equal to $\frac{a_0}{1-r} = \frac{3}{2}\frac{1}{3^{n+1}} = \frac{1}{2}\frac{1}{3^n}$. Hence $g(x) \leq a + \frac{1}{2}\frac{1}{3^n} < a + \frac{1}{3^n}$, and the proof is finished. □

Exercise. Argue that the argument for g being 1-1 would work for any factor $r = \frac{1}{n}$, where n is a natural number ≥ 3. Argue that the argument would not work with $r = \frac{1}{2}$. [Hint. Study the validity of the inequality in (2.58) for different factors.]

Exercise.* Notice that $g : {}^{\omega}2 \to \mathbb{R}$ has its range included in the closed interval $[0, \frac{3}{2}]$. Let n be any natural number. Modify the definition of $F(\sigma)$ slightly so that the function g is still 1-1 and has its range included in $[0, \frac{1}{n}]$. [Hint. Start the geometrical progression in the construction of $F(\sigma)$ not at $1 = \frac{1}{3^0}$, but at some $\frac{1}{3^m}$ for a suitable m.]

*Exercise** Note that for any $r \in \mathbb{R}$, the function $h_r : \mathbb{R} \to \mathbb{R}$ which maps $x \mapsto x + r$ is 1-1. Use the previous exercise and the existence of such h_r to argue that if $r_1 < r_2$ are two real numbers, then

$$(r_1, r_2) \approx [r_1, r_2) \approx (r_1, r_2] \approx [r_1, r_2] \approx \mathbb{R}.$$

In words, a non-trivial interval (i.e. an interval determined by some $r_1 \neq r_2$) on the real line has the same size as the whole real line.

2.4 The definition of size

We end this introductory section on comparing sizes with the following two apparently simple questions:

Question 2.18 *Is the preordering $\preceq$ linear? That is, given two sets x, y, is it the case that $x \preceq y$ or $y \preceq x$?*

Question 2.19 *Can we assing to each set x another set $|x|$ which will measure its size in the following sense: For every x, y*

(i) $x \approx y \leftrightarrow |x| = |y|$;

(ii) $|x| \approx x$.

We will show that under AC, the Axiom of Choice, the answer to these questions is YES. See the following Section 3.

3 Axiom of Choice and its equivalents

3.1 Axiom of Choice, AC

Given set x we call f a *choice function, výběrová funkce* on x if the domain of f is the set of all non-empty elements of x, i.e. $\mathrm{dom}(f) = x - \{\emptyset\}$ and $f(y) \in y$ for every $y \in \mathrm{dom}(f)$, i.e. f will choose exactly one element from every non-empty element of x.

Recall that *Axiom of Choice* (axiom výběru) is the following statement:

On every set x there exists a choice function.

AC has many equivalent formulations. To state some of these formulations, we first define new notions.

Let r be a binary relation on a set $x \times y$. We say that a function $f \subseteq r$ *uniformizes*(uniformizuje) r if $\mathrm{dom}(f) = \mathrm{dom}(r)$.

Given a finite number of sets $x_0, \ldots, x_n$, recall that $x_0 \times \ldots \times x_n$ is called the *Cartesian product* of $x_0, \ldots, x_n$. $x_0 \times \ldots \times x_n$ contains all $n+1$-tuples $(q_0, \ldots, q_n)$ such that $q_i \in x_i$ for every $0 \leq i \leq n$. We will generalize this notion to infinite families. Let $f : I \to a$ be a function from I onto a. It is a matter of convention that f can also be written as $\langle a_i \,|\, i \in I\rangle$ to indicate the fact that the elements in a are "enumerated" by the elements in I (the letter "I" is for "index set"). This is the indexing already mentioned in (1.25). Notice that $\{a_i \,|\, i \in I\}$ is not the same as $\langle a_i \,|\, i \in I\rangle$: $\{a_i \,|\, i \in I\} = \mathrm{rng}(f) = a$, while $\langle a_i \,|\, i \in I\rangle = f$. Assume now that $I \neq \emptyset$ and $\langle a_i \,|\, i \in I\rangle$ is a function such that a_i is non-empty for every $i \in I$. We define *the Cartesian product* of $\langle a_i \,|\, i \in I\rangle$, denoted $\prod\langle a_i \,|\, i \in I\rangle$ (or $\prod_{i\in I} a_i$), by

$$\prod\langle a_i \,|\, i \in I\rangle = \{f \,|\, f : I \to \bigcup\{a_i \,|\, i \in I\}, (\forall i \in I) f(i) \in a_i\}. \tag{3.59}$$

Lemma 3.1 *The following statements are equivalent:*

(i) Axiom of choice.
(ii) Every binary relation can be uniformized.
(iii) The product $\prod\langle a_i \,|\, i \in I\rangle$ is non-empty for every sequence $\langle a_i \,|\, i \in I\rangle$ such that I is non-empty and every a_i is non-empty.

Proof. (i)→(ii). Let a relation $r \subseteq x \times y$ be given. Clearly, the family $s = \{r''q \,|\, q \in \mathrm{dom}(r)\}$ contains just non-empty set. Let f be a choice function for s; then $\bar{f}$ with domain $\mathrm{dom}(r)$ defined by

$$\bar{f}(q) = f(r''q), \text{ for every } q \in \mathrm{dom}(r) \tag{3.60}$$

uniformizes r.

(ii)→(iii). Let $\langle a_i \,|\, i \in I\rangle$ with I non-empty and every a_i non-empty be given. Let r be a binary relation on $I \times \bigcup\{a_i \,|\, i \in I\}$ defined by

$$\langle i, q\rangle \in r \text{ iff } i \in I \wedge q \in a_i, \text{ for every } \langle i, q\rangle \in I \times \bigcup\{a_i \,|\, i \in I\}. \tag{3.61}$$

It is immediate that every f which uniformizes r is an element of $\prod\langle a_i \,|\, i \in I\rangle$. (ii) thus guarantees that the product $\prod\langle a_i \,|\, i \in I\rangle$ is non-empty.

(iii)→(i). Let x be a set. Let $y = x - \{\emptyset\}$, and assume y is non-empty. Form the product

$$\prod\langle z \,|\, z \in y\rangle = \{f \,|\, f : y \to \bigcup y, (\forall z \in y)\; f(z) \in z\}. \tag{3.62}$$

By our assumption, $\prod\langle z \,|\, z \in y\rangle$ is non-empty. It is easy to see that any function in $\prod\langle z \,|\, z \in y\rangle$ is a choice function on x. □

3.2 Well-ordering principle, WO

Recall that $\langle M, \leq \rangle$ is *partially ordered set*, if $\leq$ is reflexive, transitive and anti-symmetric. We call $\leq$ a non-strict ordering.

Remark 3.2 * *For interested students.* In this subsection, we will consider just sets, not classes (although we will use capital letters to denote these sets for typographical reasons). However, the concepts and results below also apply to some special classes $\langle A, \leq \rangle$ which have the following property: the family $\{a \in A \mid a \leq b\}$ for every $b \in A$ is just a set (even when A is a proper class). Such orders are called *set-like, úzká* and will be sufficient for our needs.

Let us say that $<$ is a strict ordering if it is antireflexive and transitive.

Observation 3.3 *Given $<$, then $\leq$ defined by $a \leq b$ iff $a < b \vee a = b$ is non-strict ordering. And conversely, given $\leq$ then $<$ defined by $a < b$ iff $a \leq b \wedge a \neq b$ is a strict ordering.*

Proof. Exercise. □

Recall the terminology concerning partial orders, given $\langle M, \leq \rangle$ a partial order and $X \subseteq M$:

(3.63)

– x is the *least* (nejmenší) element of X iff $x \in X$ and for all $y \in X$, $x \leq y$.
– x is a *minimal* (minimální) element of X iff $x \in X$ and there is no $y \in X$ such that $y < x$.
– x is a *lower bound* (dolní závora) of X iff for all $y \in X$, $x \leq y$.
– x is the *infimum* of X iff x is the greatest lower bound of X.
– Similarly for *greatest*, *maximal*, *upper bound*, and *supremum.*

Definition 3.4 *We say that partially ordered set $\langle M, \leq \rangle$ is* well-ordered (dobře uspořádaná) *if every non-empty set $x \subseteq M$ has the least element in $\leq$.*

Example. The set $\langle \omega, \leq \rangle$ is well-ordered. The set $\langle \mathbb{Z}, \leq \rangle$ is not well-ordered.

Lemma 3.5 *Every well-ordered set $\langle M, \leq \rangle$ is also linearly ordered.*

Proof. Let $x, y \in M$ be given. $\{x, y\}$ is a subset of M and so must have the least element. Assume x is the least element, then $x \leq y$. If y is the least element, then $y \leq x$. □

Well-ordered sets are useful because they can be easily compared.

Definition 3.6 *Let $\langle A, \leq \rangle$ and $\langle B, \trianglelefteq \rangle$ be two partially ordered sets. We say that they are* isomorphic *and write it as $\langle A, \leq \rangle \cong \langle B, \trianglelefteq \rangle$, or just $A \cong B$ if the orderings $\leq$ and $\trianglelefteq$ are obvious from the context, if there is a bijection $f : A \to B$ such that for all a_0, a_1 in A*

$$a_0 \leq a_1 \leftrightarrow f(a_0) \trianglelefteq f(a_1). \tag{3.64}$$

Definition 3.7 *With $\langle A, \leq \rangle$ and $\langle B, \trianglelefteq \rangle$ as in the previous definition, we say that $f : A \to B$ is an* embedding (vnoření) *if f is 1-1 and for all a_0, a_1 in A*

$$a_0 \leq a_1 \leftrightarrow f(a_0) \trianglelefteq f(a_1). \tag{3.65}$$

Notice that f being an embedding is almost as strong as f being an isomorphism: the only (important) difference is that if f is an embedding, it does not have to be onto.

Exercise. Verify that if $f : A \to B$ is an isomorphism between $\langle A, \leq\rangle$ and $\langle B, \trianglelefteq\rangle$, then $f^{-1} : B \to A$ is an isomorphism between $\langle B, \trianglelefteq\rangle$ and $\langle A, \leq\rangle$.

We say that $a \subseteq A$ is a *initial segment* (with respect to $\leq$) (dolní množina) if for every $b \in a$ and every $c \in A$:

$$c \leq b \to c \in a. \tag{3.66}$$

We say that $f : A \to B$ is an *initial embedding* (počátkové vnoření) if $\mathrm{dom}(f) \subseteq A$ is the initial segment in A and $\mathrm{rng}(f) \subseteq B$ is an initial segment in B, and f is an isomorphism between $\langle \mathrm{dom}(f), \leq\rangle$ and $\langle \mathrm{rng}(f), \trianglelefteq\rangle$.

Lemma 3.8 *Let $\langle A, \leq\rangle$ and $\langle B, \trianglelefteq\rangle$ be wellordered sets. And let f and g be initial embeddings from A to B. Then $f \subseteq g$ or $g \subseteq f$.*

Proof. Since f and g are initial embeddings, it must be true that $\mathrm{dom}(f) \subseteq \mathrm{dom}(g)$, or conversely. W.l.o.g. assume that $\mathrm{dom}(f) \subseteq \mathrm{dom}(g)$ is the case. We will argue that $f \subseteq g$. Assume for contradiction that $f \not\subseteq g$, i.e. that there is $x \in \mathrm{dom}(f)$ such that $f(x) \neq g(x)$. Since $\langle A, \leq\rangle$ is well-ordering, we can take the least x such that $f(x) \neq g(x)$; so fix this x for the rest of the argument. Since $\langle B, \trianglelefteq\rangle$ is well-ordered, it is in particular linearly ordered, and so $f(x) \lhd g(x)$ or $g(x) \lhd f(x)$. W.l.o.g. assume that $f(x) \lhd g(x)$ is the case. Because f and g are initial embeddings, $f(x)$ is in the range of g; let $y \in A$ be such that $g(y) = f(x)$. Because g is an isomorphism on its domain and $g(y) \lhd g(x)$, it must be the case that $y < x$. However, because $y < x$ it must hold that $f(y) < f(x)$, and so $f(y) \neq g(y)$. This contradicts the assumption that x is the least element where f and g are different. □

The following theorem shows that well-ordered sets can be easily compared which will be a key to the importance of ordinal numbers (see Section 4.2).

Theorem 3.9 *Let $\langle A, \leq\rangle$ and $\langle B, \trianglelefteq\rangle$ be well-ordered sets. Then there exists a unique isomorphism F such that F is an isomorphism either between $\langle A, \leq\rangle$ and an initial segment of $\langle B, \trianglelefteq\rangle$ or between an initial segment of $\langle A, \leq\rangle$ and $\langle B, \trianglelefteq\rangle$.*

Proof. Let $\mathscr{S}$ be the following set:

$$\mathscr{S} = \{f \mid f \text{ is an initial embedding from } \langle A, \leq\rangle \text{ to } \langle B, \trianglelefteq\rangle\}. \tag{3.67}$$

We will argue that $\bigcup \mathscr{S} = F$ is the desired isomorphism.

Clearly $F \subseteq A \times B$, and so F is a relation. Also, $\mathrm{dom}(F)$ is an initial segment because $\mathrm{dom}(F) = \bigcup\{\mathrm{dom}(f) \mid f \in \mathscr{S}\}$, and the union of initial segments is always an initial segment. The same argument applies to $\mathrm{rng}(F) = \bigcup\{\mathrm{rng}(f) \mid f \in \mathscr{S}\}$.

We claim that F is a function. Assume for contradiction that F is not a function; then there must be functions $f, g \in \mathscr{S}$ such that $f(x) \neq g(x)$ for some x. However, by Lemma 3.8, if f, g are in $\mathscr{S}$, then either $f \subseteq g$ or $g \subseteq f$. In either case it follows that $f(x) = g(x)$.

We now show that f is 1-1. If x, y are in the domain of F, then x, y must be in the domain of some $f \in \mathscr{S}$ (either $x \leq y$ or $y \leq x$; if $x \leq y$ and $y \in \mathrm{dom}(f)$, then $x \in \mathrm{dom}(f)$ because $\mathrm{dom}(f)$ is an initial segment; similarly for $y \leq x$). This is used to show that F is 1-1: assume $x, y \in \mathrm{dom}(F)$ and $F(x) = F(y)$, then for some $f \in \mathscr{S}$, $F(x) = f(x) = f(y) = F(y)$. Because f is 1-1, $x = y$. In fact, since f is an initial embedding, it follows that $x \leq y \leftrightarrow f(x) \trianglelefteq f(y)$ and so $x \leq y \leftrightarrow F(x) \trianglelefteq F(y)$. This shows that F is an initial embedding.

We now show that either $\mathrm{dom}(F) = A$, or $\mathrm{rng}(F) = B$. Assume for contradiction that $\mathrm{dom}(F) \neq A$ and $\mathrm{rng}(F) \neq B$. We will argue that F can be extended into a strictly larger initial F' embedding from A to B. However, this F' must already be in $\mathscr{S}$ and this will be a contradiction.

Let x be the least element of $A - \mathrm{dom}(F)$ and y the least element of $B - \mathrm{rng}(F)$. It is immediate that $F' = F \cup \{\langle x, y\rangle\}$ is in $\mathscr{S}$ and is strictly bigger than F.

It remains to show that such F is unique. Assume for contradiction there is some $F' \neq F$ which also satisfies the conditions of the Theorem. By Lemma 3.8, it must be the case that $F \subseteq F'$ or $F' \subseteq F$ because both F and F' are initial embeddings. However F' cannot be strictly smaller than F because this would imply that $\mathrm{dom}(F') \neq A$ and $\mathrm{rng}(F') \neq B$. If F' were strictly bigger than F, then $\mathrm{dom}(F')$ would need to bigger than A, or $\mathrm{rng}(F')$ would need to be bigger than B. However, this would mean that F' does not satisfy the conditions of the Theorem. □

Corollary 3.10 *If x and y are sets which can be well-ordered (i.e. there is some $\leq$ such that $\langle x, \leq\rangle$ is a well-ordered set, and some $\leq'$ such that $\langle y, \leq'\rangle$ is a well-ordered set), then x are y are comparable in the relation $\preceq$ comparing sizes:*

$$x \preceq y \text{ or } y \preceq x. \tag{3.68}$$

In other words, the relation $\preceq$ is linear on the class of all well-orderable sets.

Proof. By Theorem 3.9, there is a bijection F between x and an initial segment of y or between an initial segment of x and y. In the first case $F : x \to y$ shows that $x \preceq y$; in the second case $F^{-1} : y \to x$ shows that $y \preceq x$. □

Because we have shown that the concept of a well-ordered set is very useful, we will formulate as a new candidate for an axiom.

Definition 3.11 Well-ordering Principle, WO (princip dobrého uspořádání) *is the following statement*

Every set can be well-ordered.

It is easy to show that this principle implies the Axiom of Choice.

