Department of Logic FFUK

Set theory

Radek Honzík

University text for a two-semester course in Set theory at Department of Logic, Faculty of Arts, Charles University in Prague.

Version: May 30, 2022

Note. Sections 1–5 are covered in the first term of the course. The remaining sections are covered in the second term of the course.

Contents

1 An axiomatic framework for set theory

1.1 Axioms of set theory

An excellent and detailed account can be found in [1].[1] We just review the basic 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, in this more concrete context: "Why an axiomatic set theory, i.e. why a 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 the "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 axiomatisation we attempt to list all properties 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.

[ZF1] 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).

[ZF2] Extensionality.

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

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

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

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

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

Exercise. Realize that

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

[ZF2] Pairing.

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

[1] Relevant parts of the book are available as PDF copies on my webpage; you need to know the password to access them.

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

Exercise. Show

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

[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. 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 we require of an ordered pair.

Lemma 1.2 *The following holds*

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

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

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

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 holds that $\langle x, y\rangle \neq \langle v, w\rangle$. 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:

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

Let us shorten the expression in (1.5) as A & 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 identities 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. □

By induction[2] we can define an ordered n-tuple as follows: $\langle a_0\rangle = a_0$, and then $\langle a_0, a_1, \ldots, a_k\rangle = \langle\langle a_0, a_1, \ldots, a_{k-1}\rangle, a_k\rangle$. The analogue of Lemma 1.2 is shown by induction.

Fact 1.3

$$\text{ZF1}, \text{ZF2} \vdash (\forall x_0, \ldots, y_0, \ldots) \\ [\langle x_0, \ldots, x_k\rangle = \langle y_0, \ldots, y_k\rangle \leftrightarrow x_0 = y_0 \;\&\; \ldots \;\&\; x_k = y_k]. \tag{1.6}$$

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 or Comprehension scheme. Let $\varphi(q, p)$ be a formula with two free variables q and p. Then (1.7) is an axiom of ZF (hence there are infinitely many axioms in (1.7) – one for each $\varphi(p, q)$).

$$(\forall x, p)(\exists z)(\forall q)[q \in z \leftrightarrow q \in x \;\&\; \varphi(q, p)]. \tag{1.7}$$

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

$$z = \{q \mid q \in x \;\&\; \varphi(q, p)\} \tag{1.8}$$

for z in (1.7).

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

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

There are good reasons to forbid the quantification over formulas: consider the *Berry's paradox*: Since there are infinitely many natural numbers, there are certainly natural numbers which cannot be defined by any combination of 100 letters or less, and we can take the least such. If n is the "least number which cannot be defined by any combination of 100 letters or less", then n *is* defined by less than 100 letters after all (the definition appears between "..." above), 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.

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

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: $\langle p_0, \dots, p_n \rangle$ is just a single set):

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

$$(\forall x, p_0, \dots, p_n)(\exists z)(\forall q)[q \in z \leftrightarrow q \in x \; \& \; \psi(q, p_0, \dots, p_n)]. \tag{1.9}$$

We now 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*

$$z = \{q \,|\, q \in x \; \& \; q \notin q\}.$$

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 Foundation (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, i.e. $V = \{x \,|\, x = x\}$. Then we do obtain contradiction from the existence of set $z = \{q \,|\, q \in V \; \& \; q \notin q\}$ 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 $x \cap y = \{q \,|\, q \in x \; \& \; q \in y\}$, difference of two sets $x \setminus y = \{q \,|\, q \in x \; \& \; q \notin y\}$.[3]
- Emptyset: $\emptyset = \{q \,|\, q \in x \; \& \; q \neq q\}$, 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 $z = \{q \,|\, q \in x \; \& \; q \neq q\}$ above satisfies $\varphi(z)$; (ii) it can be shown that there can be at most one set with no element: 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* if $x \cap y = \emptyset$.

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

– Intersection (generalisation of intersection):

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

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

$$\bigcap x = \{q \mid q \in y \ \& \ (\forall q')(q' \in x \to q \in q')\}. \tag{1.11}$$

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.

$$(\forall x)(\exists z)(\forall q)[q \in z \leftrightarrow (\exists y)(y \in x \ \& \ q \in y)]. \tag{1.12}$$

We introduce the following abbreviations:

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

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

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

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.

$$(\forall x)(\exists z)(\forall q)(q \in z \leftrightarrow q \subseteq x). \tag{1.13}$$

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

We can form a new unary operation:

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

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

$$\mathscr{P}(x) \subseteq x \tag{1.14}$$

Fix such an x. Consider the set z defined in Lemma 1.6. z is clearly a subset of x. By our assumption (1.14) 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:

  (1.15) $x \times y = \{q \,|\, q \in \mathscr{P}(\mathscr{P}(x \cup y)) \;\&\; (\exists q_0, q_1)(q = \langle q_0, q_1 \rangle \;\&\; q_0 \in x \;\&\; q_1 \in y)\}.$

  The product is a set because $\mathscr{P}(\mathscr{P}(x \cup y))$ 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** 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** of r as

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

  and similarly we define **range** of r as

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

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

  We also define the **inverse** of r as

  (1.18) $$r^{-1} = \{\langle q, q' \rangle \,|\, \langle q', q \rangle \in r\}.$$

  and if $a \subseteq x$ we define the **image of r over a**:

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

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

  (1.20) $$r \restriction a = \{\langle q, q' \rangle \,|\, \langle q, q' \rangle \in r \;\&\; q \in a\}.$$

  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).[4] 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 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) $x''(y \cup z) = x''y \cup x''z$,

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

(c)

$$x''(y \cap z) \subseteq x''y \cap x''z \text{ and } x''y \setminus x''z \subseteq x''(y \setminus z). \tag{1.21}$$

Give an example where the converse inclusion $\supseteq$ does not hold in (1.21). Compare with (1.25).

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

  $$r \circ s = \{\langle x, z\rangle \mid (\exists y)\langle x, y\rangle \in r \ \&\ \langle y, z\rangle \in s\}. \tag{1.22}$$

  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:

  $$(\forall x, y, \bar{y})(\langle x, y\rangle \in r \ \&\ \langle x, \bar{y}\rangle \in r \to y = \bar{y}). \tag{1.23}$$

  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 $f : x \to y$ and $g : y \to z$ 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 $f \circ g : x \to z$ such that for each $q \in x$, $f \circ g\ (q) = g(f(q))$.

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

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

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

  – $f : x \to y$ is *onto* y if $\mathrm{rng}(f) = y$ (if f is onto, we also call it *surjective*).

  Exercises.

  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):

$$(f^{-1})''(x \cap y) = (f^{-1})''x \cap (f^{-1})''y \quad \text{and } (f^{-1})''(x \setminus y) = (f^{-1})''x \setminus (f^{-1})''y. \tag{1.24}$$

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

$$f''(x \cap y) = f''x \cap f''x \text{ and } f''(x \setminus y) = f''x \setminus f''y. \tag{1.25}$$

[Notice that this is a strengthening of the results in (1.21); you can use the results in (1.21) here.]

- **Index sets**. 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

$$a = \{\bar{a}(j) \,|\, j \in i\} \tag{1.26}$$

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

$$a = \{a_i \,|\, i \in I\}. \tag{1.27}$$

Compare with (1.38).

[ZF6] Axiom of infinity.

$$(\exists x)[\emptyset \in x \;\&\; (\forall q)(q \in x \to q \cup \{q\} \in x)]. \tag{1.28}$$

Under all reasonably definitions of *finiteness*, a set in the axiom ZF6 is infinite.

Note that a set x in (1.28) is note determined uniquely, there are more sets satisfying (1.28). If x satisfies (1.28), we say that x is *inductive*. We will define the set of natural numbers $\mathbb{N}$, also denoted as ω, as follows:

$$\omega = \mathbb{N} = \bigcap \{x \,|\, x \text{ is inductive}\}. \tag{1.29}$$

Remark 1.9 We are not entitled to use the operation $\bigcap$ here according to (1.10) 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*. S is a proper class. Another example of a proper class is the *universe of all sets*, 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 not everything we use with sets is meaningful with classes: for instance we can write $x \in A$, but not $A \in B$, and most importantly we *must not* quantify over proper classes. More about classes is in Subsection 1.2.

[5] So if $\varphi(x, p_0, \ldots p_n)$ is a formula and $p_0, \ldots, p_n$ are sets then $\{q \,|\, \varphi(q, p_0, \ldots, p_n)\}$ 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 generalised so that it can be applied 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

$$\bigcap A = \{q \mid q \in y \,\&\, (\forall q')(\varphi_A(q', p_0, \dots, p_n) \to q \in q'\} \tag{1.30}$$

and so $\bigcap A$ is a set by the axiom of separation (because of the expression "$q \in y$" in the definition of $\bigcap A$ in (1.30)).

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

For the following Lemma, we define the notion of the least set in $\subseteq$: We say that x is the least set in the inclusion relation in a class A if for all $a \in A$, $x \subseteq a$.

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

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 $\mathbb{N} = \bigcap\{x \mid x \text{ is inductive}\} \subseteq y$ 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. We say that a formula $\varphi(u, v, p)$ determines a function (compare with (1.23)) if

$$(\forall p, u, v_0, v_1)[\varphi(u, v_0, p) \,\&\, \varphi(u, v_1, p) \to v_0 = v_1]. \tag{1.31}$$

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 $\{\langle u, v\rangle \mid \varphi(u, v, p)\}$; 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 for the formula $\varphi(u, v, p)$:

$$(\forall p)(\forall x)(\exists z)(\forall q)[q \in z \leftrightarrow (\exists q' \in x)\varphi(q', q, p)]. \tag{1.32}$$

For each formula $\varphi(u, v, p)$ which determines a function, the formula in (1.32) 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:

$$\text{For every set } x, F[x] \text{ is a set.} \tag{1.33}$$

Note the following properties:

- As in the scheme of Separation, we can show that more parameters as in the formula $\varphi(u, v, p_0, \dots, p_n)$ 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 $\psi(u, v, p) = \varphi(u, p) \ \& \ u = v$. We can show that Axiom of Replacement applied to $\psi(u, v, p)$ proves the existence of a set $a = \{q \in x \mid \varphi(q, p)\}$.

- 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 $(u = \emptyset \ \& \ v = a) \vee (u = \{\emptyset\} \ \& \ v = b)$ 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 $\{v \mid u \neq v\}$ 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.

$$(\forall x)[x \neq \emptyset \to (\exists q)(q \in x \ \& \ x \cap q = \emptyset)]. \tag{1.34}$$

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 $y_0 \in y_1 \in \dots \in y_n \in y_0$, for the same reason.
– There can be no infinite $\in$-chain: $y_0 \ni y_1 \ni y_2 \ni \dots$. (If there were such, then ZF8 would fail for $x = \{y_i \mid i \in \omega\}$).

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).

$$(\forall x)(\exists f)[(f \text{ is a function with } \mathrm{dom}(f) = x - \{\emptyset\}) \ \& \\ (\forall q)[(q \in x \ \& \ q \neq \emptyset) \to f(q) \in q)]]. \tag{1.35}$$

Such f is called a *choice function* (for x). 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 $p_0, \dots, p_n$; if $\varphi(u, p_0 \dots p_n)$ 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, $q \in A \leftrightarrow q \in B$.
- $A \cap B$, $A \cup B$, $A - B$.
- The 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*.
- $A \times B = \{\langle a, b\rangle \mid a \in A \ \& \ b \in B\}$.