Theorem 3.12 WO *implies* AC. *That is*

$$\mathrm{ZF} \vdash \mathrm{WO} \to \mathrm{AC}. \tag{3.69}$$

Proof. Let x be given. We want to find a choice function f on x. By WO, fix a well-ordering $\leq$ of the set $\bigcup x$. Note that if $a \in x$ then $a \subseteq \bigcup x$. We define for each non-empty $a \in x$:

$$f(a) = \text{ the } \leq\text{-least element of } a. \tag{3.70}$$

It is immediate that f is a choice function. □

We will later show that AC and WO are in fact equivalent. However the proof we will use will require the notion of an ordinal number, and so it will be given in Section 5.3.1.

Corollary 3.13 WO *implies that the relation $\preceq$ for comparing sizes is linear on the universe V.*

Proof. Immediate by Corollary 3.10. □

***

Exercise. Show that if a set x can be wellordered, and $x \approx y$, then also y can be wellordered. [Hint. Let $f : x \to y$ be a bijection. If $<_x$ wellorders x, then $<_y$ defined by $q <_y q' \leftrightarrow f^{-1}(q) <_x f^{-1}(q')$ for $q, q' \in y$ wellorderes y. In fact, $\langle x, <_x\rangle$ and $\langle y, <_y\rangle$ are isomorphic.]

3.3 Principle of Maximality, PM

In this section we formulate yet another form of "choice principle". Its origins are more algebraical.

Let $\langle A, \leq \rangle$ be a partially ordered say. We say that $X \subseteq A$ is a *chain (řetězec)* if the ordering $\leq$ is linear on X, i.e. for all $x, y \in X$, $x \leq y$ or $y \leq x$.

Recall the terminology in (3.63).

Definition 3.14 Principle of Maximality, PM, (princip maximality) *is the following statement. Let $\langle A, \leq \rangle$ be a partially ordered set. Assume further that every non-empty chain $X \subseteq A$ has an upper bound in the ordering $\leq$. Then the following holds: For every $x \in A$, there is a maximal element in $\langle A, \leq \rangle$ above x.*

Example. The condition that every chain must have an upper bound is essential. Consider the set of natural numbers with the usual ordering, $\langle \omega, \leq \rangle$. Then ω itself is a chain which however does not have an upper bound in ω. It follows that we cannot conclude that there a maximal element above every $n \in \omega$ (and indeed there is no maximal element above n).

PM is sometimes referred to as *Zorn's lemma* in the honour of the American mathematician (algebraist and group theorist) Max A. Zorn who first used this principle in 1935.

PM is stronger than WO (see Theorem 3.15), although the proof is a bit less straightforward than the proof that WO implies AC. As with AC and WO, we will later show that WO and PM are in fact equivalent. See Section 5.3.1. The equivalance of all these independently discovered notions is for practical considerations a powerful reason for believing that these notions are intuitively valid (true).

Theorem 3.15 PM *implies* WO. *That is*

$$\mathrm{ZF} \vdash \mathrm{PM} \to \mathrm{WO}. \tag{3.71}$$

The proof is included in Section 3.4.2

3.4 Optional topics

3.4.1 An application of PM – ultrafilters

We show another application of PM (or equivalently of AC) in the construction of very useful and important objects, the so called *ultrafilters.*

Definition 3.16 *Let A be a set. A system $F \subseteq \mathscr{P}(A)$ is called a* filter *iff:*

(i) $A \in F$,

(ii) If $X \in F$ and $X \subseteq Y$ then $Y \in F$,

(iii) If $X \in F$ and $Y \in F$ then $X \cap Y \in F$.

A filter F is called a *proper* (vlastní) filter iff $\emptyset \notin F$.

Exercise. Let F be a filter, then: F is not proper iff $\emptyset \in F$ iff $F = \mathscr{P}(A)$.

Lemma 3.17 $F \subseteq \mathscr{P}(A)$ *is a filter iff:*

(i) $A \in F$.

(ii) For all $X, Y \subseteq A$,

$$X \cap Y \in F \leftrightarrow X \in F \wedge Y \in F. \tag{3.72}$$

Proof. If F is a filter according to Definition 3.16, then we need to show that $X \cap Y \in F$ implies that $X \in F$ and $Y \in F$. But clearly, $X \cap Y \subseteq X$ and $X \cap Y \subseteq Y$, and so by (iii) of Definition 3.16, $X \in F$ and $Y \in F$.

Conversely, if F satisfies conditions (i) and (ii) of the Lemma, we need to show that if $X \in F$ and $X \subseteq Y$ then $Y \in F$. But clearly, $X = X \cap Y \in F$ and so by (ii), both X and Y must be in F. □

Example. The following set $\mathscr{F} \subseteq \mathscr{P}(\omega)$ is an important filter, the so called *Frachet* filter:

$$\mathscr{F} = \{X \subseteq \omega \,|\, \omega \setminus X \text{ is finite}\}. \tag{3.73}$$

Exercise. Verify that $\mathscr{F}$ is indeed a proper filter.

Definition 3.18 *We say that a system $E \subseteq \mathscr{P}(A)$ has the* finite intersection property, FIP *(E je centrovaný systém) if for every $n \in \omega$ and every family $e_0, \ldots, e_n \subseteq E$ it holds that*

$$e_0 \cap \ldots \cap e_n \neq \emptyset. \tag{3.74}$$

Lemma 3.19 *Every $E \subseteq \mathscr{P}(A)$ with FIP can be extended into a proper filter.*

Proof. Define F as follows:

$$F = \{X \subseteq A \,|\, (\exists n \in \omega)(\exists e_0, \ldots, e_n)\, e_0 \cap \ldots \cap e_n \subseteq X\}. \tag{3.75}$$

It is immediate that F contains E and F is a proper filter (Exercise). [Hint. To verify that F is closed under intersection, i.e. that for $X, Y \in F$ we have that $X \cap Y \in F$, argue that if $X \supseteq e_0 \cap \ldots \cap e_n$ and $Y \supseteq e'_0 \cap \ldots \cap e'_m$ then $X \cap Y \supseteq e_0 \cap \ldots \cap e_n \cap e'_0 \cap \ldots \cap e'_m$.] □

Definition 3.20 *A proper filter $F \subseteq \mathscr{P}(A)$ is called an* ultrafilter *if F is a filter and moreover:*

$$\textit{For all } X \subseteq A, \textit{ either } X \textit{ or } A \setminus X \textit{ is in } F. \tag{3.76}$$

Example. The Frachet filter $\mathscr{F}$ on ω is not an ultrafilter. [Hint. Consider the set of all even numbers.]

We say that a proper filter F is *maximal* if it is a maximal proper filter with respect to the relation $\subseteq$: i.e. there is no proper filter F' such that $F' \supseteq F$ and $F' \neq F$.

Lemma 3.21 *Let $F \subseteq \mathscr{P}(A)$ be a proper filter. Then the following are equivalent:*

(i) F is maximal.
(ii) For every $X \subseteq A$ with $X \notin F$ there is some $Y \in F$ such that $X \cap Y = \emptyset$.

Proof. (i)→(ii). So let X be a subset of A which is not in F. For contradiction assume that for all $Y \in F$, the intersection $X \cap Y$ is non-empty. Then the set $F \cup \{X\}$ has FIP because if $X_1, \ldots, X_n$ are elements from F, then also $X_1 \cap \ldots \cap X_n$ is in F (because F is a filter), and by our assumption $(X_1 \cap \ldots \cap X_n) \cap X \neq \emptyset$. By Lemma 3.21, there is a proper filter F' which contains $F \cup \{X\}$. Since $F' \supseteq F$ and $F' \neq F$, F' contradicts the initial assumption that F is maximal. There it follows that there must exists some $Y \in F$ such that $X \cap Y = \emptyset$.

(ii)→(i). Assume F is a filter and $F' \supseteq F$ is also a filter. We will show that F' is non-proper, and hence F is maximal. If $X \in F' \setminus F$, then by our assumption there is some $Y \in F$ such that $X \cap Y = \emptyset$. Because F' is a filter, $\emptyset = X \cap Y \in F'$. This means that F' is not a proper filter, and so F is a maximal proper filter. □

Lemma 3.22 *For every $F \subseteq \mathscr{P}(A)$ the following are equivalent:*

(i) F is an ultrafilter.
(ii) F is maximal.

Proof. (i)→(ii). Let F be an ultrafilter. We want to show that F is maximal. Let X not in F be given. By above Lemma 3.21, it suffices to find $Y \in F$ such that $X \cap Y = \emptyset$. Since F is an ultrafilter, it follows that $-X = A \setminus X$ is in F, and $X \cap -X = \emptyset$.

(ii)→(i). Let F be maximal. Assume that X is not in F. We want to show that $-X = A \setminus X$ must be in F. By Lemma 3.21, there is some $Y_X \in F$ such that $Y_X \cap X = \emptyset$, which is equivalent to $Y_X \subseteq -X$. This immediately implies that $-X$ is in F (by the definition of filter). □

Theorem 3.23 *(PM) Every $E \subseteq \mathscr{P}(A)$ with FIP can be extended into an ultrafilter.*

Proof. Let us denote

$$\mathbb{F} = \{F \mid F \text{ is a proper filter on } A\}. \tag{3.77}$$

We first show that $(\mathbb{F}, \subseteq)$ satisfies the condition that every $\subseteq$-chain has an upper bound. Let $\mathscr{C} \subseteq \mathbb{F}$ be a chain, i.e. a linearly ordered subfamily of $\mathbb{F}$. We will argue that

$$F = \bigcup \mathscr{C} \tag{3.78}$$

is a proper filter which is the upper bound (in fact a supremum) of $\mathscr{C}$. Clearly: $A \in F$, $\emptyset \notin F$, and $X \in F$ implies $Y \in F$ for every X, Y. It remain to show the intersection property. Let X, Y be in F and fix F_X and F_Y in $\mathscr{C}$ such that $X \in F_X$ and $Y \in F_Y$; since $\mathscr{C}$ is a chain, we have either $F_X \subseteq F_Y$ or $F_Y \subseteq F_X$. Without loss of generality, assume that $F_X \subseteq F_Y$ is true. Then X, Y are in F_Y, and since F_Y is a filter $X \cap Y$ is in F_Y and then also in F.

Let $E \subseteq \mathscr{P}(A)$ be a system with FIP. By Lemma 3.19, E can be extended into a proper filter F. By Principle of Maximality (PM), there is a maximal element above F in the ordering $(\mathbb{F}, \subseteq)$ of all proper filters on A. Let $U \supseteq F$ be a maximal element (there may be more of them). By Lemma 3.22, this U is the desired ultrafilter extending E. □

*Exercise** Show that a proper filter F on A is an ultrafilter iff it satisfies for all $X, Y \subseteq A$:

$$X \cup Y \in F \leftrightarrow X \in F \vee Y \in F. \tag{3.79}$$

3.4.2 PM implies WO

Theorem 3.24 PM *implies* WO*. That is*

$$\mathrm{ZF} \vdash \mathrm{PM} \to \mathrm{WO}. \tag{3.80}$$

Proof. Let A be a set. We want to find an ordering $\leq$ with $\mathrm{dom}(\leq) = A$ such that $\leq$ is a well-ordering.

Define

$$\mathscr{S} = \{R \subseteq A \times A \mid R \text{ is a well-ordering on } \mathrm{dom}(R)\} \tag{3.81}$$

and the ordering $\trianglelefteq$ on $\mathscr{S}$ by: $R \trianglelefteq R'$ iff $R \subseteq R'$ and R' end-extends R, i.e. all elements in $\mathrm{dom}(R') - \mathrm{dom}(R)$ come after all elements of $\mathrm{dom}(R)$ in the ordering R'. This means that whenever $x \in \mathrm{dom}(R)$ and $y \in \mathrm{dom}(R') - \mathrm{dom}(R)$, we have $\langle x, y \rangle \in R'$.

$\mathscr{S}$ is non-empty because it contains at least $\emptyset$ ($\emptyset$, being empty, is trivially a well-ordering on its domain $\emptyset$). We want to apply PM to $\langle \mathscr{S}, \trianglelefteq \rangle$.

To apply PM we need to check that every (non-empty) chain in $\langle \mathscr{S}, \trianglelefteq \rangle$ has an upper bound. Let a chain $X \subseteq \mathscr{S}$ be given. We argue that $\bigcup X$ is in $\mathscr{S}$ and is an upper bound of X in the ordering $\trianglelefteq$.

First notice that $\bigcup X$ is a binary relation on A since it is a union of binary relations; and also $\mathrm{dom}(\bigcup X)$ is the union

$$\bigcup\{\mathrm{dom}(R) \mid R \in X\}. \tag{3.82}$$

By reflexivity of R's, we also have that $\mathrm{dom}(R) = \mathrm{rng}(R)$ and so $\mathrm{dom}(\bigcup X) = \mathrm{rng}(\bigcup X)$.

To show that $\bigcup X$ is in $\mathscr{S}$ and an upper bound of X in the ordering $\trianglelefteq$ we need to verify:

(a) *Exercise.* Verify that $\bigcup X$ is indeed a partial order on $\mathrm{dom}(\bigcup X)$.

(b) We need to show that $\bigcup X$ is a well-ordering on $\mathrm{dom}(\bigcup X)$. Let a non-empty $Y \subseteq \mathrm{dom}(\bigcup X)$ be given. Choose arbitrary R in X such that $\mathrm{dom}(R) \cap Y$ is non-empty. Because R is a well-ordering on its domain, $\mathrm{dom}(R) \cap Y$ has the least element in the ordering R; denote this element r:

$$r = \text{ the } R\text{-least element in } \mathrm{dom}(R) \cap Y. \tag{3.83}$$

We argue that r is in fact the least element in Y in the ordering $\bigcup X$. Let $y \in Y$ be arbitrary, we want to show that $\langle r, y \rangle$ is in $\bigcup X$. Let R' be a relation such that $y \in \mathrm{dom}(R')$. If $R' \subseteq R$, then $y \in \mathrm{dom}(R)$ and so by (3.83), $\langle r, y \rangle \in R$ and hence $\langle r, y \rangle \in \bigcup X$. If $R \subsetneq R'$, and $y \notin \mathrm{dom}(R)$, then we use the fact that R' end-extends R to conclude that $\langle r, y \rangle \in R'$ and so $\langle r, y \rangle \in \bigcup X$.

(c) Lastly, we need to check that $\bigcup X$ is an upper bound of X in the relation $\trianglelefteq$. If $R \in X$, then $R \subseteq \bigcup X$. It is also easy to check that $\bigcup X$ end-extends the relation R, and so $R \trianglelefteq \bigcup X$.

By PM there is a maximal element above $\emptyset \in \mathscr{S}$. Let $R \in \mathscr{S}$ be one such element. We will argue that $\mathrm{dom}(R) = A$, and this will prove the theorem. Assume for contradiction that there is some $a \in A$ not in $\mathrm{dom}(R)$. Define R' by

$$R' = \{\langle x, y \rangle \mid \langle x, y \rangle \in R \vee (x \in \mathrm{dom}(R) \wedge y = a) \vee (x = a \wedge y = a)\}, \tag{3.84}$$

or equivalently, where we denote $B = \mathrm{dom}(R) \cup \{a\}$:

$$R' = R \cup (B \times \{a\}). \tag{3.85}$$

It is easy to check that $R' \in \mathscr{S}$ and that R' is strictly bigger than R in the ordering $\trianglelefteq$. This is a contradiction with R being a maximal element in $\mathscr{S}$. □

4 Ordinal numbers

4.1 Motivation

Recall the definition of a wellordered set in Definition 3.4. Wellordered sets are very important in set theory because they can be used to generalize constructions by recursion or induction on ω (natural numbers) to transfinite lengths (i.e. longer than ω). Such constructions – constructions by *transfinite recursion*, see Section 5.2 – are made possible by Theorem 3.9, which is very special and holds only for wellordered sets.

4.2 Definition of an ordinal number and the class of ordinal numbers

The goal, and the motivation, in the definition of an ordinal number is the following.

(1) A ordinal will be a set of the form $\langle \alpha, < \rangle$, where $<$ is a strict wellordering on α.

(2) We will also require that
$$\alpha = \{\beta < \alpha \,|\, \beta \text{ is an ordinal}\}.$$

(3) We will want that the class $\{\alpha \,|\, \alpha \text{ is an ordinal}\}$ is itself wellordered by a strict ordering $<$. And moreover the ordering $<$ on the class of all ordinals will be universal in the sense that for every ordinal $\langle \alpha, < \rangle$, the ordering on α is just the restriction of the ordering on the class of all ordinal numbers.

It turns out that the simplest way how to do this is to use the relation $\in$ as the wellordering.

Before we start to define the notion of an ordinal, we will state a simple, but useful lemma. But a definition first.

Definition 4.1 *Let R be a binary relation. The restriction of the relation R to a class X, in symbols R_X, is a relation on X defined by*

$$R_X = \{(x, y) \,|\, x \in X \wedge y \in X \wedge (x, y) \in R\} = R \cap X^2.$$

Lemma 4.2 *Let $<$ be a strict wellordering on a class A. Then for every $B \subseteq A$, the restriction $<_B$ of $<$ to B is a strict wellordering on B.*

Proof. Recall that $<$ is a strict wellordering on A if it is a irreflexive ($a \not< a$ for every $a \in A$) and transitive relation, and moreover every non-empty subset $x \subseteq A$ has the least element in the ordering $<$.

Clearly $b \not<_B b$ for every $b \in B$ because b is also in A and $b \not< b$. If a, b, c are in B and $a <_B b$ and $b <_B c$, then also $a < b$ and $b < c$ and hence $a < c$. Since both a and c are in B, $a <_B c$.

If $x \subseteq B$ is a non-empty subset of B, then it is also a subset of A, and hence has the least element $a \in x$ in the ordering $<$. This a is the least element of x in $<_B$, which is easy to verify. □

Definition 4.3 *We say that a class X is* transitive, tranzitivní *if*

$$(\forall x)\,[x \in X \rightarrow x \subseteq X]. \tag{4.86}$$

This means that if X is transitive, and for sets x, y holds $x \in y \in X$, then $x \in X$.

Examples. $\emptyset$, and V are transitive classes.

Lemma 4.4 *X is transitive iff $\bigcup X \subseteq X$.*

Proof. ($\to$). If $x \in \bigcup X$, then there is $y \in X$ such that $x \in y$, this by transitivity of X implies $x \in X$.

($\leftarrow$). We wish to show that $x \in y \in X$ implies $x \in X$. If $x \in y \in X$, then $x \in \bigcup X$, and hence by the assumption $\bigcup X \subseteq X$, x must be in X. □

Lemma 4.5 *(1) If X, Y are transitive classes then $X \cap Y$ and $X \cup Y$ are transitive classes.*
(2) If every $x \in X$ is itself a transitive set, then $\bigcup X$ and $\bigcap X$ are transitive classes.
(3) If X is a transitive class then the relation $\in$ restricted to X is transitive iff every $x \in X$ is a transitive set.

Proof. Ad (1). If x is in $X \cap Y$ and $y \in x$, then $y \in X$ and $y \in Y$ by transitivity of Y, X, and so $y \in X \cap Y$. If x is in $X \cup Y$, then it is at least in one of the two sets X, Y. Assume that $x \in X$, and $y \in x$, then $y \in X$, and so also $y \in X \cup Y$.

Ad (2). This is just a generalization of (1). If $x \in \bigcap X$ and $y \in x$, then because every $z \in X$ is transitive, y must be in every z, and hence $y \in \bigcap X$. Similarly for $\bigcup X$.

Ad (3). Let $\in_X$ be the restriction of $\in$ to X. Notice that if X is transitive and $x \in X$, then for all y:

$$y \in x \leftrightarrow y \in_X x. \tag{4.87}$$

If X fails to be transitive, we have just the following for $x \in X$ and arbitrary y:

$$y \in_X x \to y \in x. \tag{4.88}$$

Assume that $\in_X$ is transitive on X: that is if x, y, z are elements of X, and $x \in_X y \wedge y \in_X z$, then also $x \in_X z$. Let $x \in X$ be arbitrary, we want to show that x is transitive: this means that if $y_0 \in y_1 \in x$, then $y_0 \in x$. However, this follows from the transitivity of the relation $\in_X$ on X, and the fact that y_0 and y_1 are in X (by transitivity of X), and so in the domain of $\in_X$.

Conversely, let elements x, y, z in X satisfy $x \in_X y \wedge y \in_X z$, we want to show that $x \in_X z$. However, this follows from the transitivity of z. □

Exercise.

1. Show that the following weakening of Lemma 4.5(3) is not true: "If X is transitive and $y \in X$, then y is transitive." [Hint: for example $X = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ and $y = \{\{\emptyset\}\}$.]
2. Every transitive set must contain the emptyset $\emptyset$. [Hint: Let X be a transitive class. From the Axiom of Foundation we know that X must have a minimal element – some $a \in X$ such that no $b \in a$ is in X. However any such b shows that X is not transitive. It follows that a must not contain any elements, hence a is the emptyset $\emptyset$.]

The following definition of an ordinal can be motivated as follows. We wish to define a representatives of wellordered sets of a certain type, that is a pair $\langle a, < \rangle$, where a is a set and $<$ a wellordering on a: that is $<$ should be irreflexive and transitive, and every non-empty subset of a should have the least element in $<$. The simplest way one might consider is to take a pair of the form $\langle a, \in_a \rangle$ for some non-empty set a. The relation $\in$ is in general not a partial ordering, but it does satisfy some properties which suggest that with a bit of tweaking it might work. Realize:

Motivation.

(1) The relation $\in$ is irreflexive by the Axiom of Foundation, that is for all x, $x \notin x$
(2) The relation $\in$ is not in general transitive, but recall Lemma 4.5(3): if a is transitive and every element of a is itself transitive, then $\in_a$ is transitive.
(3) Again by the Axiom of Foundation we know that every non-empty subset b of a has a minimal element in the relation $\in$ (recall that b can have more minimal elements – so this is weaker then having the least element, which is alway unique). However, with transitivity, we will be able to show that in fact this minimal element is unique, and it is the least element.

Remark 4.6 **For interested students.* The above Motivation it really just a motivation: Ordinal numbers can – and we will do it this way – be developed without the Axiom of Foundation. It follows that all properties below which we show about ordinals are provable in ZF minus the Axiom of Foundation. This means that the class ORD is well-founded even if the universe V is not. This has the nice consequence that even if the Axiom of Foundation is not true, we can still build inside of V the class ORD and show that it generates a subclass $V' \subseteq V$, where V' satisfies the Axiom of Foundation. See Theorem 5.21. This in particular implies that if the theory (ZF $-$ Axiom of Foundation) is consistent, so is ZF.

Definition 4.7 *A set x is called an* ordinal, ordinál, *or* ordinal number, ordinální číslo *if x is a transitive set and the restriction*

$$\in_x = \in \cap x^2$$

is a strict wellordering on x.

The class of all ordinals will be denoted ORD:

$$\text{ORD} = \{x \,|\, x \text{ is an ordinal number}\}$$

Exercise. Show that $\emptyset$ is an ordinal.

In order to show some other examples, we will verify some simple properties of ordinal numbers.

Lemma 4.8 *Let x be an ordinal number, then:*

(1) (ZF $-$ Axiom of Foundation) $x \notin x$.
(2) $x \cup \{x\} =_{df} x + 1$ is an ordinal number.

Proof. Ad (1). This is trivially implied by the Axiom of Foundation. We shall however show that (1) is true even without the Axiom of Foundation. We will argue by contradiction. We know that $\in$ a strict ordering on x, this means that for every $y \in x$,

$$y \notin y. \tag{4.89}$$

If we assume for contradiction that x is in x, it must satisfy (4.89), and hence $x \notin x$. A contradiction.

Ad (2). We first verify transitivity of $x + 1$. Fix arbitrary z, y such that $z \in y \in x + 1$. If y is in $x + 1$, then it is either in x or is equal to x. In the first case, we have $z \in x$ by the transitivity of x, and so $z \in x + 1$. In the second case we get immediately $z \in y = x$, and again $z \in x + 1$.

Now we verify that $\in_{x+1}$ is a wellordering on $x + 1$. First check that

$$(\forall a)[a \in x \leftrightarrow a \in_{x+1} x], \text{ and } \in_{x+1} = \in_x \cup \{(a, x) \,|\, a \in x\}. \tag{4.90}$$

The first claim in (4.90) holds because $x + 1$ is transitive. The second claim in (4.90) holds because $x + 1$ contains only one more element besides the elements in x – the set x itself (and this set x can only occur on the right side of the $\in$ relation: if $x \in a$ were true for some $a \in x$, then we obtain $x \in x$, which contradicts (1) of the present lemma). The relation $\in_{x+1}$ is clearly irreflexive (by (1) of the present lemma), and also transitive (by Lemma 4.5(3)), and so it is a strict ordering. (4.90) implies that $\in_{x+1}$ is an ordering which extends the ordering $\in_x$ by adding the new element x above all the elements in x (above in the sense of $\in_{x+1}$). This implies that $\in_{x+1}$ is a wellordering: if $y \subseteq x + 1$ is non-empty, then either $x \cap y$ is also non-empty, in which case the least element of $x \cap y$ in $\in_x$ is also the least element in $\in_{x+1}$, or $y = \{x\}$, in which case x is the least element. □

Examples of ordinal numbers. Recall our definition of natural numbers

$$\mathbb{N} = \omega = \bigcap \{x \,|\, x \text{ is an inductive set}\}$$

in (1.28). We will show that all elements $n \in \omega$ are ordinal numbers.

Lemma 4.9 *All elements in $\mathbb{N}$ are ordinal numbers.*

Proof. We will show that if a is an inductive set, then $\mathrm{ORD} \cap a$ is an inductive set. It follows that $\mathbb{N} = \bigcap\{x \,|\, x \text{ is an inductive set}\} \subseteq a \cap \mathrm{ORD}$, and hence every element of $\mathbb{N}$ is in ORD.[8]

First realize that $\emptyset$ is an ordinal number, and so $\emptyset \in \mathrm{ORD} \cap a$. If q is in $\mathrm{ORD} \cap a$, then by definition of inductiveness, $q \cup \{q\} \in a$; we need to show that also $q \cup \{q\} \in \mathrm{ORD}$. But this follows from Lemma 4.8(2). □

We will now verify some other properties of ORD.

Lemma 4.10 ORD *is a transitive class.*

Proof. Let x be an ordinal. We wish to show that for an arbitrary y: if $y \in x$, then $y \in \mathrm{ORD}$. In words, every element of an ordinal is itself an ordinal.

Since $\in_x$ on x is a transitive relation (because every ordering is a transitive relation), we know by Lemma 4.5(3) that every element of x is a transitive set, in particular y is transitive.

We finish the proof by arguing that $\in_y$, the restriction of $\in$ to y, is a strict wellordering of y. The relation $\in_x$ is by our assumption a strict wellordering of x. By transitivity of x we know that $y \subseteq x$, and by Lemma 4.2, we can conclude that $\in_x$ restricted to y is a strict wellordering. Now it is enough to argue that $\in_x$ restricted to y is in fact the relation $\in_y$. Consider the following fact, which follows from the transitivity of y and x: for every $a \in y$ and every b,

$$b \in_y a \leftrightarrow b \in_x a \leftrightarrow b \in a. \tag{4.91}$$

We want to show that $\{\langle c, d\rangle \,|\, c \in y \wedge d \in y \wedge c \in_x d\}$ (the restriction of $\in_x$ to y) is identical with $\{(c, d) \,|\, c \in y \wedge c \in y \wedge c \in d\}$ (the restriction of $\in$ to y). However, this follows from (4.91). □

The following Fact and Theorem are given without proofs. Proofs can be found in Section 4.5.1.

Fact 4.11 *The following holds for all ordinals x and y:*

(1) $x \cap y$ is an ordinal,
(2) $x \in y \leftrightarrow x \subsetneq y$

Theorem 4.12 *The relation $\in_{\mathrm{ORD}} = \in \cap \mathrm{ORD}^2$ is a strict wellordering of the class* ORD.

Corollary 4.13 *The class* ORD *is a proper class.*

Proof. If ORD were a set, then by Lemma 4.10 and by Theorem 4.22, ORD would be an ordinal number. This would mean that $\mathrm{ORD} \in \mathrm{ORD}$. However this is impossible by Lemma 4.8(1). □

Notation. From now on, we will denote ordinal numbers by small case Greek letters from the beginning of the alphabet: $\alpha, \beta, \gamma, \ldots$. We shall also write $<$ for the ordering on the ordinal numbers: we will write $\alpha < \beta$ instead of $\alpha \in_{\mathrm{ORD}} \beta$. Under this notation, the ordering $<$ restricted to α is a wellordering on α: $\langle \alpha, <_\alpha \rangle$. To further simplify notation, we will no longer keep track of the distinction between $\in_\alpha$ and $\in$, resp. $<$ and $<_\alpha$. For instance we will write $\langle \alpha, < \rangle$ instead of $\langle \alpha, \in_\alpha \rangle$.

4.3 Ordinal numbers as types or representatives of wellordered sets

In Motivation section 4.1, we have define the equivalence between wellordered sets using the notion of an isomorphism between sets. In this section we show that ordinal numbers are natural

[8] We do not know yet whether ORD is a set or not (in fact it is not, see Corollary 4.13). However, for every set a, $a \cap \mathrm{ORD}$ is a set, and so is an element of $\{x \,|\, x \text{ is an inductive set}\}$ which implies $\mathbb{N} \subseteq a \cap \mathrm{ORD}$.

representatives of equivalence classes for the equivalence $\cong$ (in every equivalence class of wellordered sets, there is exactly one ordinal number, and different equivalence classes are represented by different ordinals).

Lemma 4.14 *A subset $x \subseteq \mathrm{ORD}$ is an ordinal iff x is a transitive set.*

Proof. By Theorem 4.22 ORD itself is wellordered by $\in$, and as a subset, x is wellordered by the relation $\in$ as well. By definition of an ordinal applied to x wellordered by $\in$, we immediately obtain that x is a transitive set iff x is an ordinal. □

Theorem 4.15 *Every wellordered set $\langle a, <\rangle$ is isomorphic to exactly one ordinal.*

Proof. By Theorem 3.9, there is an isomorphism F from $\langle a, <\rangle$ onto an initial segment of ORD, or from $\langle \mathrm{ORD}, \in_{\mathrm{ORD}}\rangle$ onto an initial segment of $\langle a, <\rangle$. The second possibility cannot happen because this would mean that ORD is a set (recall that by Schema of Replacement, a range of a function over a set is again a set: and certainly $F^{-1}[b]$, for any $b \subseteq a$, is a set). So there is an isomorphism F from $\langle a, <\rangle$ onto an initial segment of ORD. Again by Schema of Replacement, this initial segment must be a set because a is a set. Since every initial segment of ORD is a transitive set, we obtain by Lemma 4.14 that this initial segment is an ordinal. Let as denote this ordinal α. Then $\langle a, <\rangle$ and $\langle \alpha, \in\rangle$ are isomorphic.

By Theorem 3.9, this isomorphism F is unique and consequently this α the unique ordinal isomorphic with $\langle a, <\rangle$. □

Corollary 4.16 *No two distinct ordinals are isomorphic.*

4.4 Properties of ordinal numbers

We know from Theorem 4.22 that $\in$ is a strict wellordering on the class ORD. The following lemma provides some more information about this ordering (recall that we have decided to write $<$ instead of $\in$).

Lemma 4.17 *The following properties hold:*

(1) 0 *is the least element of* ORD.
(2) If A is a set of ordinal numbers, then $\bigcup A$ is also an ordinal number and it is the supremum of A in the wellordering $<$.
(3) If α is an ordinal, then $\alpha + 1$ is the least ordinal greater than α. We call $\alpha + 1$ the successor, následník *of α, and α the* predecessor, předchůdce *of $\alpha + 1$.*

Proof. Ad (1). Obvious.

Ad (2). First note that the set $\bigcup A$ is transitive by Lemma 4.5, and by Lemma 4.14 it is an ordinal. Let us denote $\alpha = \bigcup A$. α is clearly an upper bound of A: if $\beta \in A$, then $\beta \subseteq \bigcup A$ which by Lemma 4.19 implies $\beta \leq \alpha$. We now show that α is the least upper bound. This is equivalent to showing that if $\gamma < \alpha$, then $\gamma < \beta$ for some $\beta \in A$. If $\gamma < \alpha$, then there is some $\beta \in A$ such that $\gamma \in \beta$. Since $\in$ is the ordering $<$, this means $\gamma < \beta$ as required.

Ad (3). By Lemma 4.8, $\alpha+1$ is an ordinal number. Since trivially $\alpha \in \alpha+1$, we obtain $\alpha < \alpha+1$. If $\beta < \alpha + 1$, then either $\beta \in \alpha$ (which is equivalent to $\beta < \alpha$) or $\beta = \alpha$; this means that there is no ordinal between α and $\alpha + 1$. □

We distinguish two types of ordinal numbers:

Definition 4.18 *We say that an ordinal α is a* successor (izolovaný) *ordinal if it is of the form $\beta+1$ for some β. We say that α is a* limit *ordinal if it is greater than* 0 *and has no predecessor,i.e. if $\beta < \alpha$, then $\beta + 1 < \alpha$. 0 is a special ordinal: it is not successor, nor limit.*

It follows that all ordinal numbers greater than 0 are divided into two disjoint classes: successor and limit ordinals (if α is not a successor, then $\beta < \alpha$ implies that $\beta+1 < \alpha$, otherwise $\beta+1 = \alpha$ which contradicts the assumption that α is not a successor). Note: the existence of limit ordinals follows from the existence of a inductive set in the Axiom of Infinity; see Subsection 4.5.2 for details.

4.5 Optional topics

4.5.1 Some proofs concerning ordinal numbers

Lemma 4.19 *The following holds for all ordinals x and y:*

(1) $x \cap y$ is an ordinal,
(2) $x \in y \leftrightarrow x \subsetneq y$

Proof.

Ad (1). Follows from the fact that $x \cap y$ is a transitive set by Lemma 4.5, which is a subset of x (and y), and so the ordering on $x \cap y$ is just the restriction of the ordering on x to $x \cap y$ (see (4.91) for more details).