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 & , $\vee, \neg$), we first show that the above formulas (i)–(v) hold for the propositional connectives $\&, \vee, \neg$, 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 $p \ \& \ q$ for conjunction and $p \vee q$ for disjunction. $p \ \& \ q$ 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 $\&, \vee, \neg, 0, 1$ satisfy the formulas (i)–(v) above. [Hint. These are just propositional tautologies.]

As soon as we know that $\&, \vee, \neg, 0, 1$ 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: $q \in a \ \& \ q \in b$ is equivalent to $q \in b \ \& \ q \in a$; however this is true because the conjunction & 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 $x, y \in B$:

$$x \leq_B y \leftrightarrow x \wedge y = x \leftrightarrow x \vee y = y. \tag{1.36}$$

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 $\mathscr{P}(x)$):

(1.37)

(1) $- - a = a$
(2) $-a = -b \rightarrow a = b$
(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 (1.26).

It follows we can write

$$a = \{q \mid q \in a\} = \{a_i \mid i \in I\} \tag{1.38}$$

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

$$\begin{aligned} \bigcup\{a_i \mid i \in I\} &= \bigcup a \\ \bigcap\{a_i \mid i \in I\} &= \bigcap a \end{aligned} \tag{1.39}$$

The lefthand side of (1.39) is sometimes written as $\bigcup_{i\in I} a_i$, and $\bigcap_{i\in I} a_i$.

Lemma 1.16 *Let $\{a_i \mid i \in I\}$ 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 ψ & $[(\exists x)\varphi(x)]$ is logically equivalent to $(\exists x)[\psi \;\&\; \varphi(x)]$ 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.] □

2 Comparing sizes

2.1 Relation "to be bigger than" for infinite objects

Let R be a binary relation on a class A, i.e. $R \subseteq A \times A$. We call R a *partial ordering* or shortly an *ordering* (on A) if it satisfies the following properties for all $x, y, z \in A$:
(i) $\langle x, x\rangle \in R$ (reflexivity)
(ii) $\langle x, y\rangle \in R \;\&\; \langle y, z\rangle \in R \to \langle x, z\rangle \in R$ (transitivity)
(iii) $\langle x, y\rangle \in R \;\&\; \langle y, x\rangle \in R \to x = y$ (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 $\langle y, x\rangle \in R$.

Exercise. A binary relation R' on A is called *a strict ordering* if it is an irreflexive transitive relation, where irreflexive means that $\langle x, x\rangle \notin R'$ for any x. Show that if R is an ordering on A and we define R' as follows:

$$\langle x, y\rangle \in R' \text{ iff } \langle x, y\rangle \in R \;\&\; x \neq y, \tag{2.40}$$

then R' is a strict ordering.

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 $\langle x, y\rangle \in \leq$. Similarly, a strict ordering is denoted by $<$, etc. 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. $x \subseteq y$ & $x \neq y$) 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

$$n < m \text{ iff } x \prec y. \tag{2.41}$$

2.2 Basic properties

Lemma 2.2 *(i) The relation $\approx$ is an equivalence 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 $f : x \to y$, then $f^{-1} : y \to x$ 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*. The important Theorem 2.7 connects the relations $\approx$ and $\preceq$ in the natural way and shows that although $x \preceq y$ & $y \preceq x$ 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 equivalence, this means that x and y are in the same equivalence class.

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

Before we show Theorem 2.7, 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) *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

$$C = \{u \subseteq x \mid u \subseteq H(u)\} \tag{2.42}$$

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

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

First notice that if u is in 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.

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

We will apply monotonicity of H to (2.43), obtaining $H(c) \subseteq H(H(c))$. This means that $H(c)$ is an element of C, and so in particular

$$H(c) \subseteq c. \tag{2.45}$$

(2.43) and (2.45) 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 *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 *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 $\bigcap x$. □

Remark 2.6 The notion of supremum and infimum is widely used in mathematics (for instance in analysis, we use the fact that every nonempty bounded subsets of $\mathbb{R}$ has the supremum and the infimum (in the usual ordering on $\mathbb{R}$): for instance $\{\frac{1}{n} \mid n \in \omega\}$ does not have the least element, but has the infimum which is equal to 0 in this case).

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

Theorem 2.7 (Cantor, Bernstein) *For every* x, y*:*

$$x \approx y \leftrightarrow (x \preceq y \;\&\; y \preceq x). \tag{2.46}$$

Proof. The direction from left to right in (2.46) is obvious, so we need to prove the converse: $(x \preceq y \;\&\; y \preceq x) \rightarrow x \approx y$.

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 $f[a] = \{b \in y \mid (\exists a' \in a) f(a') = b\}$; 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 $\mathscr{P}(y)$ 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:

$$H(u) = x \setminus g[y \setminus f[u]], \text{ for every } u \subseteq x. \tag{2.47}$$

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 (1.37), 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:

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

It implies that

$$x \setminus c = g[y \setminus f[c]]. \tag{2.49}$$

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. (2.49) 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 $f[c] \cup g^{-1}[x \setminus c] = y$. 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*.

Theorem 2.8 (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

$$a = \{y \in x \,|\, y \notin g(y)\}. \tag{2.50}$$

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 2.9 *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 2.18 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 2.10 *Let x, y be arbitrary sets, then we write ${}^x y$ to denote the following set:*

$${}^x y = \{f \,|\, f : x \to y\}, \tag{2.51}$$

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 $x \neq \emptyset$, then ${}^x\emptyset = \emptyset$.

Remark 2.11 The x, y in the definition of ${}^x y$ can be finite. We can take Definition 2.10 as a definition of the usual exponentiation 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:

$$f_1 \neq f_2 \Leftrightarrow \text{ there is some } q \in v \text{ such that } f_1(q) \neq f_2(q). \tag{2.52}$$

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 2.12 *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) $(x \approx x_1 \;\&\; y \approx y_1) \to (x \times y) \approx (x_1 \times y_1)$,
(iv) $x \approx y \to \mathscr{P}(x) \approx \mathscr{P}(y)$,
(v) $\mathscr{P}(x) \approx {}^{x}2$, *where* $2 = \{\emptyset, \{\emptyset\}\}$.
(vi) $x^2 \approx {}^{2}x$. *The exponentiation ${}^{y}x$ can thus be viewed as a generalisation 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 $g[a] = \{b \in y \mid (\exists q \in a) g(q) = b\} \subseteq y$.

Ad (v). If $y \subseteq x$ is a subset, then we define the *characteristic function of y* $\chi_y : x \to 2$ 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 $f : \mathscr{P}(x) \to {}^{x}2$ 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 $f : x^2 \to {}^{2}x$ by assigning to each $\langle a, b\rangle$ the function $\{\langle 0, a\rangle, \langle 1, b\rangle\}$. □

Lemma 2.13 *Let x, y, u, v be sets:*
(i) $\emptyset \neq x \preceq y \to {}^{x}u \preceq {}^{y}u$,
(ii) $u \preceq v \to {}^{y}u \preceq {}^{y}v$,
(iii) ${}^{(x \times y)}u \approx {}^{x}({}^{y}u) \approx {}^{y}({}^{x}u)$.

Proof.

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 ${}^{x}u \approx {}^{y}u$. Let $g : x \to y$ be 1-1. Define $f : {}^{x}u \to {}^{y}u$ 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 $f : {}^{y}u \to {}^{y}v$ 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 $f : {}^{x\times y}u \to {}^{x}({}^{y}u)$. The bijection between ${}^{x}({}^{y}u)$ and ${}^{y}({}^{x}u)$ 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 $h' : x \to {}^{y}u$ 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 2.14 *For all x, y, u, v:*
(i) $(x \approx y \,\&\, v \approx u) \to {}^{x}u \approx {}^{y}v$,
(ii) $(\emptyset \neq x \preceq y \,\&\, u \preceq v) \to {}^{x}u \preceq {}^{y}v$.

2.3 The size of 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) exists 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 non-empty set A is *bounded below* if it has a lower bound, and is *bounded above* if it has an upper bound. We say that A is bounded if it is bounded below and above.

Definition 2.15 *We say that a linearly ordered set $\langle X, <\rangle$ is* order-complete *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 2.16 *Let $\langle X, <\rangle$ be a linearly ordered set. 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 supremum 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.

(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 2.17 $\mathbb{R}$ *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 preparation for the proof of Theorem 2.18, we show the following lemma which concerns geometrical progressions. We call a sequence (a_n) which is of the form $a_0, a_0r, a_0r^2, \ldots$ for some a_0 in $\mathbb{R}$ and $r \in \mathbb{R}$ a geometrical progression. We will only be interested in the case when $0 < r < 1$. If (a_n) is a geometrical progression, we denote by s_n the sum of its first n elements:

$$s_n = \textstyle\sum_{i=0}^{n-1} a_i = a_0 + a_0r + a_0r^2 + \cdots + a_0r^{n-1}. \tag{2.53}$$

Assuming $0 < r < 1$ one can show

$$s_n = a_0 \cdot \frac{1-r^n}{1-r} = \frac{a_0}{1-r} - \frac{a_0}{1-r}r^n. \tag{2.54}$$

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

$$\lim(a_n + b_n) = \lim a_n + \lim b_n, \text{ and } \lim(qa_n) = q \lim a_n. \tag{2.55}$$

Using this, we can see that:

$$\text{the limit of } (s_n) \text{ exists and } \lim(s_n) = \frac{a_0}{1-r}. \tag{2.56}$$

Argue as follows: by (2.55), the limit of (s_n), if it exists, is using the expression in (2.53), equal to $\frac{a_0}{1-r} - \frac{a_0}{1-r}\lim(r^n)$. It is not difficult to check 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 2.18 *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 2.7, 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 $\{q_n \,|\, i \in \omega\}$ 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 $n \in \omega$:

$$f(x)(n) = 1 \text{ if } q_n < x, \text{ or } f(x)(n) = 0 \text{ if } x \le q_n. \tag{2.57}$$

We need to show 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 the 0's and 1's 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 $F(x \restriction n) \le F(x \restriction (n+1))$ 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 and has therefore a supremum. This follows from the claim in (2.56): 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 (2.56). 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 $x \in {}^{\omega}2$:

$$g(x) = \sup\{F(x \restriction n) \,|\, n \in \omega\}. \tag{2.58}$$

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, $g(x) \le a + \sum\{1/3^{i+1} \,|\, n \le i, i \in \omega\}$, so it remains to see that

$$\sum\{1/3^{i+1} \,|\, n \le i, i \in \omega\} < \frac{1}{3^n}. \tag{2.59}$$

Applying again (2.56) with $a_0 = \frac{1}{3^{n+1}}$ and $r = \frac{1}{3}$, the sum

$$\sum\{1/3^{i+1} \mid n \leq i, i \in \omega\}$$

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 $r = 1/n$, where n is a natural number ≥ 3. Argue that the argument would not work with $r = 1/2$. [Hint. Study the validity of the inequality in (2.59) 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 2.19 *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 2.20 *Can we assign 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* 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* (AC) 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* 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 (1.26). 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

$$\prod\langle a_i \,|\, i \in I\rangle = \{f \,|\, f : I \to \bigcup\{a_i \,|\, i \in I\}, (\forall i \in I) f(i) \in a_i\}. \tag{3.60}$$

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 sets. Let f be a choice function for s; then $\bar{f}$ with domain $\mathrm{dom}(r)$ defined by

$$\bar{f}(q) = f(r''\{q\}), \text{ for every } q \in \mathrm{dom}(r) \tag{3.61}$$

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

$$\langle i, q\rangle \in r \text{ iff } i \in I \;\&\; q \in a_i, \text{ for every } \langle i, q\rangle \in I \times \bigcup\{a_i \,|\, i \in I\}. \tag{3.62}$$

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

$$\prod\langle z \,|\, z \in y\rangle = \{f \,|\, f : y \to \bigcup y, (\forall z \in y)\; f(z) \in z\}. \tag{3.63}$$

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

Remark 3.2 In this subsection, we will consider not only sets, but also (proper) classes. However, if A is a proper class and $\leq$ is an ordering on A, we will often require that for every $a \in A$, the class of the predecessors of a is a *set*:

$$\forall a \in A \; \{b \in A \,|\, b \leq a\} \text{ is a set.}$$

If $\langle A, \leq \rangle$ satisfies this condition, we say that the ordering $\leq$ on A is *set-like.*

Definition 3.3 *We say that a partially ordered class $\langle M, \leq \rangle$ is* well-ordered *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 3.4 *Every well-ordered class $\langle M, \leq \rangle$ 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 classes are useful because they can be easily compared.