Ad (2). ($\rightarrow$). Let $x \in y$ be given. By transitivity of y, this implies $x \subseteq y$. So it remain to show that $x \neq y$. However this follows by contradiction: if $x = y$, then because we assume $x \in y$, this would imply $x \in x$. Contradiction with Lemma 4.8(1).

($\leftarrow$). Let us assume $x \subsetneq y$, and show $x \in y$. We will argue that the least element of $y - x$ (note that $y - x$ is non-empty by the fact that $x \subsetneq y$) is the set x. As the least element is in y, this will prove the lemma.

Let $z \in (y - x)$ be the least element of $y - x$. We will show $x \subseteq z$ and $z \subseteq x$, which implies $x = z$ as required.

Let $q \in x$ be given, we will show that $q \in z$ is true. Since $x \subseteq y$, q is in y. From the linearity of the ordering $\in_y$ on y ($\in_y$ is a wellordering, which implies linearity) we obtain that one these must be true:

$$q \in z \vee q = z \vee z \in q \tag{4.92}$$

$q = z$ cannot be true because $z \notin x$, while $q \in x$. $z \in q$ cannot be true either: by transitivity of x, $z \in q \in x$ would imply $z \in x$. Hence $q \in z$ must be true.

Conversely, if $q \in z$, we want to show that $q \in x$. By transitivity of y, $q \in z$ implies $q \in y$. Since $(y - x) \cup x = y$, q is either in $(y - x)$ or in x. Assume for contradiction that q is in $y - x$. Then because z is the least element in $y - x$, it must hold $z \in q$, but this is a contradiction as we showed above in (4.92). It follows that $q \in x$, which finishes the proof of the lemma. □

Before we prove Theorem 4.22, we first state the following easy lemma.

Lemma 4.20 *Let $(X, <)$ be a partially ordered class and assume that $<$ is a linear ordering. Let $Y \subseteq X$ be a subset of X. Then if Y has a minimal element, then this element is unique and is the least element of Y.*

Proof. By contradiction. If $y \neq y'$ are two minimal elements of Y, then by linearity of $<$, either $y < y'$ or $y' < y$ must be true. If $y < y'$ is true, then y' cannot be minimal; if $y' < y$ is true, then y cannot be minimal. A contradiction. □

Corollary 4.21 *If $(X,<)$ is a linear partial ordering such that every non-empty subset $Y \subseteq X$ has a minimal element, then $(X,<)$ is in fact a wellordering.*

Proof. Obvious by Lemma 4.20. □

Theorem 4.22 *The relation $\in_{\text{ORD}} = \in \cap \text{ORD}^2$ is a strict wellordering of the class* ORD.

Proof. By Lemma 4.8(1), the relation $\in_{\text{ORD}}$ is irreflexive. By Lemma 4.10, ORD is a transitive class and by the definition of an ordinal, all the elements in ORD are transitive sets; by Lemma 4.5(3) it follows that the relation $\in_{\text{ORD}}$ is transitive on ORD. This show that the relation $\in_{\text{ORD}}$ is a strict partial ordering of ORD.

By Corollary 4.21, it suffices to show the two following properties:

(1) $\in_{\text{ORD}}$ is linear,
(2) Every non-empty subset $X \subseteq \text{ORD}$ has a minimal element in $\in_{\text{ORD}}$.

Proof of (1). First recall that for all sets x, y the following holds:

$$x \subseteq y \leftrightarrow x \cap y = x \tag{4.93}$$

Now, let x, y be two ordinals; in order to verify linearity of $\in_{\text{ORD}}$ we need to show that either $x \in y$, or $y \in x$ or $x = y$. By Lemma 4.19(1), $x \cap y$ is itself an ordinal. Consider the following inclusions:

$$x \cap y \subseteq x, x \cap y \subseteq y \tag{4.94}$$

If both inclusions are strict, i.e. $x \cap y \subsetneq x$ and $x \cap y \subsetneq y$, then by Lemma 4.19(2) this means $x \cap y \in x$ and $x \cap y \in y$, which implies $x \cap y \in x \cap y$ which contradicts Lemma 4.8(1).

Hence at least one of the inclusions in (4.94) is the identity. If both inclusions are in fact identities, i.e. $x \cap y = x$ and $x \cap y = y$, then by (4.93) $x \subseteq y$ and $y \subseteq x$, and hence $x = y$. If $x \cap y = x$ is true but $x \neq y$, we obtain $x \subsetneq y$, which by Lemma 4.19(2) implies $x \in y$. Similarly, if $x \cap y = y$ and $x \neq y$, we obtain $y \in x$.

Proof of (2). Let $a \subseteq \text{ORD}$ be a non-empty subset of ordinal numbers. Let $\alpha \in a$ be an ordinal. Consider the set $\alpha \cap a$ which is a subset of α. If $\alpha \cap a$ is empty, then α is a minimal element of a in $\in$ (if there were $c \in a$ such that $c \in \alpha$, then $c \in a \cap \alpha$, but $a \cap \alpha$ should is empty). If $a \cap \alpha$ is non-empty, then because α is welloredered by $\in_\alpha$, $a \cap \alpha$ must have the least element in the ordering $\in_\alpha$. Let us denote this element as β. We will argue that β is a minimal element of a. Assume for contradiction that there is some $c \in a$ such that $c \in \beta$. By transitivity of α and the fact that $c \in \beta \in \alpha$, this would mean that $c \in \alpha \cap a$. But since β is the least element in $\in_\alpha$ in $\alpha \cap a$ and $c \in \beta \leftrightarrow c \in_\alpha \beta$, this is a contradiction. □

4.5.2 Basic properties of natural numbers

We will now apply the results we have shown for ORD to natural numbers. Recall that $\mathbb{N}$ was define as $\mathbb{N} = \bigcap\{a \mid a \text{ is an inductive set}\}$.

Lemma 4.23 *The following properties hold:*

(1) Every limit ordinal is an inductive set.
(2) If α is a limit ordinal, then $\mathbb{N} \subseteq \alpha$.

Proof. Ad (1). Clear by definition.

Ad (2). Since α is an inductive set, $\mathbb{N} = \bigcap\{a \,|\, a \text{ is an inductive set}\} \subseteq \alpha$. □

The following induction theorem is important in its own right because it captures our intuition behind the natural numbers.

Theorem 4.24 (Induction for natural numbers) *Let A be a subset of $\mathbb{N}$ such that*

(1) $0 \in A$, *and*
(2) $(\forall n \in \mathbb{N})\, n \in A \rightarrow n+1 \in A$,

then $A = \mathbb{N}$.

Proof. We will argue by contradiction: so assume it is not the case that $A = \mathbb{N}$. Let n_0 be the least element of $\mathbb{N}$ which is not in A (recall that $\mathbb{N} \subseteq \mathrm{ORD}$, and so it is wellordered by $<$). By Lemma 4.23(2), n_0 is a successor ordinal: if n_0 were a limit ordinal, then $\mathbb{N} \subseteq n_0$ leads to $n_0 \in n_0$ which is a contradiction. It follows that $n_0 = m_0 + 1$ for some m_0. Because n_0 was the least such not in A, m_0 is in A. However, this means by condition (2) above that $m_0 + 1 \in A$, contradiction. We can conclude that $A = \mathbb{N}$. □

We will apply Theorem 4.24 to show that $\mathbb{N}$ is the least limit ordinal.

Lemma 4.25 *For every $n \in \mathbb{N}$ it holds that $n \subseteq \mathbb{N}$. This means that $\mathbb{N} = \omega$ is a transitive set and hence an ordinal.*

Proof. Let us define A as the set of all $n \in \mathbb{N}$ such that $n \subseteq \mathbb{N}$:

$$A = \{n \in \mathbb{N} \,|\, n \subseteq \mathbb{N}\}.$$

We argue that $0 \in A$ and if $n \in A$, then $n+1 \in A$. $0 \in A$ is obvious. Assume $n \in A$, that is $n \in \mathbb{N}$ and $n \subseteq \mathbb{N}$. We want to show that $n + 1 = n \cup \{n\}$ is included in $\mathbb{N}$ as a subset: if q is in n, then by the induction assumption $n \subseteq \mathbb{N}$, $q \in \mathbb{N}$; if $q = n$, then $q \in \mathbb{N}$ because n is a natural number. This means that $n + 1 \in A$. By Theorem 4.24, we know that $A = \mathbb{N}$ as desired.

By Lemma 4.14, $\mathbb{N}$ is a transitive set of ordinals, and is therefore an ordinal. Now we see the motivation to denote $\mathbb{N}$ by the Greek letter ω which we use for ordinal numbers. □

Corollary 4.26 $\mathbb{N} = \omega$ *is the least limit ordinal.*

Proof. $\mathbb{N}$ is clearly a limit ordinal because it is an inductive set: $n \in \mathbb{N}$ implies $n + 1 \in \mathbb{N}$. It is the least limit ordinal because $\mathbb{N} \subseteq \alpha$ for every limit ordinal α. □

Exercise.

1. Show that $\bigcup \omega = \omega$, and hence ω is the supremum of all natural numbers. [Hint. $\bigcup \omega \subseteq \omega$ follows by transitivity of ω. The other direction follows from the fact that if $n \in \omega$, then $n + 1 \in \mathbb{N}$ and $n \in n + 1$.]

Remark 4.27 We have deferred the discussion of natural numbers to this point when we can apply the results which we have developed for ordinals numbers. All these results can be developed just for natural numbers based on the definition $\mathbb{N} = \bigcap\{a \,|\, a \text{ is an inductive set}\}$, but the proofs are basically the same as the ones for all ordinal numbers. So we have decided to do these proofs just once, and apply them retrospectively to natural numbers.

Remark 4.28 We might alternatively *define* a natural number as an ordinal smaller than the first limit ordinal. It is implicit in what we have shown that these two definitions are equivalent.

Definition 4.29 *We say that a set x is* finite *if there is a bijection between x and a natural number.*

4.5.3 ω is a model of Peano Arithmetics, PA

The following Theorem (whose proof we will postpone till Section 5.2) shows that we can uniquely defined a function on ω by postulating what happens at the successor step:

Theorem 4.30 (Construction by induction on ω) *Let g be a function from ω to ω and n_0 a natural number. Then there exists a unique function f from ω to ω such that*

$$\textit{for all } n \in \omega, f(0) = n_0, f(n+1) = g(f(n)).$$

The set of natural numbers – as represented in ZFC – shall be from now on denoted by ω. We will show that we can define in ZFC basic arithmetical operations such as $+, \cdot$ and relations such as $\leq$ in such a way that the resulting structure satisfies all axioms of arithmetics.

The *Peano Arithmetics, PA* is a theory in the language $\{+, \cdot, 0, S, \leq, <\}$ with the following axioms:

1. $(\forall x, y)(S(x) = S(y) \to x = y)$,
2. $(\forall x)(S(x) \neq 0)$,
3. $(\forall x)(x \neq 0 \to (\exists y) x = S(y))$,
4. $(\forall x)(x + 0 = x)$,
5. $(\forall x, y)(x + S(y) = S(x + y))$,
6. $(\forall x)(x \cdot 0 = 0)$,
7. $(\forall x, y)(x \cdot S(y) = x \cdot y + x)$,
8. $(\forall x, y)(x \leq y \leftrightarrow (\exists v) v + x = y)$,
9. $(\forall x, y)(x < y \leftrightarrow (\exists v)(S(v) + x = y)$,
10. (Schema of Induction) For every formula $\varphi(x, \bar{x})$ with free variables x and $\bar{x} = x_0, \dots, x_{n-1}$, we add the following axiom:

$$(\forall \bar{x})(\varphi(0, \bar{x}) \wedge [(\forall x)(\varphi(x, \bar{x}) \to \varphi(S(x), \bar{x}))] \to (\forall x)\varphi(x, \bar{x})).$$

From these axioms one can show all the usual properties of the operations, for instance commutativity of $+$ and $\cdot$.

Notice the way the operation $+$ is defined by induction from the function S: for every n, there is a unique function $f_n(x) = n + x$ which is defined by induction $f_n(0) = n + 0 = n$, and $f_n(x+1) = S(f_n(x))$. Hence g from Theorem 4.30 is in this case the function S for arguments > 0 and $n_0 = n$. Multiplication is similarly defined from the $+$ function: $F_n(x) = n \cdot x$, defined by $F_n(0) = n \cdot 0 = 0$, and $F_n(x+1) = F_n(x) + n = f_n(F_n(x))$ (providing that we already know that the operation $+$ is commutative). Hence g from Theorem 4.30 is the function f_n for arguments > 0 and $n_0 = 0$.[9]

We will show that we can define operations $S, +, \cdot$ and relation $\leq, <$ in ZFC, so that ω together with these operations satisfies all axioms of PA. We call this structure the *arithmetics as built in ZFC*.

We denote as $<_l$ the following *lexicographical ordering* defined on ω^2:

$$(n_0, m_0) <_l (n_1, m_1) \leftrightarrow n_0 < n_1 \vee (n_0 = n_1 \wedge m_0 < m_1). \tag{4.95}$$

Lemma 4.31 *The partial ordering $<_l$ is a strict wellordering on ω^2.*

Proof. It is clearly antireflexive and transitive (exercise). We will show that it is a wellordering. So let $A \subseteq \omega^2$ be a non-empty set, we wish to show that there is some pair (n_0, m_0) such that (n_0, m_0) is the $<_l$-least element in A. Let us define:

$$n_0 = \text{the } < \text{-least element of } \{n \mid (\exists m)\, (n, m) \in A\} = \mathrm{dom}(A),$$

[9] In set theory, Theorem 4.30 is a theorem in ZF, if we work in PA, this theorem is a consequence of the Schema of Induction (where g is given by a formula).

and

$$m_0 = \text{the } < \text{-least element of } \{m \mid (n_0, m) \in A\}.$$

This definition is correct because both sets are non-empty if A was non-empty; it is also obvious that $(n_0, m_0) \in A$. If $(n, m) \neq (n_0, m_0)$ and $(n, m) \in A$, then by the construction of n_0, m_0, we have that either $n_0 < n$, in which case $(n_0, m_0) <_l (n, m)$, or $n_0 = n$ but $m_0 < m$, and hence again $(n_0, m_0) <_l (n, m)$. It follows that (n_0, m_0) is the $<_l$-least element of A. □

Definition of zero, 0. We set $0 = \emptyset$.

Definition of the successor, S. For $n \in \omega$ we define $S(n) = n \cup \{n\}$.

Definition of addition, $+$. For $n, m \in \omega$ define $n+m$ as the natural numbers which is isomorphic with the set

$$(\{0\} \times n) \cup (\{1\} \times m)$$

ordered by $<_l$.

Definition of multiplication, $\cdot$. For $n, m \in \omega$ define $n \cdot m$ as the natural numbers which is isomorphic with the set

$$n \times m, \text{ or equivalently } m \times n$$

ordered by $<_l$.

Remark 4.32 One might wonder how we know that the ordinal isomorphic with $(\{0\} \times n) \cup (\{1\} \times m)$ is a finite ordinal. This is shown by induction on m in $n + m$, using the fact that ω is an inductive set. Similarly for $n \cdot m$.

Definition of ordering $\leq$ and strict ordering $<$. We define for $n, m \in \omega$,

$$n < m \leftrightarrow n \in m, \text{ and } n \leq m \leftrightarrow n < m \vee n = m.$$

Theorem 4.33 *ω with the operations above satisfies all the axioms of PA.*

Proof. *Exercise. □

Corollary 4.34 *ZF proves the consistency of PA (where PA is formulated within ZF). This is denote as:*

$$\mathrm{ZF} \vdash \mathrm{Con}(\ulcorner \mathrm{PA} \urcorner),$$

where $\ulcorner$PA$\urcorner$ denotes the formalization of the usual axioms of PA within ZF.

Remark 4.35 You will learn more about these concepts in Logika II.

5 Transfinite induction

5.1 Proof by transfinite induction

Since ordinal numbers contain some limit ordinals (we have shown that ω is a limit ordinal, but there are many others), we will generalize the Induction theorem 4.24 to all ordinal numbers as follows:

Theorem 5.1 *Let A be a subclass of* ORD *such that for every ordinal α holds:*

$$\alpha \subseteq A \to a \in A,$$

then $A = \mathrm{ORD}$.

Proof. For contradiction assume that α is the least ordinal which is not in A. Then $\alpha \subseteq A$, and by assumption $\alpha \in A$. Contradiction. □

This can be reformulated for successor and limit ordinals:

Theorem 5.2 *Let A be a subclass of* ORD *such that for every ordinal α holds:*

(i) $0 \in A$,
(ii) $\alpha \in A \to \alpha + 1 \in A$,
(iii) $[(\forall \beta < \alpha)\beta \in A] \to \alpha \in A$,

then $A = \mathrm{ORD}$.

It follows that if we want to show that some property A holds for all ordinal numbers, it is enough to argue by Theorem 5.1 or 5.2. For examples of use see the following sections.

5.2 Construction of a function by transfinite induction

Theorem 5.3 *Let G be a class function from V to V. Then there is a unique function F from* ORD *to V such that for every $\alpha \in \mathrm{ORD}$:*

$$F(\alpha) = G(F \restriction \alpha).$$

Proof.

We will construct the required F as a union of partial approximation. Set

$$X = \{f \mid (\exists \alpha)\mathrm{dom}(f) = \alpha \wedge (\forall \beta < \alpha) f(\beta) = G(f \restriction \beta)\}.$$

The system X has the following properties:

(i) For every $f \in X$ and $\beta \in \mathrm{dom}(f)$, the restriction $f \restriction \beta$ is also in X,
(ii) If f, f' are in X and $\alpha \in \mathrm{dom}(f) \cap \mathrm{dom}(f')$ then $f(\alpha) = f'(\alpha)$,
(iii) Every $\alpha \in \mathrm{ORD}$ is the domain of some $f \in X$.

We argue that these conditions are true.

(i) is obvious (if $\gamma < \beta$ then $f(\gamma) = f \restriction \beta(\gamma) = G(f \restriction \gamma)$).

(ii) Let functions f, f' in X be given. Clearly, the intersection $\mathrm{dom}(f) \cap \mathrm{dom}(f')$ is some ordinal, let us denote it as γ. We want to show that for every $\alpha < \gamma$, $f(\alpha) = f'(\alpha)$. That is we want to show that

$$A = \{\alpha < \gamma \mid f(\alpha) = f'(\alpha)\}$$

is equal to γ. Using Induction theorem 5.1 (localized to γ), it suffices to show that if $\alpha < \gamma$ is such that $\alpha \subseteq A$, then $\alpha \in A$. The fact $\alpha \subseteq A$ means that $f \restriction \alpha = f' \restriction \alpha$; by definition of X, $f(\alpha) = G(f \restriction \alpha) = G(f' \restriction \alpha) = f'(\alpha)$. It follows that $A = \gamma$.

(iii) Again we use the Induction theorem, this time Theorem 5.2 to make things more clear. Let

$$A = \{\alpha \mid (\exists f \in X)\alpha = \mathrm{dom}(f)\}.$$

(Successor step) If $\alpha \in A$ given and f is such that $\mathrm{dom}(f) = \alpha$, then the function f' defined by

$$f' = f \cup \{(\alpha, G(f))\}$$

is a function in X with domain $\alpha + 1 \in A$.

(Limit step) If α is a limit ordinal such that $\alpha \subseteq A$, then for every $\beta < \alpha$ there is some $f_\beta \in X$ such that $\beta = \mathrm{dom}(f_\beta)$. By (ii), the union $\bigcup_{\beta<\alpha} f_\beta$ is a function with domain α, which we will denote as g. If $\gamma \in \mathrm{dom}(g)$, then by definition of g there is some f_β such that $\gamma \in \mathrm{dom}(f_\beta)$; it follows that $g \in X$ because

$$g(\gamma) = f_\beta(\gamma) = G(f_\beta \restriction \gamma) = G(g \restriction \gamma).$$

It follows that $\alpha \in A$, and by Theorem 5.2, $A = \mathrm{ORD}$ as desired.

Properties (i)–(iii) suffice the prove the theorem. By (ii), $F = \bigcup X$ is a function, and by (iii) the domain of F is ORD. We show that for every α, $F(\alpha) = G(F \restriction \alpha)$. If α is an ordinal, then by (ii) there is some $f \in X$ such that $\alpha \in \mathrm{dom}(f)$ and

$$f(\alpha) = F(\alpha), f \restriction \alpha = F \restriction \alpha.$$

It follows that

$$F(\alpha) = G(f \restriction \alpha) = G(F \restriction \alpha),$$

as desired.

By transfinite induction 5.1 one can also easily show that if F' is a function satisfying the definition of F, then $F = F'$. It follows that F is unique. *Exercise.* □

Recall that if f is a function and $x \subseteq \mathrm{dom}(f)$, then $f[x] = f''x$ denotes the set $\{b \mid (\exists a \in x) f(a) = b\}$.

The above theorem has several variants.

Theorem 5.4 *Let G be a function from V to V. Then there is a unique $F : \mathrm{ORD} \to V$ such that for every α:*

$$F(\alpha) = G(F[\alpha]).$$

Proof. Define a function G' by setting: $G'(x) = G(\mathrm{rng}(x))$ if x is a binary relation, or $\emptyset$ if x is not a binary relation. Then apply Theorem 5.3 to G': there is a unique F such that

$$F(\alpha) = G'(F \restriction \alpha) = G(F[\alpha])$$

□

Another variant (compare with Theorem 5.1 and 5.2):

Theorem 5.5 *Let G_1 and G_2 be two function from V to V and a a set. Then there is a unique $F : \mathrm{ORD} \to V$ such that:*

(i) $F(0) = a$,
(ii) $F(\alpha + 1) = G_1(F(\alpha))$,

(iii) $F(\lambda) = G_2(F[\lambda])$*, where* λ *is a limit ordinal.*

Proof. Exercise.* Hint: define G by:

$$\begin{aligned} G(x) &= G_1(x(\alpha)), \text{ if } x \text{ is a function and } \mathrm{dom}(x) = \alpha + 1, \\ &= G_2(\mathrm{rng}(x)), \text{ if } x \text{ is a function and } \mathrm{dom}(x) = \lambda, \text{for } \lambda \text{ a limit ordinal} \\ &= a, \text{ otherwise} \end{aligned}$$

Apply Theorem 5.3 to this G. □

5.3 Aplications

5.3.1 AC, WO, and PM are all equivalent

Recall that Theorem 3.12 shows that WO (Well-ordering principle) implies AC (Axiom of Choice), and Theorem 3.15 shows that PM (Principle of Maximality) implies WO (Well-ordering Principle). To complete the equivalence between these principles, we will show now that AC implies PM.

Theorem 5.6 *Axiom of Choice (AC) implies Principle of Maximality (PM):*

(5.96)
$$ZF \vdash AC \to PM.$$

Corollary 5.7 *AC, PM, WO are all equivalent, i.e.*

(5.97)
$$ZF \vdash AC \leftrightarrow WO \leftrightarrow PM$$

Proof. (of Theorem 5.6) Let $(P, \leq)$ be a partially ordered set which satisfies the condition that every chain in P has an upper bound (see Definition 3.14 for the formulation of PM and for the meaning of *chain*). Using AC, we want to show that above every element $p \in P$ there is a maximal element a in the ordering $\leq$, i.e. $p \leq a$ and there is no $b \in P$ such that $a < b$.

By AC, we can fix a choice function C on $\mathscr{P}(P)$. Let $p \in P$ be given. We will find a maximal element above p using the transfinite induction.

Define by induction the following function $F : \mathrm{ORD} \to P$:

(5.98)
$$\begin{aligned} F(0) &= p \\ F(\alpha+1) &= C(\{q \in P \,|\, F(\alpha) < q\}), \text{ if } \{q \in P \,|\, F(\alpha) < q\} \text{ is non-empty} \\ &= p, \text{ otherwise} \\ F(\lambda) &= C(\{q \in P \,|\, (\forall \beta < \lambda) F(\beta) < q\}), \text{ for } \lambda \text{ limit ordinal} \end{aligned}$$

Note that for every limit λ, the set $\{F(\beta) \,|\, \beta < \lambda\}$ is a chain and by our assumption it has an upper bound, and so the limit stage of F is correctly defined.

The following claims hold, proving the theorem:

(i) There is some $\alpha + 1 \in \mathrm{ORD}$ such that for all $\beta < \beta' < \alpha + 1$ it holds that $F(\beta) < F(\beta')$, but for all $\gamma \geq \alpha + 1$ it holds that $F(\gamma) = p$. Let as denote this ordinal $\alpha_0 + 1$. (In words, the values of F are all distinct elements of P up to α_0, but from $\alpha_0 + 1$ the values will equal to p).

(ii) The maximal element of P above p is equal to $F(\alpha_0)$ (note that $\alpha_0 = 0$ is possible, in which case p is maximal in P).

Ad (i). This follows from the Schema of Replacement: if there were no such α_0, then F would be a 1-1 function from ORD into P. The inverse function F^{-1} would then be a 1-1 function from a subset of a set P onto ORD. This contradicts Schema of Replacement which says that an image of a set by a function is also a set (and ORD is not a set).

Ad (ii). This is obvious from the definition of F. □

Remark 5.8 A simple modification of the proof shows directly that AC implies WO. (Exercise.)

5.3.2 Ordinal arithmetics – definition of operations

We can define the usual operations on ordinal numbers, such as $\alpha+\beta, \alpha\cdot\beta$, and α^β. The definition of $+$ and $\cdot$ has a "geometric" motivation as follows:

Recall the lexicographical ordering (4.95) defined on natural numbers. We extend this ordering straightforwardly to pairs of ordinal numbers (α_0, β_0) and (α_1, β_1) as follows:

(5.99) $$(\alpha_0, \beta_0) <_l (\alpha_1, \beta_1) \leftrightarrow \alpha_0 < \alpha_1 \vee (\alpha_0 = \alpha_1 \wedge \beta_0 < \beta_1).$$

As in Lemma 4.31, one can show:

Lemma 5.9 *The partial ordering $<_l$ is a strict wellordering on* ORD^2.

Proof. Exercise. [Hint. Generalize proof of Lemma 4.31.] □

We define:

Addition.

(5.100) $$\alpha + \beta = \text{ the unique ordinal isomorphic to } (\{0\} \times \alpha \cup \{1\} \times \beta, <_l).$$

Multiplication.

(5.101) $$\alpha \cdot \beta = \text{ the unique ordinal isomorphic to } (\beta \times \alpha, <_l).$$

Note that these operations are not commutative:

Exercise:

1. Verify the following:
   (i) $1 + \omega = \omega$,
   (ii) $1 + \omega < \omega + 1$,
   (iii) $2 \cdot \omega = \omega$,
   (iv) $\omega + \omega = \omega \cdot 2$
2. Show that in the definition (5.101), it *does* matter (unlike in the definition of these operations for natural numbers) whether we write $\beta \times \alpha$ or $\alpha \times \beta$.

Definition by transfinite induction

Given a function $f : \alpha \to X$, we can view f as a function indexing elements in X by ordinals in α. We can visualize this by using the following notation for f:

(5.102) $$f = \langle x_\xi \,|\, \xi < \alpha \rangle,$$

where it is understood that for each $\xi < \alpha$, $f(\xi) = x_\xi$. Note that $\mathrm{rng}(f) = \{x_\xi \,|\, \xi < \alpha\}$, so differentiate the notation $\langle\cdot\,|\,\cdot\rangle$ and $\{\cdot\,|\,\cdot\}$.

If the range of f is included in ORD, i.e. $f : \alpha \to \mathrm{ORD}$, we often write

$$f = \langle \gamma_\xi \,|\, \xi < \alpha \rangle. \tag{5.103}$$

Let $\langle \gamma_\xi \,|\, \xi < \alpha \rangle$ be a non-decreasing sequence of ordinals (i.e. $\xi < \zeta < \alpha$, then $\gamma_\xi \leq \gamma_\zeta$). We define the *limit* of the sequence $\langle \gamma_\xi \,|\, \xi < \alpha \rangle$ by

$$\lim_{\xi \to \alpha} \gamma_\xi = \sup(\{\gamma_\xi \,|\, \xi < \alpha\}). \tag{5.104}$$

We say that a sequence of ordinals $\langle \gamma_\alpha \,|\, \alpha \in \mathrm{ORD} \rangle$ is *normal* if it is increasing (we also call this *monotonous*) (i.e. $\xi < \zeta$ implies $\gamma_\xi < \gamma_\zeta$) and for every limit α, $\gamma_\alpha = \lim_{\xi \to \alpha} \gamma_\xi$. We say that an ordinal α is a *fixed point* of $\langle \gamma_\alpha \,|\, \alpha \in \mathrm{ORD} \rangle$ if $\gamma_\alpha = \alpha$.

The following is a simple lemma concerning normal functions:

Lemma 5.10 *If $F = \langle \gamma_\alpha \,|\, \alpha \in \mathrm{ORD} \rangle$ is a normal function from* ORD *to* ORD*, then*

$$\textit{for every } \alpha, \alpha \leq F(\alpha). \tag{5.105}$$

Proof. We proceed by induction: assume for contradiction that α_0 is the least ordinal such that

$$F(\alpha_0) < \alpha_0 \tag{5.106}$$

Using the fact that F is monotonous, we obtain from (5.106):

$$F(F(\alpha_0)) < F(\alpha_0) < \alpha_0, \tag{5.107}$$

which contradicts that α_0 was the least such that (5.106) holds. □

We now show that the notion of normality ensures that normal function have unbounded fixed points:

Lemma 5.11 (Fixed-point lemma) *If $F = \langle \gamma_\alpha \,|\, \alpha \in \mathrm{ORD} \rangle$ is a normal function from* ORD *to* ORD*, then it has arbitrarily large fixed points (in other words, there is a proper class of fixed points of $\langle \gamma_\alpha \,|\, \alpha \in \mathrm{ORD} \rangle$).*

Proof. We will show that for every ordinal α there is some $\beta \geq \alpha$ such that $\gamma_\beta = \beta$. Define the following function f by induction on ω:

$$\begin{array}{rcl} f(0) & = & \alpha, \\ f(n+1) & = & F(f(n)), \text{ for every } n \in \omega. \end{array}$$

Let

$$\beta = \lim_{n \to \omega} f(n). \tag{5.108}$$

We claim that β is a fixed point, i.e.

$$\beta = \gamma_\beta. \tag{5.109}$$

By (5.105), we have that $\alpha = f(0) \leq f(1) \leq \ldots$ and so $\beta \geq \alpha$. Realize that exactly one of the following is true:

(i) α is a fixed point and $\alpha = \beta$.
(ii) α is not a fixed point, and $\alpha < \beta$ and β is a limit ordinal.

To see (i), note that if $\beta = \alpha$, then in particular $\alpha = f(1) = F(\alpha)$, which is equivalently written as $\alpha = \gamma_\alpha$.

(ii) If α is not a fixed point, then $F(\alpha) > \alpha$ by (5.105), and by induction (using monotonicity) $\alpha < F(\alpha) < F(F(\alpha)) < \dots$. This implies that $\alpha < \beta$ and β is a limit ordinal.

We now show that in case (ii), β is a fixed point. By the definition of a normal function at a limit ordinal, we have:

$$\gamma_\beta = \lim_{\xi \to \beta} \gamma_\xi = \lim_{n \to \omega} f(n) = \beta, \tag{5.110}$$

where the middle identity $\lim_{\xi \to \beta} \gamma_\xi = \lim_{n \to \omega} f(n)$ follows from the fact that for every $\xi < \beta$, there is some n such that $\xi < f(n)$ (because β is the limit of $f(n)$'s), and consequently $\gamma_\xi = F(\xi) < F(f(n)) = f(n+1) < \beta$. The last identity follows from (5.108). □

Definition 5.12 (Addition.) *For all ordinal numbers α we define by induction on β in $\alpha + \beta$ the addition as follows:*

(i) $\alpha + 0 = \alpha$,
(ii) $\alpha + (\beta + 1) = (\alpha + \beta) + 1$,
(iii) $\alpha + \beta = \lim_{\xi \to \beta}(\alpha + \xi)$, for limit β.

Definition 5.13 (Multiplication.) *For all ordinal numbers α we define by induction on β in $\alpha \cdot \beta$ the multiplication as follows:*

(i) $\alpha \cdot 0 = 0$,
(ii) $\alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha$,
(iii) $\alpha \cdot \beta = \lim_{\xi \to \beta}(\alpha \cdot \xi)$, for limit β.

Exercises.

1. * Verify by transfinite induction on β that the geometric definitions of addition and multiplication are equivalent to the definition by transfinite induction.
2. Verify that for a fixed α, the functions f_α and g_α defined on ORD are normal, where: $f_\alpha(\beta) = \alpha + \beta$ and $g_\alpha(\beta) = \alpha \cdot \beta$. We say that addition and multiplication is a normal function in the second variable β.

Transfinite induction allows us to define also the exponentiation of ordinal numbers:

Definition 5.14 (Exponentiation.) *For all ordinal numbers α we define by induction on β in α^β the exponentiation as follows:*

(i) $\alpha^0 = 1$,
(ii) $\alpha^{(\beta+1)} = (\alpha^\beta) \cdot \alpha$,
(iii) $\alpha^\beta = \lim_{\xi \to \beta}(\alpha^\xi)$, for limit β.

Exercises.

1. Verify that for a fixed α, the function r_α defined on ORD is normal, where: $r_\alpha(\beta) = \alpha^\beta$. We say that exponentiation is a normal function in the second variable β.
2. * Try to visualize the following ordinal numbers: $\omega < \omega \cdot 2 < \omega \cdot 3 < \omega \cdot \omega = \omega^2 < (\omega^2) + \omega < \omega^3 < \omega^\omega < \omega^{\omega^\omega} = \omega^{(\omega^\omega)} < \epsilon_0 = \lim_{n \to \omega} n(\omega)$, where $n(\omega)$ is the iteration $\omega^{\omega^{\omega^{\cdots}}}$ of height n. [All these numbers still have the size ω.]
3. Using Lemma 5.11 argue that there are β, γ, δ such that:
   (i) $\omega + \beta = \beta$; argue using the construction in Lemma 5.11, that the least such $\beta \geq \omega$ is the ordinal $\omega \cdot \omega$.
   (ii) $\omega \cdot \gamma = \gamma$; argue using the construction in Lemma 5.11, that the least such $\beta \geq \omega$ is the ordinal ω^ω.