Definition 3.5 *Let $\langle A, \leq \rangle$ and $\langle B, \trianglelefteq \rangle$ be two partially ordered classes. 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*[7] *$f : A \to B$ 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 3.6 *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 *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 $\leq$) 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* 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 $\langle \mathrm{rng}(f), \trianglelefteq \rangle$.

[7] A careful reader might correctly ask whether we can quantify over proper classes (if A and B are proper classes, so is f). Strictly speaking, we cannot quantify over classes, and the definition should be formulated only for sets; however, often the bijection f may itself be definable in which case the definition makes sense even for proper classes.

Lemma 3.7 *Let $\langle A, \leq \rangle$ and $\langle B, \trianglelefteq \rangle$ be well-ordered classes with set-like ordering. 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 $f(x) \triangleleft g(x)$ or $g(x) \triangleleft f(x)$. W.l.o.g. assume that $f(x) \triangleleft g(x)$ 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 $g(y) \triangleleft g(x)$, 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 set-like classes can be easily compared.

Theorem 3.8 *Let $\langle A, \leq \rangle$ and $\langle B, \trianglelefteq \rangle$ be well-ordered classes with set-like ordering. 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 3.7, 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 3.7, 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 $\operatorname{dom}(F') \neq A$ and $\operatorname{rng}(F') \neq B$. If F' were strictly bigger than F, then $\operatorname{dom}(F')$ would need to be bigger than A, or $\operatorname{rng}(F')$ would need to be bigger than B. However, this would mean that F' does not satisfy the conditions of the Theorem. □

Corollary 3.9 *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 3.8, 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 it as a new axiom.

Definition 3.10 Well-ordering Principle, WO *is the following statement*

Every set can be well-ordered.

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

Theorem 3.11 WO *implies* AC. *That is*

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

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 6.2.

Corollary 3.12 WO *implies that the relation $\preceq$ for comparing sizes is linear on the universe V.*

Proof. Immediate by Corollary 3.9. □

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

3.2.1 A bit of history

Definition of a well-ordered set: Burali-Forti. The notion of a well-ordered set was formulated by Cantor. Due to its second-order character, it was first not understood very well. In particular, Burali (another mathematician) tried to reformulate it: Burali thought that his definition of a perfectly ordered set is stronger than Cantor's, while it was in fact weaker. (Note that Burali in a short time realised his error.) It is instructive to try to find out where Burali made a mistake. For more details see [FG], p.105.

We say that $(P, <)$ is *perfectly ordered* if

(i) P is linearly ordered.
(ii) P has the least element.
(iii) Every $p \in P$ which has a successor has an immediate successor.
(iv) Every p satisfies the following: if p has an immediate predecessor, then there exists $q < p$ which has no immediate predecessor and the number of z such that $q < z < p$ is finite.

Exercises.

(1) Argue that every finite $(P, <)$ is perfectly ordered.
(2) Argue that every well-ordered set is perfectly ordered. [Hint. by contradiction: given α, go down finding immediate predecessors; this process must end after finitely many steps, otherwise we find a subset which has no least element]
(3) Find an example of $(P, <)$ which is perfectly ordered, yet not well-ordered. It follows that the notion of a perfectly ordered set is strictly weaker than the notion of a well-ordered set. [Hint. Take $(\mathbb{Z}, <)$ and replace each number with a copy of $(\omega, <)$.]

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 set. We say that $X \subseteq A$ is a *chain* if the ordering $\leq$ is linear on X, i.e. for all $x, y \in X$, $x \leq y$ or $y \leq x$.

Definition 3.13 Principle of Maximality, PM *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 called *Zorn's lemma* in honour of the American mathematician (algebraist and group theorist) Max A. Zorn who first used this principle in 1935.

PM implies WO (see Theorem 3.14), 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 6.2. The equivalence of all these independently discovered notions is for practical considerations a powerful reason for believing that these notions are intuitively valid (true).

Theorem 3.14 PM *implies* WO*. That is*

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

The proof is included in Section 3.3.3

3.3.1 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 3.15 *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* 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 3.16 *$F \subseteq \mathscr{P}(A)$ is a filter iff:*

(i) $A \in F$.
(ii) *For all $X, Y \subseteq A$,*

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

Proof. If F is a filter according to Definition 3.15, 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 3.15, $X \in F$ 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 *Fréchet* 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 3.17 *We say that a system $E \subseteq \mathscr{P}(A)$ has the* finite intersection property, FIP *if for every $n \in \omega$ and all sequences $e_0, \ldots, e_n$ of elements in E it holds that*

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

Lemma 3.18 *Every $E \subseteq \mathscr{P}(A)$ with FIP can be extended into a proper filter.*

Proof. Define F as follows:

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

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 3.19 *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 Fréchet 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 3.20 *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 3.20, 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 $F' \supseteq F$, $F' \neq F$, 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 3.21 *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 3.20, 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 3.20, 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 3.22 *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 if $X \in F$ and $X \subseteq Y$, then $Y \in F$ for every X, Y. It remains 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 3.18, 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 3.21, 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 $X, Y \subseteq A$:

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

3.3.2 Ramsey theorem and ultrafilters

As an illustration of the use of ultrafilters, we will prove the following theorem.[8]

Theorem 3.23 (Ramsey) *For every partion Q of $[\omega]^r$, $r \geq 1$, to finitely many pieces k, $k \geq 1$, there exists an infinite set A such that $[A]^r \subseteq q$ for some q in Q.*

Proof. (For more details, see Balcar and Stepanek, p. 277.) The proof will be by induction on r. For $r = 1$ and arbitrary $1 \leq k < \omega$, this is just Dirichlet's principle: if ω is written as a finite union of subsets of ω, then at least one such subset must be infinite.

So assume the theorem holds $r \geq 1$, we show it for $r + 1$. Fix a uniform ultrafilter[9] U on ω. Let $f : [\omega]^{r+1} \to k$ determine the partition.

(i) For any $u \in [\omega]^r$ and $i < k$, set

$$X(u, i) = \{x \in \omega - u \mid f(u \cup \{x\}) = i\}.$$

Let $g(u)$ be the unique $i < k$ such that $X(u, g(u)) \in U$.

[8] The theorem can be proved also without the ultrafilters, but the ultrafilter construction provides more control, and can be generalised.

[9] An ultrafilter on ω is *uniform* if all its elements are infinite.

(ii) Hence $g : [\omega]^r \to k$ determines a partition.

(iii) Using g, we construct $Y \subseteq \omega$ which will contain a homogenous set A for f. Choose first r elements of Y arbitrarily. Suppose for $n \geq r$, we have $Y_n = \{y_j \mid j < n\}$ already constructed. For every $u \in [Y_n]^r$, $X(u, g(u)) \in U$, and therefore

$$X_n = \bigcap \{X(u, g(u)) \mid u \in [Y_n]^r\}$$

is in U and is therefore infinite. Therefore $X_n - \{p \mid p \leq y_{n-1}\}$ is non-empty. Let y_n be the least element of this set. Y is the union of Y_n's, $n < \omega$.

(iv) Consider $g : [Y]^r \to k$. By induction assumption, there is $A \subseteq Y$ homogeneous for g.[10]

(v) Suppose $g(x) = 0$ for all $x \in [A]^r$. We show that $g(x) = 0$ for all $[A]^{r+1}$, proving the theorem (if $g(x)$ is not 0 but other number, the proof is the same). Let $v \in [A]^{r+1}$ be arbitrary. We know that $A \subseteq Y$. Let y be the greatest element of v and $v = u \cup \{y\}$. Then $g(u) = 0$ since $u \in [A]^r$, and $y \in X(u, 0)$, and therefore $f(v) = f(u \cup \{y\}) = 0$.

The proof is finished. □

3.3.3 PM implies WO

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

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

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:

[10] If $h : Y \to \omega$ is a bijection, then h can be used to define a partition g' on $[\omega]^r$ such that if A is homogeneous for g', $h^{-1}[A]$ is homogeneous for g.

(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 $R \subseteq R'$, 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

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

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 Introduction to ordinal numbers

4.1 Basic properties

The notion of well-ordering is very important because it allows us to define new objects by a *recursive construction* – for instance one can define the function $n \mapsto 2^n$ by recursion on ω as follows: $2^0 = 1$ and $2^{n+1} = 2 \cdot 2^n$.[11] If we wish (and we do wish it) to use recursion to construct new objects in an infinite number of steps, we need something "longer" than ω, yet with the same nice structure – in other words, we wish generalize ω to a well-ordered class which contains ω as its initial segment. This leads to the definition of *ordinal numbers.* In this section, we will focus mostly on the definitions and will not give the proofs.

We wish to define ordinals so that they satisfy the following properties:

[11] The words "induction" and "recursion" are often used with the same meaning. However, there is tendency to prefer "recursion" for constructions of various objects and "induction" for proofs (as in a "proof by induction").

(1) An ordinal will be a set of the form $\langle \alpha, < \rangle$, where $<$ is a strict well-ordering 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 well-ordered by the 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 well-ordering.

Before giving the definition of the ordinal number in Definition 4.5, we review briefly the notion of restriction of a relation to a set (Definition 4.1) and of a transitive set (Definition 4.3). These notions are important for other areas of set theory as well; we therefore give some simple properties as well (Lemma 4.2, and 4.4).

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*

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

To get familiar with the definition, let us prove the following:

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

Proof. Recall that $<$ is a strict well-ordering 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 *if*

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

Examples. $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$ and V are transitive classes. In general, if x and y are transitive, so are $x \cap y$ and $x \cup y$. Is $\{\{\emptyset\}\}$ transitive?

There is another, equivalent, way of saying when x is transitive:

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

Proof. ($\rightarrow$). 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. □

Now we come to the important definition of an ordinal number.

Definition 4.5 *A set x is called an* ordinal*, or* ordinal number *if x is a transitive set and the restriction*

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

is a strict well-ordering on x.

The class of all ordinals will be denoted ORD:

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

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 often write $<$ instead of $\in$ for the ordering on the ordinal numbers (mostly because there is the convention to denote orderings by $<$; however do not forget that the ordering is in fact $\in$).

Here are some important facts about ordinal numbers:[12]

Fact 4.6 *Let α, β, γ be ordinal numbers.*

(i) $\alpha \notin \alpha$.

(ii) $\alpha \in \beta \;\&\; \beta \in \gamma \rightarrow \alpha \in \gamma$.

(iii) $\alpha \cup \{\alpha\} =_{df} \alpha + 1$ *is an ordinal number, and* $\alpha + 1$ *is the immediate successor of* α *in the ordering* $\in$ *on* ORD.

(iv) If $A \subseteq \mathrm{ORD}$ *is a non-empty class, then there exists* $\alpha \in A$ *which is the least element of* A *in* ORD.

(v) If a *is a subset of* ORD*, then* $\bigcup a$ *is the supremum of* a *in* ORD.

(vi) ORD *is a proper class.*

Fact (4.6) says that $\in$ is a strict ordering (i,ii) on the ordinal numbers which satisfies that for every α there is an immediate successor $\alpha + 1$ (iii), and moreover $\in$ well-orderes ORD in a very strong sense: not only every non-empty set has the least element, it also holds that every non-empty class has the least element (iv). (v) says that the ordering on ORD has suprema for all sets of ordinal numbers.

Let us state the following Corollary:

Corollary 4.7 $\in$ *is a set-like well-ordering in* ORD.

Here are some more facts:

Fact 4.8 *(i) All natural numbers* $n \in \omega$ *are ordinal numbers, and* ω *itself is an ordinal number.* ω *is the supremum of* $\omega \subseteq \mathrm{ORD}$. ω *is the initial segment of ordinal numbers.*

[12] We will not give proofs for the Facts concerning ordinal numbers. However, an interested student should read the proofs in [1].

(ii) For every α in ORD*, α is an initial segment of ordinal numbers and $\langle \alpha, \in \rangle$ is a well-ordered set.*

Ordinals can be divided into three types:

Fact 4.9 (Three types of ordinals) *Every ordinal α falls into exactly one of the following three types:*

(i) $\alpha = \emptyset$; in this case α is the least element of ORD *and is often denoted simply as* 0.

(ii) α is of the form $\beta + 1$ for some $\beta < \alpha$.

(iii) If α is not 0, nor equal to $\beta + 1$, then we call α a limit ordinal*. For instance ω is a limit ordinal.*

Exercise. Show that α is a limit ordinal iff for every $\beta < \alpha$, $\beta + 1 < \alpha$.

All elements of ORD smaller than ω are called *finite* ordinals. Finite ordinals can be identified with natural numbers $0, 1, 2, 3, \ldots$. We say that a set x is *finite* if there is a bijection between x and some finite ordinal.

The following fact captures one of the most important properties of ORD which are used in practice.

Fact 4.10 (Representation of well-ordered sets)
For every well-ordered set $\langle x, < \rangle$ there exists one and only one ordinal α such that

$$\langle x, < \rangle \text{ and } \langle \alpha, \in \rangle \text{ are isomorphic.} \tag{4.87}$$

Remark 4.11 All the properties of ordinal numbers can be shown just in ZF without the Axiom of Foundation and the Axiom of Choice. This is not important for us right now, but will have some significance in more advanced lectures. This contrasts with the properties of cardinal numbers (introduced below) which heavily rely on AC.

4.2 Addition and multiplication of ordinal numbers

We can define the usual operations on ordinal numbers, such as $\alpha + \beta, \alpha \cdot \beta$, and α^β. The definitions of $+$ and $\cdot$ have a "geometric" motivation. To explain this motivation, we first define a well-ordering of pairs of ordinal numbers:

$$(\alpha_0, \beta_0) <_l (\alpha_1, \beta_1) \leftrightarrow \alpha_0 < \alpha_1 \vee (\alpha_0 = \alpha_1 \;\&\; \beta_0 < \beta_1). \tag{4.88}$$

We call $<_l$ the *lexicographical* ordering of ORD^2.

Lemma 4.12 *The ordering $<_l$ is a strict well-ordering on* ORD^2.

Proof. It is clearly antireflexive and transitive (exercise). We will show that it is a well-ordering. So let $A \subseteq \mathrm{ORD}^2$ be a non-empty set, we wish to show that there is some pair (α_0, β_0) such that (α_0, β_0) is the $<_l$-least element in A. Let us define:

$$\alpha_0 = \text{the } < \text{-least element of } \{\alpha \mid (\exists \beta)\, (\alpha, \beta) \in A\},$$

and

$$\beta_0 = \text{the } < \text{-least element of } \{\beta \mid (\alpha_0, \beta) \in A\}.$$

This definition is correct because both sets are non-empty if A was non-empty; it is also obvious that $(\alpha_0, \beta_0) \in A$. If $(\alpha, \beta) \neq (\alpha_0, \beta_0)$ and $(\alpha, \beta) \in A$, then by the construction of α_0, β_0, we have that either $\alpha_0 < \alpha$, in which case $(\alpha_0, \beta_0) <_l (\alpha, \beta)$, or $\alpha_0 = \alpha$ but $\beta_0 < \beta$, and hence again $(\alpha_0, \beta_0) <_l (\alpha, \beta)$. It follows that (α_0, β_0) is the $<_l$-least element of A. □

We define:
Addition.

$$\alpha + \beta = \text{ the unique ordinal isomorphic to } \langle (\{0\} \times \alpha) \cup (\{1\} \times \beta), <_l \rangle. \tag{4.89}$$

Multiplication.

$$\alpha \cdot \beta = \text{ the unique ordinal isomorphic to } \langle \beta \times \alpha, <_l \rangle. \tag{4.90}$$

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 (4.90), it *does* matter whether we write $\beta \times \alpha$ or $\alpha \times \beta$.

The definition of exponentiation α^β has no immediate geometric intuition; it is defined using a transfinite recursion; see Section 6.3.

4.3 Transfinite induction

Ordinal numbers are a generalisation of natural numbers, and therefore make it possible to formulate a more general version of the induction-based arguments available for natural numbers. Since the induction on ORD goes beyond all $n < \omega$ into infinite stages, we call it a *transfinite induction.*

The following theorem generalizes the induction theorem for natural numbers which states that if $A \subseteq \omega$ contains $\emptyset$ and is closed under the successor operation, then $A = \omega$:

Theorem 4.13 *Let A be a subclass of* ORD *such that for every ordinal α holds:*

$$\alpha \subseteq A \rightarrow \alpha \in A,$$

then $A =$ ORD.

Proof. For contradiction assume that α is the least ordinal which is not in A. Then $\alpha \subseteq A$ (because every β smaller than α is in A, and "smaller" in ORD means $\in$), and by assumption $\alpha \in A$. Contradiction. □

This can be reformulated for successor and limit ordinals:

Theorem 4.14 *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) For a limit ordinal α: $[(\forall \beta < \alpha)\beta \in A] \to \alpha \in A$,
then $A =$ ORD.

Recall the inductive construction of 2^n for $n < \omega$: we define $2^0 = 1$, and $2^{n+1} = 2 \cdot 2^n$. The correctness of this definition follows from the following theorem on construction by induction on ω: if $G : \omega \to \omega$ is a function and $a \in \omega$, then there is a unique $F : \omega \to \omega$ such that $F(0) = a$, and $F(n+1) = G(F(n))$. In our example, setting $a = 1$, and $G = 2n$, we get the function 2^n. This theorem generalizes as follows:

Theorem 4.15 *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 $\alpha \in$* ORD*:*

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

The following simpler form of Theorem 6.1 is often used:

Theorem 4.16 *Let G be a class function from V to V and a an arbitrary set. Then there is a unique function F from* ORD *to V such that*
(i) $F(0) = a$,
(ii) For every α, $F(\alpha + 1) = G(F(\alpha))$,
(iii) For every limit α, $F(\alpha) = G(F \restriction \alpha)$.

For the proof and applications, see Sections 6 and 6.2.

5 Introduction to cardinal numbers

Cardinal numbers are sets which we assign to other sets to measure their size. This works similarly as for the finite sets: to a set containing say three elements a, b, c we assign number 3, etc.

It turns out that with the Axiom of Choice, the best way to measure the size of sets is to use *ordinal numbers* introduced above. In particular, Theorem 4.10 seems to be helpful in measuring sets: if we can well-order a set x, we can "measure" it with an ordinal number. But there is a slight problem with measuring size according to Theorem 4.10: the ordinal which measures the set is determined by the well-order. Thus it is possible that x is well-ordered by two orderings $<_1$ and $<_2$, $\langle x, <_1 \rangle$ and $\langle x, <_2 \rangle$, but the unique ordinals are different:

$$\langle x, <_1 \rangle \cong \langle \alpha_1, \in \rangle \text{ and } \langle x, <_2 \rangle \cong \langle \alpha_2, \in \rangle \text{ and } \alpha_1 \neq \alpha_2. \tag{5.91}$$

Exercise. Consider for instance ω, which can be ordered by an ordering isomorphic to $\omega + 1$ but also to ω (more options are possible).

Exercise If x is finite, then the ordinal α is unique (i.e. does not depend on the well-ordering) and is equal to some $n < \omega$.

The undesirability of the non-uniqueness of the ordinal in (5.91) is solved by taking the least ordinal which can measure x in the sense of (5.91); see Definition 5.1.

5.1 The definition

Our intuition (extrapolated from the finite sets) says that the cardinal numbers should 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 $|x| = |y|$.

As we mentioned above, it is not convenient to measure size of set by an ordinal given by Theorem 4.10 because the ordinal is not unique. To ensure properties (i)–(iii) above, we make the following definition:

Definition 5.1 (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:

$$|x| = \text{ the least ordinal } \alpha \text{ such that } \alpha \approx x. \tag{5.92}$$

5.2 Basic properties

The Axiom of Choice is very important in dealing with cardinals. Some of the results below are provable without AC, but we will not keep track of this. From this time on we assume AC and often use it without mentioning.

Lemma 5.2 *(i) $|x|$ is defined for every x and is unique.*

(ii) For every set x, the size $|x|$ is an element of CARD*.*

(iii) For every set x, $|x| \approx x$.

(iv) For all sets x, y, $x \approx y \to |x| = |y|$.

Proof. Ad (i,ii). Let $(x, <)$ be any well-ordering of x (there is one by AC which implies WO). By Theorem 3.8 there is a unique ordinal α such that $(x, <) \cong (\alpha, \in)$, which implies $x \approx \alpha$. Since ORD is well-ordered, 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 (iii). By definition of $|x|$.

Ad (iv). 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 $|x| = n$. If there is no $n < \omega$ such that $|x| = n$, then x is called *infinite*.

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

Lemma 5.3 *The following hold about cardinals:*

(i) ω is the least infinite cardinal.

(ii) Every infinite cardinal is a limit ordinal.

(iii) For every cardinal κ there is a cardinal λ such that $\lambda > \kappa$.

(iv) If $\langle \kappa_\xi \,|\, \xi < \alpha \rangle$ is an increasing sequence of cardinals for α a limit ordinal, then the supremum $\bar{\kappa} = \sup(\{\kappa_\xi \,|\, \xi < \alpha\})$ is a cardinal.

Proof. Ad (i). Every natural number $n < \omega$ is a cardinal because one can show – by induction – that there can be no bijection between m, n for $m \neq n$. ω is the supremum of natural numbers, and it is a cardinal because there can be no bijection between ω and a natural number.

Ad (ii). 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 (iii). 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 (iv). 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 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 (iii) and (iv) together imply that cardinal numbers are unbounded in the ordinal numbers, i.e. for every $\alpha \in \mathrm{ORD}$, 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^+.

5.3 Definition of addition, multiplication, and exponentiation on cardinals

We define the following basic operations for cardinal numbers:

Definition 5.4 (Addition.) *Let κ and λ be cardinals. We define*

$$\kappa + \lambda = |\kappa + \lambda|, \tag{5.93}$$

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 7.3).

Definition 5.4 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 5.5 (Multiplication.) *Let κ and λ be cardinals. We define*

$$\kappa \cdot \lambda = |\kappa \cdot \lambda|, \tag{5.94}$$

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 7.3).