(iii) $\omega^\delta = \delta$. The least δ above ω such that $\omega^\delta = \delta$ (constructed using Lemma 5.11) is denoted as ϵ_0 (see above).

4. * Argue that if $\delta = \omega^\delta$, then:
   (i) $\omega + \delta = \delta$,
   (ii) $\omega \cdot \delta = \delta$,
   (iii) $\omega^\delta = \delta$.
   [Hint: $\delta \leq \omega + \delta \leq \omega \cdot \delta \leq \omega^\delta = \delta$.]

We will state the following important normal form theorem without proof:

Theorem 5.15 (Cantor's Normal Form Theorem.) *Every ordinal $\alpha > 0$ can be represented uniquely in the form:*

$$\alpha = \omega^{\beta_1} \cdot k_1 + \ldots + \omega^{\beta_n} \cdot k_n, \tag{5.111}$$

where n is a natural number $\geq 1, \alpha \geq \beta_1 > \ldots > \beta_n$, and $k_1, \ldots, k_n$ are non-zero natural numbers.

5.4 Optional topics

5.4.1 The well-founded universe

Using Theorem 5.5, let us define a class function V and a class WF, where WF is abbreviation for "well-founded" (fundovaný):

$$\begin{array}{rcl} V_0 & = & \emptyset \\ V_{\alpha+1} & = & \mathscr{P}(V_\alpha) \\ V_\lambda & = & \bigcup_{\alpha<\lambda} V_\alpha, \text{ if } \lambda \text{ is a limit ordinal} \\ WF & = & \bigcup_{\alpha\in\mathrm{ORD}} V_\alpha \end{array} \tag{5.112}$$

Note that we write (as is customary) V_α instead of $V(\alpha)$. We will show below, see Theorem 5.21, that the class WF contains all sets if we assume the Axiom of Foundation. It follows that under the Axiom of Foundation, the universe of all sets has a very simple and elegant description (5.112).

We first show some simple properties of WF:

Lemma 5.16 *(1) For each α, V_α is a transitive set.*
(2) For each α, and every $\beta < \alpha$, $V_\beta \subseteq V_\alpha$.
(3) For each α, $\alpha \subseteq V_\alpha$, and ORD $\subseteq WF$.

Proof. Claims (1) and (2) will by shown together by induction following Theorem 5.2. Let us denote

$$A = \{\alpha \in \mathrm{ORD} \,|\, V_\alpha \text{ is transitive and } (\forall \beta < \alpha) V_\beta \subseteq V_\alpha\}. \tag{5.113}$$

We show that $A = \mathrm{ORD}$. We need to show:

(i) $\emptyset \in A$,
(ii) If $\alpha \in A$, then $\alpha + 1 \in A$,
(iii) If all $\beta < \lambda$ are in A, then $\lambda \in A$ (for λ limit).

Ad (i). Clearly, $\emptyset \in A$.

Ad (iii). First note that if λ is a limit ordinal, then for every $\beta < \lambda$, $V_\beta \subseteq V_\lambda$ because $V_\lambda = \bigcup_{\beta<\lambda} V_\beta$.

If λ is a limit ordinal, and $x \in V_\lambda$, there is some $\beta < \lambda$ such that $x \in V_\beta$. Since V_β is by the induction assumption transitive, we obtain $x \subseteq V_\beta \subseteq V_\lambda$. It follows that V_λ is transitive. This shows that $\lambda \in A$.

Ad (ii). Let us assume that $\alpha \in A$, we will show that $\alpha + 1 \in A$. Since V_α is transitive, we obtain that $V_\alpha \subseteq V_{\alpha+1}$: if $x \in V_\alpha$, then $x \subseteq V_\alpha$, and hence $x \in V_{\alpha+1} = \mathscr{P}(V_\alpha)$. This suffices to show that $V_{\alpha+1}$ is transitive: if $x \in V_{\alpha+1}$, then $x \subseteq V_\alpha \subseteq V_{\alpha+1}$, and hence $x \subseteq V_{\alpha+1}$, which shows that $V_{\alpha+1}$ is transitive.

If $\beta < \alpha$, then $V_\beta \subseteq V_\alpha \subseteq V_{\alpha+1}$ by the induction assumption. If $\beta = \alpha < \alpha + 1$, then this means $V_\alpha \subseteq V_{\alpha+1}$, which we have already shown. This implies that $\alpha + 1 \in A$.

Combining (i)–(iii), we conclude that $A = \mathrm{ORD}$ as desired.

Ad (3). This is again shown by induction: It holds for $\emptyset$ and λ limit (to show that $\lambda \subseteq V_\lambda$ use the fact that $\alpha \in \lambda$ implies by induction assumption that $\alpha \subseteq V_\alpha$, and so $\alpha \in V_{\alpha+1} \subseteq V_\lambda$). To argue that $\alpha + 1 = \alpha \cup \{\alpha\} \subseteq V_{\alpha+1}$, note that $\alpha \subseteq V_\alpha$ implies $\alpha \in V_{\alpha+1}$, and so $\alpha + 1 \subseteq V_{\alpha+1}$. □

Corollary 5.17 *WF is a transitive class.*

Proof. If $x \in WF$, then there is some α such that $x \in V_\alpha$. By transitivity of V_α we obtain $x \subseteq V_\alpha$ and because $V_\alpha \subseteq WF$, we conclude $x \subseteq WF$. □

We make the following useful definition.

Definition 5.18 *The* rank *of a set $x \in \mathrm{WF}$, in symbols $\mathrm{rank}(x)$, is the least $\alpha \in \mathrm{ORD}$ such that $x \in V_{\alpha+1}$. Equivalently, $\mathrm{rank}(x)$ is the least $\alpha \in \mathrm{ORD}$ such that $x \subseteq V_\alpha$.*

It follows that if $\alpha = \mathrm{rank}(x)$, then $x \subseteq V_\alpha$, $x \notin V_\alpha$, and $x \in V_\beta$ for every $\beta > \alpha$. Note that $\mathrm{rank}(x)$ can be a limit ordinal.

We sum up the basic properties of the rank function:

Lemma 5.19 *Basic properties of the rank function:*

(i) *For any α, $V_\alpha = \{x \in \mathrm{WF} \,|\, \mathrm{rank}(x) < \alpha\}$.*

(ii) *If $y \in \mathrm{WF}$, then*

(a) $\forall x \in y(x \in \mathrm{WF} \wedge \mathrm{rank}(x) < \mathrm{rank}(y))$,

(b) $\mathrm{rank}(y) = \sup\{\mathrm{rank}(x) + 1 \,|\, x \in y\}$.

Proof. Ad (i). If $x \in V_\alpha$, then by definition of the rank, $\mathrm{rank}(x) < \alpha$. Conversely, if $\mathrm{rank}(x) < \alpha$, then $x \in V_\beta$ for some $\beta \leq \alpha$, and so $x \in V_\alpha$.

Ad (ii)(a). If $x \in y \in \mathrm{WF}$, then by transitivity of WF, $x \in \mathrm{WF}$. As $y \subseteq V_{\mathrm{rank}(y)}$ and $\mathrm{rank}(y)$ is the least such, $x \in y$ implies $x \in V_{\mathrm{rank}(y)}$, and so $\mathrm{rank}(x) < \mathrm{rank}(y)$.

Ad (ii)(b). Denote $\bar{\alpha} = \sup\{\mathrm{rank}(x) + 1 \,|\, x \in y\}$. First we show that $\mathrm{rank}(y) \leq \bar{\alpha}$. Clearly, $y \subseteq V_{\bar{\alpha}}$, because $\mathrm{rank}(x) < \bar{\alpha}$ for each $x \in y$ and so $x \in V_{\bar{\alpha}}$ by (i) of the present lemma. This implies that $\mathrm{rank}(y) \leq \bar{\alpha}$. Conversely we want to show that $\bar{\alpha} \leq \mathrm{rank}(y)$. For each $x \in y$, $\mathrm{rank}(x) < \mathrm{rank}(y)$ and so $\mathrm{rank}(x) + 1 \leq \mathrm{rank}(y)$. It follows that $\mathrm{rank}(y)$ is the upper bound of the set $\{\mathrm{rank}(x) + 1 \,|\, x \in y\}$, and so the supremum $\bar{\alpha}$ is less or equal to $\mathrm{rank}(y)$. □

We can thus view WF is as the universe constructed by recursion from simpler sets: for instance it cannot happen that there is set x in WF such that $x \in x$ because this would imply $\mathrm{rank}(x) < \mathrm{rank}(x)$.

We first state the following simple lemma:

Lemma 5.20 *Let x be a set and $x \subseteq WF$, then there is $\alpha \in \mathrm{ORD}$ such that $x \in V_\alpha$, and hence $x \in WF$.*

Proof. Consider the following class:

$$\bar{x} = \{\mathrm{rank}(y) \mid y \in x\} \subseteq \mathrm{ORD}. \tag{5.114}$$

By Schema of Replacement $\bar{x}$ must be a set (because it is a range of a function assigning ranks with domain restricted to x). Since the class ORD is a proper class, $\bar{x}$ (being a set) cannot be unbounded in ORD, so there must be some $\alpha \in \mathrm{ORD}$ such that $\bar{x} \subseteq \alpha$. It follows that $x \subseteq V_\alpha$ and consequently $x \in V_{\alpha+1}$. □

The following theorem claims that with the Axiom of Foundation, WF is the universe of all sets V. We denote this fact by the expression $WF = V$, which is a shorthand for the formula $(\forall x)(\exists \alpha \in \mathrm{ORD}) x \in V_\alpha$.

Theorem 5.21 *Let* F *denote the Axiom of Foundation and* ZF − F *the theory* ZF *without* F. *Then*

$$\mathrm{ZF} - \mathrm{F} \vdash \mathrm{F} \leftrightarrow V = \mathrm{WF}. \tag{5.115}$$

Proof. ($V = \mathrm{WF} \to \mathrm{F}$). We need to show that every x which is non-empty has a minimal element in the relation $\in$. Let x be a non-empty set. Consider the following set of ordinals

$$\bar{x} = \{\mathrm{rank}(y) \mid y \in x\}. \tag{5.116}$$

Let α be the least element of $\bar{x}$ and y some element of x such that $\mathrm{rank}(y) = \alpha$. We argue that y is a $\in$-minimal element of x: if $z \in y$, then $\mathrm{rank}(z) < \mathrm{rank}(y)$ by Lemma 5.19 (ii)(a). This contradicts the fact that α is the least element of $\bar{x}$.

($\mathrm{F} \to V = \mathrm{WF}$). Assume for contradiction that $X = V - \mathrm{WF} \neq \emptyset$. If X is a set, then we can argue straightforwardly: by F, there is some $y \in X$ which is $\in$-minimal, i.e. $y \subseteq \mathrm{WF}$ (no element $z \in y$ can be in X, which implies that z must be in WF). However, by Lemma 5.20, this means that $y \in \mathrm{WF}$, contradiction.

The general case (when X is a proper class) will follow from the following claim:

$$\text{F implies that every non-empty class has an } \in \text{-minimal element.} \tag{5.117}$$

We make the following false start: let X be a non-empty class. Pick any $x \in X$: if x is not minimal, consider the set $x \cap X$ (which is non-empty if x is not minimal). $x \cap X$ is a set and hence must have a minimal element, say z. Is z minimal in X? Well, it does not have to be: if $z' \in z \in x \cap X$, then $z' \notin x \cap X$, but $z' \in X$ is still possible. However if $z' \in x$, then it must hold $z' \notin X$ (otherwise $z' \in x \cap X$, which contradicts the minimality of z). This leads us to the idea to include x in a transitive set x^* to ensure that $z' \in z \in x^* \cap X$ implies $z' \in x^*$.

We define the *transitive closure* of a set.

Definition 5.22 *Let x be set, we define the* transitive closure *of x by recursion*

$$\begin{array}{rcl} \mathrm{trcl}_0(x) & = & x \\ \mathrm{trcl}_{n+1}(x) & = & \bigcup \mathrm{trcl}_n(x) \\ \mathrm{trcl}(\{x\}) & = & \bigcup_{n\in\omega} \mathrm{trcl}_n(x) \end{array} \tag{5.118}$$

Intuitively, $\mathrm{trcl}(x) = x \cup (\bigcup x) \cup (\bigcup\bigcup x) \cup \ldots$. In particular $x \subseteq \mathrm{trcl}(x)$.

Exercise. Show that for every x, the set $\mathrm{trcl}(x)$ is transitive. Also show that if x is transitive, then $\mathrm{trcl}(\{x\}) = x$, and that $x \subseteq y$ implies $\mathrm{trcl}(\{x\}) \subseteq \mathrm{trcl}(\{y\})$. Notice that this implies that $\mathrm{trcl}(\{x\})$ behaves as a closure operator: $x \subseteq \mathrm{trcl}(\{x\})$, $\mathrm{trcl}(\{\mathrm{trcl}(\{x\})\}) = \mathrm{trcl}(\{x\})$ and $x \subseteq y \to \mathrm{trcl}(\{x\}) \subseteq \mathrm{trcl}(\{y\})$ for every x, y.

We now finish the proof of (5.117). Let $x \in X$ be arbitrary. If x is not minimal, then $\operatorname{trcl}(\{x\}) \cap X$ is a non-empty set. By F, there is a minimal element $z \in \operatorname{trcl}(\{x\}) \cap X$. If z' is arbitrary and $z' \in z \in \operatorname{trcl}(\{x\}) \cap X$, then $z' \in \operatorname{trcl}(\{x\})$ by transitivity of $\operatorname{trcl}(\{x\})$, and so $z' \notin X$. It follows that z is minimal in X. □

Note that we can use the technique of Theorem 5.21 to argue that $\mathrm{CON}(\mathrm{ZF}-\mathrm{F}) \Rightarrow \mathrm{CON}(\mathrm{ZF})$, i.e. that by adding Axiom of Foundation to our system, we will not add contradiction (viz přednáška Modely teorie množin).

Remark 5.23 All mathematics can be defined in WF: $\omega = \mathbb{N} \subseteq V_\omega$, and so $\omega \in V_{\omega+1}$. $\omega \times \omega \subseteq V_\omega$, and because $\mathbb{Q}$ is a partition of $\omega \times \omega$, $\mathbb{Q} \subseteq \mathscr{P}(V_\omega) = V_{\omega+1}$, and so $\mathbb{Q} \in V_{\omega+2}$. Real numbers $\mathbb{R}$ are identified with certain subsets of $\mathbb{Q}$ (the so called *Dedekind cuts* (Dedekindovské řezy)), and so $\mathbb{R} \subseteq \mathscr{P}(\mathbb{Q}) \subseteq V_{\omega+2}$, which makes $\mathbb{R}$ an element of $V_{\omega+3}$, etc. In fact, it is safe to regard all "classical mathematics" to take place in $V_{\omega+\omega}$.

Remark 5.24 * *For interested students.* If $V = WF$ and A is a proper class, then we can uniformly choose a *non-empty subset* $a \subseteq A$ by considering the collection of elements in A of the least rank:

$$a = \{b \in A \mid (\forall c \in A)\operatorname{rank}(b) \leq \operatorname{rank}(c)\}. \tag{5.119}$$

This provides formal argument for comments concerning the definition of cardinals by means of equivalence classes.

6 Cardinal numbers

We assume AC in this section.

6.1 Basic operations

Cardinal numbers, denoted CARD, will be some special ordinal numbers, i.e. CARD $\subseteq$ ORD. Cardinal numbers will have the following basic properties:

(i) For each set x there exists a unique element of CARD denoted $|x|$,
(ii) $|x| \approx x$,
(iii) $x \approx y$, then $|x| = |y|$.

Definition 6.1 (Cardinals.) *We say that an ordinal α is a* cardinal *if there is no $\beta < \alpha$ such that $\beta \approx \alpha$. The class of cardinals is denoted* CARD*.*

If x is a set, then we define the size of $|x|$ as follows:

$$|x| = \text{ the least ordinal } \alpha \text{ such that } \alpha \approx x. \tag{6.120}$$

Lemma 6.2 *(1) By AC, $|x|$ is defined for every x and is unique.*
(2) For every set x, the size $|x|$ is an element of CARD*.*
(3) For every set x, $|x| \approx x$.
(4) For all sets x, y, $x \approx y \to |x| = |y|$.

Proof. Ad (1,2). Let $(x, <)$ be any wellordering of x (there is one by AC which implies WO). By Theorem 3.9 there is a unique ordinal α such that $(x, <) \cong (\alpha, \in)$, which implies $x \approx \alpha$. Since ORD is wellordered, there is the least ordinal $\beta \leq \alpha$ such that $\beta \approx \alpha$ (by transitivity of $\approx$, β is the least ordinal such that $\beta \approx x$). It follows that β is a cardinal number and $|\alpha| = |x| = \beta$.

Ad (3). By definition of $|x|$.