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 5.6 *For all cardinal numbers κ, λ such that $1 < \kappa, \lambda$:*

$$\kappa + \lambda \leq \kappa \cdot \lambda \tag{5.95}$$

Proof. Given X and Y such that $|X| > 1$ and $|Y| > 1$, we need construct a 1-1 function from the disjoint union of X and Y to the product $X \times Y$. This is easy, Exercise. □

Definition 5.7 (Exponentiation.) *Let κ and λ be cardinals. We define*

$$\kappa^\lambda = |\{f \mid f : \lambda \to \kappa\}|, \tag{5.96}$$

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

$$\kappa^\lambda = |{}^\lambda\kappa|. \tag{5.97}$$

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 7.7.

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 5.8 *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:

$$g(x) = \chi_x, \text{ where } \chi_x(\xi) = 1 \text{ if } b^{-1}(\xi) \in x, \text{ and } \chi_x(\xi) = 0 \text{ otherwise,} \tag{5.98}$$

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. □

5.4 Alephs

We have learned above that the class of all cardinal numbers is unbounded and continuous[13] in the class of all ordinal numbers. Since ordinal numbers are well-ordered, we can enumerate all cardinal numbers, one by one, giving them ordinal numbers as indices. This enumeration is called the function *aleph*, and denoted $\aleph$. For α in the domain of

[13] By continuity we mean that if $\kappa_0 < \kappa_1 < \ldots$ is an increasing sequence of cardinals, then their limit (in the ordinal numbers), is a cardinal; see Lemma 5.3(iv).

$\aleph$, we write $\aleph_\alpha$ instead of $\aleph(\alpha)$. For practical reasons, we start the enumeration at ω, ignoring the finite ordinals. Thus $\aleph_0 = \omega$ (the first infinite cardinal), $\aleph_1$ = the second infinite cardinal, $\aleph_2$ = the third infinite cardinal, ..., $\aleph_\omega$ = the ω-th infinite cardinal, etc.

How do we define $\aleph_\alpha$ formally for all ordinals α? We can use for instance transfinite recursion, see Theorem 4.16: By recursion define function $\aleph : \mathrm{ORD} \to \mathrm{CARD}$:

(5.99)

$$\begin{aligned}\aleph_0 &= \omega,\\ \aleph_{\alpha+1} &= \text{the least cardinal strictly greater than } \aleph_\alpha,\\ \aleph_\lambda &= \text{the supremum of the cardinals } \{\aleph_\beta \,|\, \beta < \lambda\}, \text{ for } \lambda \text{ limit.}\end{aligned}$$

Notice that the definition of the limit is correct because the supremum of cardinals is a cardinal (continuity of cardinals) by Lemma 5.3(iv).

5.5 Cardinal addition and multiplication

The following theorem will be proved in Section 7.1. It says that addition and multiplication of infinite cardinals is very simple. This contrasts with the situation for exponentiation, see Section 5.6.

Theorem 5.9 *For all cardinals κ and λ, at least one of them infinite:*

$$\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa, \lambda). \tag{5.100}$$

5.6 Exponentiation and the Continuum Hypothesis

There is very little one can prove about exponentiation κ^λ in ZFC. Some more results above those in Lemma 5.10 can be proved, see Section 7.6, but in general the cardinal exponentiation is a very complex subject.

Lemma 5.10 *Let κ, λ be infinite cardinals. Then the following hold:*

(i) $\kappa < 2^\kappa$.

(ii) $\kappa < \lambda \to 2^\kappa \leq 2^\lambda$.

Proof. Ad (i). By Cantor's theorem $\kappa < |\mathscr{P}(\kappa)|$. By Lemma 5.8, $|\mathscr{P}(\kappa)| = 2^\kappa$.

Ad (ii). This follows from the fact that from the assumption $\kappa < \lambda$, one can find a 1-1 function from $\mathscr{P}(\kappa)$ to $\mathscr{P}(\lambda)$. Note that in general, we can only show $2^\kappa \leq 2^\lambda$, but not $2^\kappa < 2^\lambda$. □

Recall that there is a bijection between $\mathbb{R}$ and $\mathscr{P}(\omega)$. This size of real numbers therefore equals to the size of $\mathscr{P}(\omega)$, which is $2^{\aleph_0}$. By Lemma 5.10(i), $\aleph_1 \leq 2^{\aleph_0}$. Can we say more about $2^{\aleph_0}$?[14]

David Hilbert formulated in 1900 the following conjecture, called the *Continuum Hypothesis,* CH:

$$2^{\aleph_0} = \aleph_1. \tag{5.101}$$

[14] Since $2^{\aleph_0}$ is a cardinal, there must be some ordinal α such that $2^{\aleph_0} = \aleph_\alpha$. The question is, can we find out which α it is?

It took 62 years to find out (Göedel, Cohen) that CH is not provable nor disprovable from ZFC if ZFC is consistent. CH is thus the most famous *independent sentence* over ZFC.[15] In other words, ZFC does not provide enough information to decide which α is such that $2^{\aleph_0} = \aleph_\alpha$.

The Continuum Hypothesis can be generalised to the so called *General Continuum Hypothesis,* GCH:

$$(\forall \alpha \in \mathrm{ORD})\ 2^{\aleph_\alpha} = \aleph_{\alpha+1}. \tag{5.102}$$

GCH is also independent over ZFC.

6 More on ordinal numbers and transfinite recursion

6.1 Transfinite recursion theorem

Above, we have stated a theorem about transfinite recursion without giving a proof, see Theorems 4.15 and 4.16. We now give the proof:

Theorem 6.1 *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 approximations. Set

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

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 4.13 (localised 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 4.14 to make things more clear. Let

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

[15] There are many more independent sentences over ZFC; by the way, AC itself is independent over ZF.

(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 4.14, $A = \mathrm{ORD}$ as desired.

Properties (i)–(iii) suffice to finish the proof of 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

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

as desired.

By transfinite induction 4.13 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 $\{b \,|\, (\exists a \in x) f(a) = b\}$.

The above theorem has several variants.

Theorem 6.2 *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 6.1 to G': there is a unique F such that

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

□

Another variant (compare with Theorem 4.13 and 4.14):

Theorem 6.3 *Let G_1 and G_2 be two functions 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 6.1 to this G. □

6.2 AC, WO, and PM are all equivalent

Recall that Theorem 3.11 shows that WO implies AC, and Theorem 3.14 shows that PM implies WO. To complete the equivalence between these principles, we will show now that AC implies PM.

Theorem 6.4 *Axiom of Choice implies Principle of Maximality:*

$$\mathsf{ZF} \vdash \mathsf{AC} \rightarrow \mathsf{PM}. \tag{6.103}$$

Corollary 6.5 AC*,* PM*,* WO *are all equivalent, i.e.*

$$\mathsf{ZF} \vdash \mathsf{AC} \leftrightarrow \mathsf{WO} \leftrightarrow \mathsf{PM}. \tag{6.104}$$

Proof. (of Theorem 6.4) Let $(P, \leq)$ be a partially ordered set which satisfies the condition that every chain in P has an upper bound (see Definition 3.13 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 transfinite recursion.

Define by transfinite recursion function $F : \mathrm{ORD} \rightarrow P$:

$$\begin{aligned} F(0) &= p, \\ F(\alpha+1) &= C(\{q \in P \,|\, F(\alpha) < q\}), \text{ if } \{q \in P \,|\, F(\alpha) < q\} \text{ is non-empty,} \\ &= p, \text{ otherwise,} \\ \lambda \text{ limit, } F(\lambda) &= C(\{q \in P \,|\, (\forall \beta < \lambda) F(\beta) < q\}), \\ &\quad \text{if } \{F(\beta) \,|\, \beta < \lambda\} \text{ is a strictly increasing chain,} \\ &= p, \text{ otherwise.} \end{aligned}$$

The following claims hold, proving the theorem:

- There is unique $\alpha \in \mathrm{ORD}$ such that $F(\alpha+1)$ is equal to p, and for all β, $0 < \beta \leq \alpha$, $F(\beta) \neq p$, and $\{F(\beta) \,|\, 0 \leq \beta \leq \alpha\}$ is a strictly increasing chain of elements in P. Denote this α by α_0.
- The maximal element in P above p is equal to $F(\alpha_0)$.

The proof of these claims is left as an exercise for the reader. □

Remark 6.6 A simple modification of the proof shows directly that AC implies WO. (Exercise.)

6.3 Ordinal arithmetics – definition of operations

We have seen above in Section 4.2 that addition and multiplications have geometrical motivations. In this section we define addition and multiplication also by recursion; recursion has the advantage that it can be used to define the ordinal exponentiation α^β as well.

Definition 6.7 (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) $\alpha + \beta = \sup\{\alpha + \xi \mid \xi < \beta\}$, for limit β.

Definition 6.8 (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) $\alpha \cdot \beta = \sup\{\alpha \cdot \xi \mid \xi < \beta\}$, 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.

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

Definition 6.9 (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) $\alpha^\beta = \sup\{\alpha^\xi \mid \xi < \beta\}$, for limit β.

Exercises.

1. * Try to visualize the following ordinal numbers: $\omega < \omega \cdot 2 < \omega \cdot 3 < \omega \cdot \omega = \omega^2 < (\omega^2) + \omega < \omega^3 < \omega^\omega < \omega^{\omega^\omega} = \omega^{(\omega^\omega)} < \epsilon_0 = \sup\{g(n) \mid n < \omega\}$, where $g(n)$ is defined by recursion as follows: $g(0) = \omega$, and $g(n+1) = \omega^{g(n)}$. [All these numbers still have the size ω.]

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

Theorem 6.10 (Cantor's Normal Form Theorem.) *Every ordinal $\alpha > 0$ can be represented uniquely in the form:*

$$\alpha = \omega^{\beta_1} \cdot k_1 + \ldots + \omega^{\beta_n} \cdot k_n, \tag{6.105}$$

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.

Remark 6.11 Notice that the normal form for ordinals is reminiscent of the decadic "normal form" for numbers: any number $l \in \omega$ can be uniquely written as $l = 10^{n_1} \cdot k_1 + \ldots + 10^0 \cdot k_n$. In Cantor's normal form theorem, the base is ω (instead of 10 as in the decadic representation).

6.4 Normal functions and fixed points

Definition 6.12 *We say that a function* $f : \mathrm{ORD} \to \mathrm{ORD}$ *is* normal *if the following hold:*

(i) f *is increasing:* $(\forall \alpha, \beta)\alpha < \beta \to f(\alpha) < f(\beta)$.

(ii) f *is continuous: for every limit ordinal* γ,

$$f(\gamma) = \sup\{f(\beta) \mid \beta < \gamma\}. \tag{6.106}$$

Addition and multiplication are normal functions in the second variable (see Exercises below); but not in the first variable. Exponentiation is continuous in the exponent (see Exercises below).

Definition 6.13 *Let* $f : \mathrm{ORD} \to \mathrm{ORD}$ *be given. We say that* α *is a fixed point of* f *if*

$$f(\alpha) = \alpha. \tag{6.107}$$

Remark 6.14 Fixed points are easy to find for normal functions (see Lemma 6.16 below). Note that fixed points for addition and multiplication are possible only because we deal with infinite numbers (limit ordinals). If we add finite numbers, there are of course no n and k greater than 0 such that $n + k = k$.