Ad (4). If $x \approx y$ and $\beta = |x|$, then by transitivity of $\approx$ we obtain $y \approx x \approx \beta \to y \approx \beta$, and so $|y| = \beta$. □

Natural numbers are all cardinals, and are called *finite cardinals.* We say that x is *finite* if there is $n \in \omega$ such that $|x| = n$; otherwise the cardinal is called *infinite.*

Cardinals (usually infinite) are denoted by Greek letters $\kappa, \lambda, \mu \dots$.

Lemma 6.3 *The following holds about cardinals:*

(1) ω is the least infinite cardinal.
(2) Every infinite cardinal is a limit ordinal.
(3) For every cardinal κ there is a cardinal λ such that $\lambda > \kappa$.
(4) If $\langle \kappa_\xi \mid \xi < \alpha \rangle$ is an increasing sequence of cardinals for α a limit ordinal, then the supremum $\bar{\kappa} = \sup(\{\kappa_\xi \mid \xi < \alpha\})$ is a cardinal.

Proof. Ad (1). Every natural numbers is a cardinal. ω is the supremum of natural numbers, and it is a cardinal because there can be no bijection between ω and a natural number (by induction on natural numbers show that no two natural numbers can have the same size, and consequently there can be no bijection between ω and a natural number).

Ad (2). For every infinite ordinal $\alpha \geq \omega$, one can easily construct a bijection between $\alpha + 1$ and α: for instance set $i(\alpha) = 0$, $i(n) = n + 1$ for $n \in \omega$, and $i(\beta) = \beta$ for $\omega \leq \beta < \alpha$. It is easy to check that $i : \alpha + 1 \to \alpha$ is a bijection. It follows that no cardinal greater than ω can be of the form $\alpha + 1$ for some ordinal α.

Ad (3). Use AC for simplicity.[10] If κ is a cardinal, then by Cantor's theorem $\kappa \prec \mathscr{P}(\kappa)$. By AC, $|\mathscr{P}(\kappa)|$ exists and must be bigger than κ.

Ad (4). By contradiction. Assume that there is a bijection $b : \bar{\kappa} \to \alpha$ for some $\alpha < \bar{\kappa}$. Since $\bar{\kappa}$ is the supremum of κ_ξ's, there is some κ_ξ such that $\alpha < \kappa_\xi$. We reach contradiction by arguing that there is a bijection between α and κ_ξ, contradicting the fact that κ_ξ is a cardinal. By Cantor-Bernstein's theorem, it suffices to show $\alpha \preceq \kappa_\xi$ and $\kappa_\xi \preceq \alpha$. The first inequality is obvious, because $\alpha < \kappa_\xi$. We will show the the second inequality. It suffices to find a 1-1 function from κ_ξ into α. However, this is easy: clearly b restricted to κ_ξ is such a function.

Note that (3) and (4) together imply that cardinal numbers are unbounded in the ordinal numbers, i.e. for every $\alpha \in$ ORD, there is a cardinal κ such that $\alpha \leq \kappa$. □

If κ is a cardinal, then the least cardinal above κ is denoted as κ^+.

We define the following basic operations for cardinal numbers:

Definition 6.4 (Addition.) *Let κ and λ be cardinals. We define*

$$\kappa + \lambda = |\kappa + \lambda|, \tag{6.121}$$

where the sign + on the righthand side denotes the addition on ordinal numbers.

Caution: This means that + for CARD and ORD is not the same operation. For instance $\omega + \omega = \omega \cdot 2 > \omega$ if we sum ordinal numbers, but $\omega + \omega = \omega$ if we sum cardinal numbers (we will show this later, see Corollary 6.23).

Definition 6.4 can be rephrased as follows. The sum $\kappa + \lambda$ is the size of disjoint union $X \cup Y$, where X and Y are disjoint and $|X| = \kappa$ and $|Y| = \lambda$ (if X, Y are disjoint, we call $X \cup Y$ the *disjoint union of X and Y*). Note that this definition does not depend on the particular sets X, Y which we choose: if $|X'| = |X|$ and $|Y'| = |Y|$, and X' and Y' are disjoint, then the size of the union of X, Y is the same as the size of the union of X', Y' (Exercise).

Definition 6.5 (Multiplication.) *Let κ and λ be cardinals. We define*

$$\kappa \cdot \lambda = |\kappa \cdot \lambda|, \tag{6.122}$$

where the sign $\cdot$ on the righthand side denotes the multiplication on ordinal numbers.

Caution: This means that $\cdot$ for CARD and ORD is not the same operation. For instance $\omega \cdot \omega = \omega^2 > \omega$ if we multiply ordinal numbers, but $\omega \cdot \omega = \omega$ if we multiply cardinal numbers (we will show this later, see Corollary 6.23).

Similarly as for the addition, we can view $\kappa \cdot \lambda$ is the size of a Cartesian product $X \times Y$, where $|X| = \kappa$ and $|Y| = \lambda$. Note that this time X and Y are not required to be disjoint.

Lemma 6.6 *For all cardinal numbers κ, λ such that $1 < \kappa, \lambda$:*

$$\kappa + \lambda \leq \kappa \cdot \lambda \tag{6.123}$$

Proof. Given X and Y such that $|X| > 1$ and $|Y| > 1$, we need contruct a 1-1 function from the disjoint union of X and Y to the product $X \times Y$. This is easy, Exercise. □

[10] One can show this without AC, see [BS].

Definition 6.7 (Exponentiation.) *Let κ and λ be cardinals. We define*

$$\kappa^{\lambda} = |\{f \mid f : \lambda \to \kappa\}|, \tag{6.124}$$

where $f : \lambda \to \kappa$ denotes a function with domain λ and range included in κ.

It is customary to write ${}^{\lambda}\kappa$ to denote the set $\{f \mid f : \lambda \to \kappa\}$, and so

$$\kappa^{\lambda} = |{}^{\lambda}\kappa|. \tag{6.125}$$

Caution. The ordinal exponentiation and the cardinal exponentiation are defined differently. Cardinal exponentiation is a complicated notion which is not completely determined by the axioms of ZFC. See Section 6.5.

Special case of cardinal exponentiation is 2^{κ} for a cardinal κ. This cardinal measures the size of the powerset of sets of size κ:

Lemma 6.8 *If A is a set and $|A| = \kappa$, then $|\mathscr{P}(A)| = 2^{\kappa}$.*

Proof. Let b be a bijection from A onto κ.

We define a bijection g from $\mathscr{P}(A)$ onto ${}^{\kappa}2$ as follows:

$$g(x) = \chi_x, \text{ where } \chi_x(\xi) = 1 \text{ if } b^{-1}(\xi) \in x, \text{ and } \chi_x(\xi) = 0 \text{ otherwise}, \tag{6.126}$$

for every $x \subseteq A$. It is easy to check that g is indeed a bijection and so $\mathscr{P}(A) \approx {}^{\kappa}2$ and so $|\mathscr{P}(A)| = 2^{\kappa}$. The function χ_x is called the *characteristic function* of x. □

6.2 Alephs

We state a simple Corollary of Theorem 3.9.

Corollary 6.9 *The following holds:*

(1) Assume that A is a proper class wellordered by a relation $<$ such that for every $a \in A$, the class of all predecessors of a in $<$ is a set:

$$\{b \in A \mid b < a\} \textit{ is a set for every } a \in A \tag{6.127}$$

Then $(A, <)$ is isomorphic to $(\mathrm{ORD}, \in)$.

(2) Assume that A is an infinite set wellordered by a relation $<$ such that for every $a \in A$, the set of all predecessors of a in $<$ is finite:

$$\{b \in A \mid b < a\} \textit{ is finite for every } a \in A \tag{6.128}$$

Then $(A, <)$ is isomorphic to $(\omega, \in)$.

Proof. Exercise. [Hint: Ad (1). First realize that for every $\alpha \in \mathrm{ORD}$, the class of all predecessors $\{\beta \in \mathrm{ORD} \mid \beta \in \alpha\}$ is a set because $\alpha = \{\beta \mid \beta \in \alpha\}$. (Recall that $\in$ is the wellordering of ORD.) Then generalize the proof of Theorem 3.9 and argue that it cannot happen that the isomorphism is from the one proper class onto some initial segment of the other class because this would contradict the Schema of Replacement (due to the fact that every initial segment is a set in both cases). It follows that the isomorphism must be from one class onto the other.

Ad (2). This is similar to (1), just replace the word "set" by 'finite" and "proper class" by "not finite". Instead of Schema of Replacement, use the fact that there can be no bijection between a finite and an infinite set.] □

Corollary 6.10 *The class* CARD *is a proper class wellordered by* $\in$ *such that for every* $\kappa \in$ CARD *the class* $\{\lambda \in \mathrm{CARD} \mid \lambda \in \kappa\}$ *is a set. It follows that*

(1) $(\mathrm{CARD}, \in)$ *and* $(\mathrm{ORD}, \in)$ *are isomorphic,*
(2) $(\mathrm{CARD} - \omega, \in)$ *and* $(\mathrm{ORD}, \in)$ *are isomorphic,*

where $\mathrm{CARD} - \omega$ *denotes the class of all infinite cardinals.*

Proof. CARD is a proper class because it is unbounded in ORD by Lemma 6.3(3). CARD is wellordered by $\in$ because it is a subclass of ORD which is wellordered by $\in$. The rest is a simple consequence of Corollary 6.9. □

By Corollary 6.10(1) there is a bijection i which is an isomorphism with respect to the ordering on ordinal numbers from ORD onto CARD, and so cardinal numbers can be enumerated by ordinal numbers: $i(0) = 0, i(1) = 1, \dots, i(\omega) = \omega, i(\omega + 1) =$ the least cardinal above ω etc. However, it is customary to start the enumeration at the least infinite cardinal ω, and use the enumeration ensured by 6.10(2).

Definition 6.11 *We denote by* $\aleph$ *("aleph") the order-preserving bijection from* ORD *onto* $(\mathrm{CARD} - \omega, \in)$.

For conventional reasons we write $\aleph_\alpha$ instead of $\aleph(\alpha)$. Examples: $\aleph_0 = \omega, \aleph_1 =$ the least cardinal greater than ω, etc.

Lemma 6.12 $\aleph$ *is a normal function. In particular* $\aleph$ *has fixed points unbounded in* CARD.

Proof. $\alpha < \beta$ clearly implies $\aleph_\alpha < \aleph_\beta$. By Lemma 6.3(4), if $\langle \aleph_\xi \mid \xi < \alpha \rangle$ is an increasing sequence of cardinals for α a limit ordinal, then the supremum $\sup(\{\aleph_\xi \mid \xi < \alpha\})$ is a cardinal and is equal to $\aleph_\alpha$. It follows that $\aleph$ is a normal function.

By Lemma 5.11, $\aleph$ has unbounded fixed points in CARD. It follows that there is a proper class of α's in ORD such that $\aleph_\alpha = \alpha$. □

We say that a cardinal $\aleph_\beta$ is a *successor cardinal* if $\beta = \alpha + 1$ for some α. If β is a limit ordinal, then we say that $\aleph_\beta$ is a *limit cardinal.*

Notational convention. We also write ω_α to denote $\aleph_\alpha$. It is the convention that if we write ω_α, we view the cardinal $\aleph_\alpha$ as an ordinal number together with the wellordering $\in$.

6.3 Regular and singular cadinals

Definition 6.13 *Let* $(X, \leq)$ *be a partially ordered set. Then* $Y \subseteq X$ *is called* cofinal *(kofinální) if*

$$\forall x \in X, \exists y \in Y \; x \leq y.$$

Examples. If $(X, \leq)$ has the greatest element x, then $\{x\}$ is the least cofinal subset of X. In general, if $A \subseteq X$ is the set of all maximal elements of $(X, \leq)$, then A is the least cofinal subset of X (Exercise: Show that any cofinal subset of X must contain A). If $(X, \leq)$ does not have the greatest element, or the maximal elements, the notion of cofinality of $(X, \leq)$ tends to be quite complicated.

We apply the notion of cofinal subset to ordinals $(\alpha, <)$. If α is a successor ordinal, i.e. $\alpha = \beta + 1$ for some β, $\{\beta\}$ is cofinal in α because β is the greates element in α. To avoid such trivial cases, we will focus on limit ordinals α.

Definition 6.14 *We say that an ordinal β is the* cofinality *(kofinalita) of α, and write this as* $\mathrm{cf}(\alpha) = \beta$*, if β is the least ordinal γ such that there is a cofinal subset of α of order type γ.*

Clearly, for each limit α: $\omega \leq \mathrm{cf}(\alpha) \leq \alpha$.

Also: for every $X \subseteq \alpha$:

$$X \text{ is cofinal in } \alpha \leftrightarrow \sup X = \bigcup X = \alpha.$$

Note: a cofinal subset X of order type $\mathrm{ot}(X)$ in α can be identified with the range of an increasing function $f : \mathrm{ot}(X) \to \alpha$.

Example. $\mathrm{cf}(\omega + \omega) = \mathrm{cf}(\omega^\omega) = \mathrm{cf}(\aleph_\omega) = \omega$.

Example. Let α be a countable ordinal, then $\mathrm{cf}(\alpha) = \omega$. Why? Let $f : \omega \to \alpha$ be a bijection. Define $g(0) = 0$, and $g(n + 1) = \max\{g(n), f(n)\} + 1$. Then by induction g is increasing, and $\sup\{f(n) \,|\, n < \omega\} = \sup\{g(n) \,|\, n < \omega\} = \alpha$, and so $\{g(n) \,|\, n < \omega\}$ is cofinal of order type ω.

Example. Does it hold that $\mathrm{cf}(\aleph_1) = \omega$? Under AC, this is false. By Lemma 6.15 (which holds under AC), $\aleph_1$ cannot have cofinality ω for the following reason: Assume for contradiction that $\{\xi_n \,|\, n < \omega\}$ is an increasing cofinal subset of $\aleph_1$. Then $\aleph_1 = \bigcup\{\xi_n \,|\, n < \omega\}$, and so $\aleph_1$ is a countable union of at most countable sets $\{\xi_n \,|\, n < \omega\}$. This contradicts Lemma 6.15.

Lemma 6.15 *(AC) Any countable union of at most countable sets is at most countable.*

Proof. We will assume that we know that $|\omega \times \omega| = \aleph_0$ (the argument is reviewed in the next Section 6.4.1).

Let $\{X_n \,|\, n < \omega\}$ be at most countable sets (all of them non-empty to avoid trivialities). We will argue that

$$|\bigcup\{X_n \,|\, n < \omega\}| \leq |\bigcup\{\{n\} \times X_n \,|\, n < \omega\}| \leq |\omega \times \omega| = \aleph_0.$$

Note that the sets $\{n\} \times X_n$ for $n < \omega$ are disjoint, even if X_n's are not. All inequalities above are obvious except perhaps $|\bigcup\{\{n\} \times X_n \,|\, n < \omega\}| \leq |\omega \times \omega|$. By AC, we can choose 1-1 functions $f_n : X_n \to \omega$ for each $n < \omega$ (these functions exists by the assumption that each X_n is at most countable). Define $g : \bigcup\{\{n\} \times X_n \,|\, n < \omega\} \to \omega \times \omega$ as follows

$$g(\langle n, x\rangle) = \langle n, f_n(x)\rangle.$$

It is easy to verify (Exercise) that g is 1-1, and so $|\bigcup\{\{n\} \times X_n \,|\, n < \omega\}| \leq |\omega \times \omega|$. □

Remark 6.16 Without AC, it is consistent that ω_1 has cofinality ω.

A general version of Lemma 6.15 formulated for all cardinals can be found below (see proof of Lemma 6.19).

Here are some basic properties of the notion of cofinality.

Lemma 6.17 *Let α be limit, then*

(i) $\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha)$,
(ii) $\mathrm{cf}(\alpha)$ *is an infinite cardinal.*

Thus, (i) and (ii) together imply that for every limit ordinal α, the cofinality $\mathrm{cf}(\alpha)$ *is a regular cardinal (see Definition 6.18 below).*

Proof. Ad (i). Denote $\beta = \mathrm{cf}(\alpha)$, $\gamma = \mathrm{cf}(\beta)$. So $\gamma \leq \beta$. Denote $f : \beta \to \alpha$ increasing cofinal, $g : \gamma \to \beta$ increasing cofinal. Then $g \circ f : \gamma \to \alpha$ is increasing cofinal, and so $\mathrm{cf}(\alpha) \leq \gamma$.

Details: cofinal let $\alpha' < \alpha$, then there is $\beta' < \beta$ such that $f(\beta') \geq \alpha'$, and $\gamma' < \gamma$ such that $g(\gamma') \geq \beta'$. So $f(g(\gamma')) \geq f(\beta') \geq \alpha$.

It follows that $\gamma < \beta$ is impossible (β is the least cofinal in α by assumption). So $\gamma = \beta$.