The following is a simple lemma concerning normal functions:

Lemma 6.15 *Let* f *be a normal function on* ORD, *then*

$$\textit{for every } \alpha, \alpha \leq f(\alpha). \tag{6.108}$$

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

$$f(\alpha_0) < \alpha_0 \tag{6.109}$$

Using the fact that f is increasing, we obtain from (6.109):

$$f(f(\alpha_0)) < f(\alpha_0) < \alpha_0, \tag{6.110}$$

which contradicts that α_0 was the least such that (6.109) holds. □

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

Lemma 6.16 (Fixed-point lemma) *If* f *is a normal function from* ORD *to* ORD, *then it has arbitrarily large fixed points, i.e. for every* γ *there is* $\alpha \geq \gamma$ *such that* $f(\alpha) = \alpha$.

Proof. Let γ be given. Define the following function g by induction on ω:

$$\begin{aligned} g(0) &= \gamma, \\ g(n+1) &= f(g(n)), \text{ for every } n \in \omega. \end{aligned}$$

We know by Lemma 6.15 that $f(\beta) \geq \beta$ for every β, and so $f(\gamma) \geq \gamma$. Now there are two possibilities.

Either $f(\gamma) = \gamma$, and γ itself is a fixed point f (and the proof is finished because we can set $\alpha = \gamma$).

Or $f(\gamma) > \gamma$. In this case one can use the fact that f is increasing to show by induction on $n < \omega$ that $g(n+1) > g(n)$ for every n. In other words, we have the following situation:

$$g(0) = \gamma < g(1) = f(\gamma) < g(2) = f(f(\gamma)) < \dots . \tag{6.111}$$

Set $\alpha = \sup\{g(n) \,|\, n < \omega\}$. We will show that $f(\alpha) = \alpha$, and this will finish the proof.

First note that α is a limit ordinal (this means that for every $\delta < \alpha$, $\delta + 1 < \alpha$; this is obvious because $\delta < g(n)$ for some n and therefore $\delta + 1 \leq g(n) < g(n+1) < \alpha$). By continuity of f (because f is normal), we know

$$f(\alpha) = \sup\{f(\beta) \,|\, \beta < \alpha\}. \tag{6.112}$$

Denote $A = \{f(\beta) \,|\, \beta < \alpha\}$ (A is thus a subset of ordinals). It suffices to show that $\alpha = \sup A$ because then $\alpha = f(\alpha)$. To argue that α is the supremum of A, we need to show two things:

(a) α is the upper bound of A, i.e. for every $\beta < \alpha$, $f(\beta) \leq \alpha$, and

(b) α is the least upper bound, i.e. for every δ which is an upper bound of A, we have $\alpha \leq \delta$.

We argue for (a) as follows: Let $\beta < \alpha$, and $f(\beta)$ be given. By the definition of α, there is some n such that $\beta < g(n) < \alpha$. It follows $f(\beta) < f(g(n)) = g(n+1) < \alpha$, and therefore α is indeed an upper bound of A.

To argue for (b), first notice that because the ordering on ORD is linear, the following are equivalent: (b) and (b'), where

(b') For every $\delta < \alpha$, there is some $\delta' \in A$, $\delta < \delta'$.

The items (b) and (b') are equivalent because with linearity the fact that δ is not an upper bound, i.e. $(\exists \delta' \in A)\ \delta' \not\leq \delta$, translates to the fact that $\delta' > \delta$.

Now (b') follows because if $\delta < \alpha$, then for some $g(n) \in A$, $\delta < g(n)$. □

Exercises.

1. * 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.

2. * 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.

3. * Using Lemma 6.16 argue that there are β, γ, δ such that:

   (i) $\omega + \beta = \beta$; argue using the construction in Lemma 6.16, that the least such $\beta \geq \omega$ is the ordinal $\omega \cdot \omega$.

   (ii) $\omega \cdot \gamma = \gamma$; argue using the construction in Lemma 6.16, 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 6.16) is denoted as ϵ_0 (see above).

4. * Using Lemma 6.16, argue that if $\delta = \omega^\delta$, then:

(i) $\omega + \delta = \delta$,
(ii) $\omega \cdot \delta = \delta$,
(iii) $\omega^\delta = \delta$.

[Hint: $\delta \leq \omega + \delta \leq \omega \cdot \delta \leq \omega^\delta = \delta$.]

6.5 ω as a domain of a model for PA

We will show that we can define in ZFC basic arithmetical operations such as $+, \cdot$ and relations such as $\leq$ on ω in such a way that the resulting structure satisfies all axioms of arithmetics.

Recall that *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 $\bar{x} = x_0, \ldots, x_{n-1}$, we add the following axiom:

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

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

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.

Recall that in Section 4.2 we have defined $+$ and $\cdot$ on ordinal numbers using the lexicographical ordering $<_l$ on ORD. Since ω is an initial segment of ORD, we can straightforwardly apply these results.

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 6.17 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 6.18 *ω with the operations above satisfies all the axioms of* PA.

The proof is in a sense obvious, but also long, and we will therefore omit it.

Corollary 6.19 ZF *proves the consistency of* PA *(where* PA *is formulated within* ZF*). This is denoted as:*

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

where $\ulcorner \mathsf{PA} \urcorner$ *denotes the formalisation of the usual axioms of* PA *within* ZF.

6.6 The well-founded universe

Using Theorem 6.3, let us define a class function V and a class WF, where WF is abbreviation for "well-founded":

$$\begin{aligned} V_0 &= \emptyset \\ V_{\alpha+1} &= \mathscr{P}(V_\alpha) \\ V_\lambda &= \textstyle\bigcup_{\alpha<\lambda} V_\alpha, \text{ if } \lambda \text{ is a limit ordinal} \\ \mathrm{WF} &= \textstyle\bigcup_{\alpha \in \mathrm{ORD}} V_\alpha \end{aligned} \tag{6.113}$$

Note that we write (as is customary) V_α instead of $V(\alpha)$. We will show below, see Theorem 6.25, 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 (6.113).

We first show some simple properties of WF:

Lemma 6.20 *(i) For each α, V_α is a transitive set.*
(ii) For each α, and every $\beta < \alpha$, $V_\beta \subseteq V_\alpha$.
(iii) For each α, $\alpha \subseteq V_\alpha$, and $\mathrm{ORD} \subseteq \mathrm{WF}$.

Proof. Claims (i) and (ii) will by shown together by induction following Theorem 4.14. Let us denote

$$A = \{\alpha \in \mathrm{ORD} \,|\, V_\alpha \text{ is transitive and } (\forall \beta < \alpha) V_\beta \subseteq V_\alpha\}. \tag{6.114}$$

We show that $A = \mathrm{ORD}$. We need to show:
(a) $\emptyset \in A$,
(b) If $\alpha \in A$, then $\alpha + 1 \in A$,
(c) If all $\beta < \lambda$ are in A, then $\lambda \in A$ (for λ limit).

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

Ad (c). 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 (b). 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 (iii). 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 6.21 WF *is a transitive class.*

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

We make the following useful definition.

Definition 6.22 *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 6.23 *Basic properties of the rank function:*

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

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

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

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

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 $\bar{\alpha} = \sup\{\mathrm{rank}(x) + 1 \mid x \in y\}$. 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 6.24 *Let x be a set and $x \subseteq \mathrm{WF}$, then there is $\alpha \in \mathrm{ORD}$ such that $x \in V_\alpha$, and hence $x \in \mathrm{WF}$.*

Proof. Consider the following class:

$$\bar{x} = \{\mathrm{rank}(y) \,|\, y \in x\} \subseteq \mathrm{ORD}. \tag{6.115}$$

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 $\mathrm{WF} = V$, which is a shorthand for the formula $(\forall x)(\exists \alpha \in \mathrm{ORD}) x \in V_\alpha$.

Theorem 6.25 *Let F denote the Axiom of Foundation and $\mathsf{ZF} - \mathsf{F}$ the theory ZF without F. Then*

$$\mathsf{ZF} - \mathsf{F} \vdash \mathsf{F} \leftrightarrow (V = \mathrm{WF}). \tag{6.116}$$

Proof. $(V = \mathrm{WF} \to \mathsf{F})$. 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

$$\bar{x} = \{\mathrm{rank}(y) \,|\, y \in x\}. \tag{6.117}$$

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 6.23 (ii)(a). This contradicts the fact that α is the least element of $\bar{x}$.

$(\mathsf{F} \to V = \mathrm{WF})$. 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 6.24, this means that $y \in \mathrm{WF}$, contradiction.

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

$$\mathsf{F} \text{ implies that every non-empty class has an } \in \text{-minimal element.} \tag{6.118}$$

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 6.26 *Let x be set, we define the* transitive closure *of x,* $\mathrm{trcl}(x)$, *by recursion*

$$\begin{aligned} \mathrm{trcl}_0(x) &= x, \\ \mathrm{trcl}_{n+1}(x) &= \bigcup \mathrm{trcl}_n(x), \\ \mathrm{trcl}(x) &= \textstyle\bigcup_{n\in\omega} \mathrm{trcl}_n(x). \end{aligned} \tag{6.119}$$

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

Note that we can use the technique of Theorem 6.25 to argue that $\mathrm{CON}(\mathsf{ZF} - \mathsf{F}) \Rightarrow \mathrm{CON}(\mathsf{ZF})$, i.e. that by adding Axiom of Foundation to our system, we will not add contradiction.

Remark 6.27 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 $\mathbb{Q}$ (*Dedekind cuts*), 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 $V_{\omega+\omega}$.

7 More on cardinal numbers

We assume AC in this section. Recall the definitions of addition and multiplications for cardinal numbers and the basic properties stated in Section 5 above. Here we provide more details.

7.1 Basic cardinal arithmetics – addition and multiplication

We define another ordering on ORD^2, which is called the *maximum-lexicographical ordering*, or the *canonical well-ordering* of ORD^2, and denoted $<_{ml}$.

We define $<_{ml}$ on ORD^2 as follows:

$$\begin{aligned} (\alpha_0, \beta_0) <_{ml} (\alpha_1, \beta_1) \leftrightarrow{} & \max(\alpha_0, \beta_0) < \max(\alpha_1, \beta_1) \vee \\ & (\max(\alpha_0, \beta_0) = \max(\alpha_1, \beta_1) \;\&\; (\alpha_0, \beta_0) <_l (\alpha_1, \beta_1)), \end{aligned} \tag{7.120}$$

where $<_l$ is defined in (4.88).

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 so

$$(\mathrm{ORD}, \in) \text{ is isomorphic with } (\mathrm{ORD}^2, <_{ml}) \tag{7.121}$$

Exercises.

1. Verify that $<_{ml}$ is a well-ordering, and that the class of all $<_{ml}$-predecessors is a set for any $(\alpha, \beta) \in \mathrm{ORD}^2$.
2. *Prove (7.121). Hint: use Theorem 3.8 (formulate it for clases).

Lemma 7.1 *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 well-ordering such that each (n, m) in $\omega \times \omega$ has finitely many predecessors. It is now easy to see that $(\omega \times \omega, <_{ml})$ and $(\omega, <)$ are isomorphic, by application of Theorem 3.8. 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, guaranteed by (7.121). In particular,

$$\Gamma(\alpha, \beta) = \text{ the order type of the set } \{(\gamma, \delta) \mid (\gamma, \delta) <_{ml} (\alpha, \beta)\} \tag{7.122}$$

We use the Γ function to prove Theorem 7.2. The proof proceeds by induction, and as the basic step uses the result proved above that $\aleph_0 \cdot \aleph_0 = \aleph_0$.

To simplify notation, let us denote for every ordinal $\alpha \in \mathrm{ORD}$,

$$\gamma(\alpha) = \Gamma(0, \alpha). \tag{7.123}$$

Note that for every (δ_0, δ_1),

$$(\delta_0, \delta_1) <_{ml} (0, \alpha) \text{ iff } \delta_0 < \alpha \text{ and } \delta_1 < \alpha. \tag{7.124}$$

It follows that $\gamma(\alpha)$ is the order-type of the set $\alpha \times \alpha$ in the ordering $<_{ml}$.