Ad (ii). Suppose $\mathrm{cf}(\alpha)$ can be mapped: $f : \kappa \to \mathrm{cf}(\alpha)$ bijection $\kappa < \mathrm{cf}(\alpha)$. Then consider $g : \kappa \to \mathrm{cf}(\alpha)$ defined $g(\xi) = \mathrm{sup} f[\xi]$. Then g is non-decreasing, and so $\mathrm{otrng}(g) \leq \kappa$, $\mathrm{rng}(g)$ cofinal. Let $h : \mathrm{cf}(\alpha) \to \alpha$ be increasing cofinal, then $h[\mathrm{rng}(g)]$ is a cofinal subset of α of order type $\leq \kappa$, contradiction. □

The notion of cofinality is used to divide all cardinals into two groups:

Definition 6.18 *We say that a cardinal κ is* regular *(regulární) if* $\mathrm{cf}(\kappa) = \kappa$. *If* $\mathrm{cf}(\kappa) < \kappa$, *we say that κ is* singular *(singulární).*

In an example above, we have shown using Lemma 6.15 that $\aleph_1$ is regular. In fact, under AC, this argument can be generalized to every successor cardinal:

Lemma 6.19 *(AC) $\aleph_{\alpha+1}$ is regular for every α.*

Proof. We will use the fact, proved below as Theorem 6.22, that for each $\beta \geq \omega$, $|\aleph_\beta \times \aleph_\beta| = \aleph_\beta$.

The first stage of the proof is analogous to Lemma 6.15: just as in Lemma 6.15, one shows that a union of at most $\aleph_\alpha$ sets of size at most $\aleph_\alpha$ is at most $\aleph_\alpha$.

This implies that $\aleph_{\alpha+1}$ is regular because no increasing sequence of ordinals in $\aleph_{\alpha+1}$ of size $< \aleph_{\alpha+1}$ can be cofinal: Let $\langle \xi_i \mid i < \eta \rangle$ be such a sequence, where $\eta < \aleph_{\alpha+1}$ is some ordinal. Then $\{\xi_i \mid i < \eta\}$ is a collection of at most $\aleph_\alpha$ sets each of size at most $\aleph_\alpha$, and so the union (supremum) $\bigcup\{\xi_i \mid i < \eta\}$ must have size at most $\aleph_\alpha$, and so is strictly smaller than $\aleph_{\alpha+1}$. □

Singular cardinals are unbouded because for every α, $\aleph_{\alpha+\omega}$ is a singular cardinal. In general note that if α is a limit ordinal, then

$$\mathrm{cf}(\aleph_\alpha) = \mathrm{cf}(\alpha).$$

If $\aleph_\alpha$ for some α limit ordinal, $\aleph_\alpha$ may or may not be regular. If $\aleph_\alpha$ is regular for some limit α, we call this cardinal weakly inaccessible. Existence of such cardinals cannot be shown inside of ZFC. They existence may be assumed however, as another higher "infinity axiom".

Note that ω is regular and limit, so such a generalization seems natural.

Definition 6.20 *κ is called weakly inaccessible (slabě nedosažitelný) if it is uncountable limit and regular.*

Note that if κ is weakly inaccessible and $\kappa = \aleph_\alpha$ for some α, then α must be limit ordinal and it must be a fixed point of the $\aleph$ function:

$$\aleph_\alpha = \mathrm{cf}(\aleph_\alpha) = \mathrm{cf}(\alpha),$$

and so $\alpha = \aleph_\alpha$. So the question of weak inaccessibility can be rehprased: are there any regular fixed points of the $\aleph$ function?

Note that by Fixed point lemma $\aleph$ function has unbounded fixed points because it is normal.

6.4 Basic cardinal arithmetics

6.4.1 Addition and multiplication

We define another ordering on ORD^2, which is called the *maximum-lexicographical ordering*, or the *canonical wellordering* of ORD^2, and denoted $<_{ml}$. Unlike $<_l$ defined before, it has the advantage that for every $(\alpha, \beta) \in \mathrm{ORD}^2$, the class of predecessors $\{(\gamma, \delta) \mid (\gamma, \delta) <_{ml} (\alpha, \beta)\}$ is a set, and by Corollary (6.9),

$$(\mathrm{ORD}, \in) \text{ is isomorphic with } (\mathrm{ORD}^2, <_{ml}) \tag{6.129}$$

We define $<_{ml}$ on ORD^2 as follows:

$$(\alpha_0, \beta_0) <_{ml} (\alpha_1, \beta_1) \leftrightarrow \max(\alpha_0, \beta_0) < \max(\alpha_1, \beta_1) \vee \\ (\max(\alpha_0, \beta_0) = \max(\alpha_1, \beta_1) \wedge (\alpha_0, \beta_0) <_l (\alpha_1, \beta_1)), \tag{6.130}$$

where $<_l$ is defined in (5.99).

Exercises.

1. Verify that $<_{ml}$ is a wellordering, and that the class of all $<_{ml}$-predecessors is a set for any $(\alpha, \beta) \in \mathrm{ORD}^2$.

Lemma 6.21 *The set $\omega \times \omega$ is countable, or equivalently:*

$$\aleph_0 \cdot \aleph_0 = \aleph_0.$$

Proof. It is easy to see that $(\omega \times \omega, <_{ml})$ is a a wellordering such that each (n, m) in $\omega \times \omega$ has finitely many predecessors. It follows by Corollary 6.9(2) that $(\omega \times \omega, <_{ml})$ and $(\omega, <)$ are isomorphic. In particular there is a bijection between $\omega \times \omega$ and ω. In other words $\aleph_0 \cdot \aleph_0 = \aleph_0$. □

Let us denote by Γ the isomorphism from ORD^2 onto ORD. In particular,

$$\Gamma(\alpha, \beta) = \text{ the order type of the set } \{(\gamma, \delta) \mid (\gamma, \delta) <_{ml} (\alpha, \beta)\} \tag{6.131}$$

We use the Γ function to prove Theorem 6.22. The proof proceeds by induction, and as the basic step uses the result proved above that $\aleph_0 \cdot \aleph_0 = \aleph_0$.

Theorem 6.22 *For every $\alpha \in \mathrm{ORD}$,*

$$\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha \tag{6.132}$$

Proof. We will show by induction that $\Gamma(\aleph_\alpha, \aleph_\alpha) = \aleph_\alpha$ for every α. Since Γ is a 1-1 function, it shows that $|\aleph_\alpha \times \aleph_\alpha| = \aleph_\alpha$. First note that a function γ defined by

$$\gamma(\alpha) = \Gamma(\alpha, \alpha) \tag{6.133}$$

is a normal function. It follows $\Gamma(\aleph_\alpha, \aleph_\alpha) \geq \aleph_\alpha$ by Lemma 5.10.

We will argue that it leads to contradiction if we assume that there is some α such that

$$\Gamma(\aleph_\alpha, \aleph_\alpha) > \aleph_\alpha. \tag{6.134}$$

Let α be the least ordinal where (6.134) occurs.

α cannot be 0 because by Lemma 6.21, $\Gamma(\aleph_0, \aleph_0) = \aleph_0$.

So $\alpha > 0$ and by the induction assumption for all $\beta < \alpha$, $\Gamma(\aleph_\beta, \aleph_\beta) = \aleph_\beta$. The assumption (6.134) implies that there are some $\delta_0, \delta_1 < \aleph_\alpha$ such that $\Gamma(\delta_0, \delta_1) = \aleph_\alpha$. Define $\delta = \max(\delta_0, \delta_1) + 1$. Since $\aleph_\alpha$ is a limit ordinal by Lemma 6.3(2), $\delta < \aleph_\alpha$, and so in particular $|\delta| < \aleph_\alpha$. Also, $(\delta_0, \delta_1) \in \delta \times \delta$. Since $(\delta_0, \delta_1) \leq_{ml} (\delta, \delta)$, we obtain $\aleph_\alpha = \Gamma(\delta_0, \delta_1) \leq \Gamma(\delta, \delta)$, which implies $|\delta \times \delta| = |\delta| \cdot |\delta| \geq \aleph_\alpha$.

However, by the induction assumption we also have that $|\delta| = |\delta| \cdot |\delta| < \aleph_\alpha$, which is a contradiction. □

Corollary 6.23 *For every $\alpha, \beta \in \mathrm{ORD}$,*

$$\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \max(\aleph_\alpha, \aleph_\beta) \tag{6.135}$$

Proof. Consider the following inequalities:

$$\max(\aleph_\alpha, \aleph_\beta) \leq \aleph_\alpha + \aleph_\beta \leq \aleph_\alpha \cdot \aleph_\beta \leq (\max(\aleph_\alpha, \aleph_\beta) \cdot \max(\aleph_\alpha, \aleph_\beta)) = \max(\aleph_\alpha, \aleph_\beta). \tag{6.136}$$

□

6.4.2 Number of finite sequences in an infinite set

If X is a set, we denote by $X^{<\omega}$ the set of all finite sequences in X:

$$X^{<\omega} = \bigcup\{X^n \,|\, n < \omega\}.$$

If $|X| = \aleph_\alpha$, then we write

$$(\aleph_\alpha)^{<\omega} = |X^{<\omega}|.$$

Lemma 6.24 *The following holds for every α:*

$$(\aleph_\alpha)^{<\omega} = \aleph_\alpha.$$

Proof. Let $|X| = \aleph_\alpha$. By induction on $n < \omega$, it holds by Theorem 6.22 that $|X^n| = \aleph_\alpha$. By the argument in the proof of Lemma 6.19, the union of at most $\aleph_\alpha$ many sets each of size at most $\aleph_\alpha$ is at most $\aleph_\alpha$, and so:

$$\aleph_\alpha \leq |X^{<\omega}| = |\bigcup\{X^n \,|\, n < \omega\}| = (\aleph_\alpha)^{<\omega} \leq \aleph_\alpha,$$

and so $\aleph_\alpha = (\aleph_\alpha)^{<\omega}$ as desired. □

Example. Let L be a first-order language with $\aleph_\alpha$ symbols. Then the number of all formulas in the language L is at most $(\aleph_\alpha)^{<\omega}$. Since $(\aleph_\alpha)^{<\omega} = \aleph_\alpha$ by the above argument, it follows that the number of all L-formulas is exactly $\aleph_\alpha$. Note that in the most common case where L has $\aleph_0$ symbols (variables $v_0, v_1, \ldots$, and finite number of functional and relational symbols), this says that there are countably many formulas in the language L.

6.5 Continuum function, continuum hypothesis (CH), and generalized continuum hypothesis (GCH)

The function which to every cardinal $\aleph_\alpha$ assigns the cardinal $2^{\aleph_\alpha}$ is called the *continuum function.* The bahaviour of this function was and is one of the central themes of set theory. We know from Cantor's theorem one thing:

$$\text{For every } \alpha, \aleph_\alpha < 2^{\aleph_\alpha}. \tag{6.137}$$

But the question is, what more can we say?

In 1900 David Hilbert, a distinguished German mathematician, listed the problem "what is the cardinal $2^{\aleph_0}$"[11] as the first problem for the next century. The originator of set theory, another German mathematician Georg Cantor, conjectured that $2^{\aleph_0}$ is the least cardinal greater than $\aleph_0$:

$$\text{Continuum hypothesis, CH: } 2^{\aleph_0} = \aleph_1. \tag{6.138}$$

[11] Recall that $2^{\aleph_0}$ denotes the size of $\mathscr{P}(\omega)$ which has the same size as $\mathbb{R}$; the Hilbert's question then reads, "how many real numbers there are"?

This can be generalized to:

(6.139) Generalized continuum hypothesis, GCH: $(\forall\alpha)2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

Using König's theorem (6.26), one can show that the cofinality of $2^{\aleph_\alpha}$ must be greater than $\aleph_\alpha$. It follows we can show the following three properties of the continuum function:

Theorem 6.25 *Continuum function satisfies for very $\alpha, \beta \in \mathrm{ORD}$:*

(1) $\alpha < \beta \to 2^{\aleph_\alpha} \leq 2^{\aleph_\beta}$;
(2) (Cantor's theorem) $\aleph_\alpha < 2^{\aleph_\alpha}$;
(3) (consequence of König's theorem) $\aleph_\alpha < \mathrm{cf}(2^{\aleph_\alpha})$.

*Exercise**

1. Because for every α, $\mathrm{cf}(\aleph_\alpha) \leq \aleph_\alpha$, the property (3) implies Cantor's theorem (2). So in fact, only two properties of the continuum function are captured in Theorem (6.25): (1),(3).

For more than 30 years mathematicians tried to prove more about the continuum function than included in Theorem 6.25, but they failed. Only in early 30's, Kurt Gödel managed to prove that if ZF is consistent, so is ZF + AC + GCH.[12] However, this just showed that is is *possible* that GCH holds.

In early 60's, Paul Cohen managed to showed[13] that if ZF is consistent so is ZF + AC + ¬CH, and ZF + ¬AC. This showed that the axioms of ZF and ZFC are too weak to decide the validity of CH and GCH.

In early 70's, William Easton finally managed to show that the properties identified in Theorem 6.25 are *the only properties* one can show about the continuum function in ZFC for *regular cardinals*[14]. For instance the following is consistent with ZFC:

(i) $2^{\aleph_0} = \aleph_2$;
(ii) $2^{\aleph_1} = \aleph_2$;
(iii) $2^{\aleph_2} = \aleph_{117}$;
(iv) $2^{\aleph_3} = \aleph_{\aleph_{\omega+1}}$, etc.

It follows that if we want to know more about the continuum function, new and more powerful axioms must be added to ZFC. This is a long process, and there is no undivided opinion about which axioms should be added. However, at least the following agreement seems to be settled among mathematicians: if anything, GCH seems to be *false* in our intuition (because it presents too neat a picture which does contradict some otherwise intuitively acceptable axioms).

6.6 Optional topics

This section will be extended in Winter 2013

6.6.1 König's theorem and infinite sums and products

If $\{\kappa_i \mid i \in I\}$ is a set of cardinal numbers for some infinite set I, we define the infinite sum

(6.140) $$\textstyle\sum_{i\in I} \kappa_i = |\bigcup_{i\in I} X_i|,$$

[12] In fact ZF proves that GCH implies AC.

[13] He developed the technique of *forcing* to prove this theorem, which since then has become the major set-theoretic tool for mathematicians if they want to derive consistency results.

[14] Situation for singular cardinals is more complex.

where $\{X_i \,|\, i \in I\}$ is a disjoint family of sets such that $|X_i| = \kappa_i$ for each $i \in I$ (one can show with AC that this definition does not depend on the choice of X_i's). One can show the following for λ an infinite cardinal and $\kappa_i > 0$ for each $i < \lambda$:

$$\textstyle\sum_{i<\lambda}\kappa_i = \lambda \cdot \sup(\{\kappa_i \,|\, i < \lambda\}). \tag{6.141}$$

Moreover, one can show: Assume that $\{X_i \,|\, i < \aleph_\alpha\}$ is a family of non-empty sets such that $|X_i| \leq \aleph_\alpha$ for each $i < \aleph_\alpha$. Then

$$\textstyle\sum_{i<\aleph_\alpha} |X_i| \leq \aleph_\alpha. \tag{6.142}$$

We can also define infinite products. Recall that if $\{X_i \,|\, i \in I\}$ is a family of non-empty sets, we define the product $\prod_{i\in I} X_i$ as follows:

$$\textstyle\prod_{i\in I} X_i = \{f \,|\, f \text{ a function} : I \to \bigcup_{i\in I} X_i \text{ such that } (\forall i \in I) f(i) \in X_i\}. \tag{6.143}$$

If $\{\kappa_i \,|\, i \in I\}$ is a family of cardinal numbers, we define the infinite product:

$$\textstyle\prod_{i\in I} \kappa_i = |\prod_{i\in I} X_i|, \tag{6.144}$$

where $\{X_i \,|\, i \in I\}$ is a family of sets such that $|X_i| = \kappa_i$ for each $i \in I$ (by AC, the definition of the product does not depend on the particular X_i's).

Let λ be an infinite cardinal and $\langle \kappa_i \,|\, i < \lambda\rangle$ a non-decreasing sequence of non-zero cardinals, then:

$$\textstyle\prod_{i<\lambda} \kappa_i = (\sup(\{\kappa_i \,|\, i < \lambda\}))^\lambda. \tag{6.145}$$

Infinite sums and infinite products are connected by the following important theorem:

Theorem 6.26 (König) *If $\kappa_i < \lambda_i$ for every $i \in I$ where I is non-empty, then*

$$\textstyle\sum_{i\in I} \kappa_i < \prod_{i\in I} \lambda_i. \tag{6.146}$$

For proof, see [BS] str. 181.

Corollaries.

(i) $2^\kappa > \kappa$ (Hint. Set $\kappa_i = 1$ for each i, and $\lambda_i = 2$, and use König's lemma.)

(ii) $\mathrm{cf}(2^\kappa) > \kappa$.
Assume $\langle \kappa_i \,|\, i < \mu\rangle$ is cofinal in 2^κ, then

$$2^\kappa = \sum_{i<\mu} \kappa_i < \prod_{i<\mu} 2^\kappa = (2^\kappa)^\mu. \tag{6.147}$$

For $\mu \leq \kappa$, this implies

$$(2^\kappa)^\mu = |{}^{(\kappa\times\mu)}2| = 2^\kappa,$$

which contradicts the strict $<$ in (6.147) above. Hence $\mu > \kappa$, and so $\mathrm{cf}(2^\kappa) > \kappa$.