Theorem 7.2 *For every $\alpha \in \mathrm{ORD}$,*

$$\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha. \tag{7.125}$$

Proof. We will show by induction that

$$\gamma(\aleph_\alpha) = \aleph_\alpha \tag{7.126}$$

for every α. Since Γ is a 1-1 function, (7.126) together with (7.124) imply that there is a bijection between $\aleph_\alpha \times \aleph_\alpha$ and $\aleph_\alpha$.

First note that γ is a normal function. It follows $\gamma(\aleph_\alpha) \geq \aleph_\alpha$ by Lemma 6.15.

We will argue that it leads to contradiction if we assume that there is some α such that

$$\gamma(\aleph_\alpha) > \aleph_\alpha. \tag{7.127}$$

Let α be the least ordinal where (7.127) occurs. α cannot be 0 because by Lemma 7.1, $\gamma(\aleph_0) = \aleph_0$.

So $\alpha > 0$ and by the induction assumption for all $\beta < \alpha$, $\gamma(\aleph_\beta) = \aleph_\beta$. The assumption (7.127) 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 5.3(ii), we get $\delta < \aleph_\alpha$, 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 $\aleph_\alpha = \Gamma(\delta_0, \delta_1) \leq \gamma(\delta)$, 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 7.3 *For every $\alpha, \beta \in \mathrm{ORD}$,*

$$\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \max(\aleph_\alpha, \aleph_\beta). \tag{7.128}$$

Proof. Consider the following inequalities:

$$\begin{aligned}\max(\aleph_\alpha, \aleph_\beta) \leq \aleph_\alpha + \aleph_\beta \leq \aleph_\alpha \cdot \aleph_\beta \leq \\ (\max(\aleph_\alpha, \aleph_\beta) \cdot \max(\aleph_\alpha, \aleph_\beta)) = \max(\aleph_\alpha, \aleph_\beta).\end{aligned} \tag{7.129}$$

□

7.2 Regular and singular cardinals

Definition 7.4 *Let $(X, \leq)$ be a partially ordered set. Then $Y \subseteq X$ is called* cofinal *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 greatest element in α. To avoid such trivial cases, we will focus on limit ordinals α.

Definition 7.5 *We say that an ordinal β is the* cofinality *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:

$$\omega \leq \mathrm{cf}(\alpha) \leq \alpha. \tag{7.130}$$

Also: for every $X \subseteq \alpha$:

$$X \text{ is cofinal in } \alpha \leftrightarrow \sup X = \bigcup X = \alpha.$$

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.

Definition 7.6 *Let $\alpha \le \beta$ be two limit ordinals. We say that $f : \alpha \to \beta$ is cofinal if its range is cofinal in β.*

Note that if the cofinality of α is γ, then we can find an increasing cofinal function $f : \gamma \to \alpha$ (simply enumerate in the increasing order the elements of a cofinal subset $X \subseteq \alpha$ of order type γ). Conversely, if $f : \gamma \to \alpha$ is increasing and cofinal, then $\mathrm{cf}(\alpha) \le \gamma$ (because the range of f is a cofinal subset of α of order type γ).

Here are some basic properties of the notion of cofinality.

Lemma 7.7 *Let α be limit, then*

(i) $\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha)$,

(ii) $\mathrm{cf}(\alpha)$ *is an infinite cardinal.*

Proof. Ad (i). Denote $\beta = \mathrm{cf}(\alpha)$, $\gamma = \mathrm{cf}(\beta)$. By (7.130), $\gamma \le \beta$. To verify $\beta \le \gamma$, argue as follows. By our assumption, there exist $f : \beta \to \alpha$ increasing cofinal and $g : \gamma \to \beta$ increasing cofinal. Then $g \circ f : \gamma \to \alpha$ is increasing cofinal (let $\alpha' < \alpha$, then there is $\beta' < \beta$ such that $f(\beta') \ge \alpha'$, and $\gamma' < \gamma$ such that $g(\gamma') \ge \beta'$. So $f(g(\gamma')) \ge f(\beta') \ge \alpha$.) Thus $\beta \le \gamma$ as required.

Ad (ii). Suppose for contradiction there is a bijection $f : \kappa \to \mathrm{cf}(\alpha)$ for some $\kappa < \mathrm{cf}(\alpha)$. Using f, define $g : \kappa \to \mathrm{cf}(\alpha)$ by $g(\xi) = \sup f[\xi]$. Then g is non-decreasing, and therefore we have that the order-type of $\mathrm{rng}(g)$ is at most κ and $\mathrm{rng}(g)$ is cofinal in $\mathrm{cf}(\alpha)$. Let $h : \mathrm{cf}(\alpha) \to \alpha$ be increasing cofinal, then $h[\mathrm{rng}(g)]$ is a cofinal subset of α of order type $\le \kappa$, contradiction. □

The notion of cofinality is used to divide all cardinals into two disjoint groups:

Definition 7.8 *We say that a cardinal κ is* regular *if $\mathrm{cf}(\kappa) = \kappa$. If $\mathrm{cf}(\kappa) < \kappa$, we say that κ is* singular.

Example. For every α, $\mathrm{cf}(\aleph_\alpha)$ is a regular cardinal. This follows from the Lemma 7.7 (i), (ii).

Example. For every α, $\aleph_{\alpha+\omega}$ is singular because its cofinality is ω and $\omega < \aleph_{\alpha+\omega}$.

Example. What is the cofinality of $\aleph_1$, or in particular of $\aleph_{\alpha+1}$? We will now show that all these cardinals are regular.

We will first prove a theorem which is useful in its own right.

Theorem 7.9 *Let $\alpha \ge 0$. Any union of $\le \aleph_\alpha$ sets each of size $\le \aleph_\alpha$ has size $\le \aleph_\alpha$.*

Proof. We know that $|\aleph_\alpha \times \aleph_\alpha| = \aleph_\alpha$. See Theorem 7.2.

Let $\{X_\beta \,|\, \beta < \aleph_\alpha\}$ be sets such that $|X_\beta| \le \aleph_\alpha$ for each $\beta < \aleph_\alpha$. We will argue that

$$|\bigcup\{X_\beta \,|\, \beta < \aleph_\alpha\}| \le |\bigcup\{\{\beta\} \times X_\beta \,|\, \beta < \aleph_\alpha\}| \le |\aleph_\alpha \times \aleph_\alpha| = \aleph_\alpha.$$

All inequalities above are obvious except perhaps $|\bigcup\{\{\beta\} \times X_\beta \,|\, \beta < \aleph_\alpha\}| \le |\aleph_\alpha \times \aleph_\alpha|$. By AC, we can choose 1-1 functions $f_\beta : X_\beta \to \aleph_\alpha$ for each $\beta < \aleph_\alpha$ (these functions exists by the assumption that each X_β has size at most $\aleph_\alpha$). Define g as follows

$$g(\langle \beta, x \rangle) = \langle \beta, f_\beta(x) \rangle.$$

It is easy to verify that g is 1-1, which finishes the proof. □

Corollary 7.10 *$\aleph_{\alpha+1}$ is regular for every α.*

Proof. Assume that $\langle \xi_i \,|\, i < \eta \rangle$ is a sequence of ordinals in $\aleph_{\alpha+1}$ and $\eta < \aleph_{\alpha+1}$. We will show that the sequence cannot be cofinal in $\aleph_{\alpha+1}$. Since the sequence is arbitrary, it follows that $\aleph_{\alpha+1}$ is regular.

We can view $\{\xi_i \,|\, i < \eta\}$ as a collection of at most $\aleph_\alpha$ sets each of size at most $\aleph_\alpha$, and so by Theorem 7.9, the union (supremum) $\bigcup\{\xi_i \,|\, i < \eta\}$ must have size at most $\aleph_\alpha$, and so is strictly smaller than $\aleph_{\alpha+1}$. □

We have shown above that every $\aleph_{\alpha+1}$ is regular, and there are many singular cardinals as well (for instance $\aleph_{\alpha+\omega}$ for every α). Is true that every $\aleph_\gamma$ is singular if γ is limit?

Perhaps surprisingly, this question has probably no answer in ZFC. See the next Section 7.3.

7.3 Weakly inaccessible cardinals

Assume γ is a limit ordinal. Then one can easily show that

$$\mathrm{cf}(\aleph_\gamma) = \mathrm{cf}(\gamma). \tag{7.131}$$

Does it imply that $\mathrm{cf}(\aleph_\gamma) < \aleph_\gamma$, namely that $\aleph_\gamma$ is singular? Not really: all we can conclude is that if $\aleph_\gamma$ is regular, then $\aleph_\gamma = \gamma$. From Lemma 6.16 we know that there are many γ's such that

$$\aleph_\gamma = \gamma. \tag{7.132}$$

The question is is there a regular γ which satisfies (7.132)?

Definition 7.11 *κ is called weakly inaccessible if it is uncountable and simultaneously a limit and regular cardinal.*

The following lemma shows that weakly inaccessible cardinals are exactly the regular fixed points of the function $\aleph$.

Lemma 7.12 *For every $\gamma > 0$: $\aleph_\gamma$ is weakly inaccessible iff ($\aleph_\gamma = \gamma$ and γ is a regular limit cardinal).*

Proof. From left to right: if $\aleph_\gamma$ is weakly inaccessible and in particular a limit cardinal, then γ must be a limit ordinal. By regularity of $\aleph_\gamma$, $\aleph_\gamma = \mathrm{cf}(\aleph_\gamma) = \mathrm{cf}(\gamma) = \gamma$.

The converse direction is obvious from the definitions. □

Why should we consider such cardinals? There are many reasons, ranging from theoretical to practical, but one of the most important is that we do have one example if we allow ω: ω is simultaneously a limit and regular cardinal. The motivation behind the definition of a weakly inaccessible cardinal is that the universe of sets should be rich enough to allow another cardinals such as ω; a weakly inaccessible cardinal $\kappa > \omega$ can thus be viewed as another, "higher" infinity.

Fact 7.13 *It is consistent relative to* ZFC *that there are no weakly inaccessible cardinals. However, no one has shown that* ZFC + *"there is a weakly inaccessible cardinal" is inconsistent. In practice weakly inaccessible cardinals are widely used.*

One can define also a *strongly inaccessible* cardinal: κ is strongly inaccessible if for every cardinal $\lambda < \kappa$, $2^\lambda < \kappa$ (so in particular κ is a limit cardinal), and κ is regular. The property of being strongly inaccessibile is generally stronger than weakly inaccessible, but they are consistently the same (under GCH, the two notions define the same cardinals).

Remark 7.14 If GCH holds and κ is weakly inaccessible, then V_κ is the model of the formal version of ZFC. Thus ZFC + GCH + "there exists a weakly inaccessible cardinal" proves the consistency of (the formal version of) ZFC. Compare with the fact that ZFC proves the consistency of (the formal version of) PA.

7.4 Cardinal exponentiation

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 7.15 *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 7.2 that $|X^n| = \aleph_\alpha$. By the argument in the proof of Corollary 7.10, 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.

For the general values of the exponent, we limit ourselves to the following simple properties:

Lemma 7.16 *For every $\kappa \geq \omega$, $2^\kappa = \kappa^\kappa$.*

Proof.

$$2^\kappa \leq \kappa^\kappa \leq (2^\kappa)^\kappa = 2^{\kappa\cdot\kappa} = 2^\kappa. \quad (7.133)$$

□

Or more generally:

Lemma 7.17 *If* $2 \leq \kappa \leq \lambda$ *and* λ *is infinite, then* $\kappa^\lambda = 2^\lambda$.

Proof.

$$2^\lambda \leq \kappa^\lambda \leq (2^\kappa)^\lambda = 2^{\kappa \cdot \lambda} = 2^\lambda. \tag{7.134}$$

□

By definition, κ^λ is the set of all λ-sequences of elements in κ. If $\lambda \leq \kappa$, and we consider not λ-sequences, but subsets of κ of size λ, nothing will change with regard to size:

Let A be a set and $|A| \geq \lambda$, the we write

$$[A]^\lambda = \{X \subseteq A \,|\, |X| = \lambda\}. \tag{7.135}$$

Lemma 7.18 *If* $|A| = \kappa \geq \lambda \geq \omega$*, then the set* $[A]^\lambda$ *has cardinality* κ^λ.

Proof. Clearly $|[A]^\lambda| \leq |A|^\lambda$ because if $X \subseteq A$ has size λ, there is a bijection $h : \lambda \to X$, and so X can identified with a function $f : \lambda \to A$.

Conversely, let $f : \lambda \to A$ be given. Then $f \subseteq (\lambda \times A)$ and $|f| = \lambda$. It follows that f is in $[(\lambda \times A)]^\lambda$, which has the same size as $[A]^\lambda$ because $|A \times \lambda| = \max(|A|, \lambda) = |A|$, thus

$$|A|^\lambda \leq |[A \times \lambda]^\lambda| \leq |[A]^\lambda|. \tag{7.136}$$

□

For more information about κ^λ in general, see Section 7.8.

7.5 Infinite sums and products

Let λ be an infinite cardinal and let $\langle X_i \,|\, i < \lambda\rangle$ be a sequence of pairwise disjoint non-empty sets such that $|X_i| = \kappa_i$ for each $i < \lambda$. We define

$$\textstyle\sum_{i<\lambda} \kappa_i = |\bigcup\{X_i \,|\, i < \lambda\}|.$$

Lemma 7.19 *Let* λ *be an infinite cardinal and* $\kappa_i > 0$ *for each* $i < \lambda$*:*

$$\textstyle\sum_{i<\lambda}\kappa_i = \lambda \cdot \sup(\{\kappa_i \,|\, i < \lambda\}). \tag{7.137}$$

Proof. Denote $\bar{\kappa} = \sup(\{\kappa_i \,|\, i < \lambda\}$.

To prove the lemma, we will show that

(i) $\sum_{i<\lambda} \kappa_i \leq \lambda \cdot \bar{\kappa}$, and

(ii) $\sum_{i<\lambda} \kappa_i \geq \lambda \cdot \bar{\kappa}$.

By definition $\sum_{i<\lambda} \kappa_i = |\bigcup\{\{i\} \times \kappa_i \,|\, i < \lambda\}|$.

Ad (i). We will define a 1-1 function g from $\bigcup\{\{i\} \times \kappa_i \,|\, i < \lambda\}$ to $\lambda \times \bar{\kappa}$. Set $f(\langle i, \xi\rangle) = \langle i, \xi\rangle$.

Ad (ii). It suffices to show separately $\lambda \leq \sum_{i<\lambda} \kappa_i$ and $\bar{\kappa} \leq \sum_{i<\lambda} \kappa_i$ because by $\lambda \geq \omega$, $\lambda \cdot \bar{\kappa} = \max(\lambda, \bar{\kappa})$. For the first inequality, define g_1 so that $g_1(i) = \langle i, 0\rangle$ for every $i < \lambda$.

For the second inequality, define g_2 so that $g_2(\xi) = \langle i, \xi\rangle$, where i is the least j such that $\xi \in \kappa_j$. □

We can also define infinite products. Recall that if $\{X_i \mid i \in I\}$ is a family of non-empty sets, we define the product $\prod_{i\in I} X_i$ as follows:

$$\prod_{i\in I} X_i = \{f \mid f \text{ a function} : I \to \bigcup_{i\in I} X_i \text{ such that } (\forall i \in I) f(i) \in X_i\}. \tag{7.138}$$

If $\{\kappa_i \mid i \in I\}$ is a family of cardinal numbers, we define the infinite product:

$$\prod_{i\in I} \kappa_i = |\prod_{i\in I} X_i|, \tag{7.139}$$

where $\{X_i \mid i \in I\}$ 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).

Lemma 7.20 *Let λ be an infinite cardinal and $\langle \kappa_i \mid i < \lambda\rangle$ a non-decreasing sequence of cardinals such that for each i, $\kappa_i > 0$. Then:*

$$\prod_{i<\lambda} \kappa_i = (\sup(\{\kappa_i \mid i < \lambda\}))^\lambda. \tag{7.140}$$

Proof. Denote $\bar\kappa = \sup(\{\kappa_i \mid i < \lambda\})$. We need to show
(i) $\prod_{i<\lambda} \kappa_i \le \bar\kappa^\lambda$, and
(ii) $\prod_{i<\lambda} \kappa_i \ge \bar\kappa^\lambda$
(i) is obvious since $\bar\kappa \ge \kappa_i$ for every $i < \lambda$.

For (ii), we first prove the following general property of products: let X_i for $i \in I$ be a system of non-empty sets ($I \neq \emptyset$), and $\{I_j \mid j \in J\}$ be some partition of I. Then

$$\prod_{i\in I} X_i \approx \prod_{j\in J}(\prod_{i\in I_j} X_i). \tag{7.141}$$

Define g with domain $\prod_{i\in I} X_i$ as follows. Given $f \in \prod_{i\in I} X_i$, set $g(f)$ to be a function F with domain J such that for each $j \in J$, $F(j)$ is a function h with domain I_j defined by $h = f \restriction I_j$. It is easy to verify that g is a bijection from $\prod_{i\in I} X_i$ onto $\prod_{j\in J}(\prod_{i\in I_j} X_i)$, thus proving (7.141).

Let us now return to the proof of (ii). Since $\lambda \ge \omega$, there is a bijection $e : \lambda \times \lambda \to \lambda$. Define a partition P of λ as follows:

$$P = \{e''(\{j\} \times \lambda) \mid j < \lambda\}; \text{ let us write } P = \{E_j \mid j < \lambda\}. \tag{7.142}$$

P is a partition of λ into λ-many pieces, each of size λ. As we assume that the sequence of κ_i's is non-decreasing, and for each $j < \lambda$, E_j unbounded in λ (otherwise we would have a bijection between λ and its bounded segment), we have that

$$\text{for each } j < \lambda,\ \bar\kappa = \sup(\{\kappa_i \mid i \in E_j\}). \tag{7.143}$$

Because each $\kappa_i > 0$, this implies that for each $j < \lambda$,

$$\prod_{i\in E_j} \kappa_i \ge \bar\kappa. \tag{7.144}$$

When we put all these pieces together, we obtain:

$$\bar\kappa^\lambda \le \prod_{j<\lambda} \bar\kappa \le \prod_{j<\lambda}(\prod_{i\in E_j} \kappa_i) = \prod_{i<\lambda} \kappa_i. \tag{7.145}$$

□

7.6 König's theorem

Infinite sums and infinite products which we reviewed in Section 7.5 are connected by the following important theorem:

Theorem 7.21 (König) *If $0 < \kappa_i < \lambda_i$ for every $i \in I$ where I is non-empty, then*

$$\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i. \tag{7.146}$$

Proof. Fix a 1-1 function $e_i : \kappa_i \to \lambda_i \setminus \{0\}$ for each $i \in I$. This is possible because $\lambda_i > \kappa_i$.

We first show $\sum_{i \in I} \kappa_i \leq \prod_{i \in I} \lambda_i$. By definition, $\sum_{i \in I} \kappa_i = |\bigcup \{\{i\} \times \kappa_i \,|\, i \in I\}|$. We will therefore find 1-1 function g from $\bigcup \{\{i\} \times \kappa_i \,|\, i \in I\}$ to $\prod_{i \in I} \lambda_i$. Let $\langle i_0, \xi_0 \rangle$ be an arbitrary element of $\bigcup \{\{i\} \times \kappa_i \,|\, i \in I\}$. Define $g(\langle i_0, \xi_0 \rangle)$ to be a function $f \in \prod_{i \in I} \lambda_i$ such that $f(i_0) = e_{i_0}(\xi_0)$, and $f(j) = 0$ for $j \neq i_0$. Define g this way for every element of $\bigcup \{\{i\} \times \kappa_i \,|\, i \in I\}$. g is correctly defined and is 1-1.

In order to show the strict inequality $<$, we will assume for contradiction that there is a bijection h from $\bigcup \{\{i\} \times \kappa_i \,|\, i \in I\}$ to $\prod_{i \in I} \lambda_i$. By diagonalization, we will find $F \in \prod_{i \in I} \lambda_i$ not in the range h, thus showing that h is not a bijection after all. For a fixed $i \in I$, let us denote

$$H(i) = \{\zeta \in \lambda_i \,|\, \exists \xi \in \kappa_i (h(\langle i, \xi \rangle)(i) = \zeta)\}. \tag{7.147}$$

Notice that for every i, $H(i)$ is a proper subset of λ_i because $\kappa_i < \lambda_i$, and $|H(i)| \leq \kappa_i$. Let us define F as follows:

$$F(i) = \min(\lambda_i \setminus H(i)), \tag{7.148}$$

for every $i \in I$. F is an element of $\prod_{i \in I} \lambda_i$ and therefore by our assumption, there must be some $\langle i_0, \xi_0 \rangle$ such that $h(\langle i_0, \xi_0 \rangle) = F$. However, $F(i_0)$ is an element of $H(i_0)$, which contradicts the definition of F in (7.148). □

Corollary 7.22 *(i) $2^\kappa > \kappa$, so König's theorem implies Cantor's theorem.*
(ii) $\mathrm{cf}(2^\kappa) > \kappa$.
(iii) $\kappa^{\mathrm{cf}(\kappa)} > \kappa$.

Proof. Ad (i). Set $I = \kappa$, $\kappa_i = 1$, and $\lambda_i = 2$ for each $i < \kappa$. By König's lemma,

$$\kappa = \sum_{i < \kappa} 1 < \prod_{i < \kappa} 2 = 2^\kappa. \tag{7.149}$$

Ad (ii). Assume $\langle \kappa_i \,|\, i < \mu \rangle$ is cofinal in 2^κ, where μ is an infinite cardinal. Then

$$2^\kappa = \sum_{i < \mu} \kappa_i < \prod_{i < \mu} 2^\kappa = (2^\kappa)^\mu. \tag{7.150}$$

If $\mu \leq \kappa$, then

$$(2^\kappa)^\mu = |{}^{(\kappa \times \mu)}2| = 2^\kappa,$$

which contradicts the strict $<$ in (7.150) above. Hence $\mu > \kappa$, and so $\mathrm{cf}(2^\kappa) > \kappa$.

Ad (iii). Let $\langle \kappa_i \,|\, i < \mathrm{cf}(\kappa) \rangle$ be some cofinal subset of κ. Then

$$\kappa = \sum_{i < \mathrm{cf}(\kappa)} \kappa_i < \prod_{i < \mathrm{cf}(\kappa)} \kappa = \kappa^{\mathrm{cf}(\kappa)}. \tag{7.151}$$

□

Corollary 7.22(iii) for instance implies that whatever the size of the real numbers is, one cannot find a countable cofinal sequence in $2^\omega = |\mathbb{R}|$ because $\mathrm{cf}(2^\omega) > \omega$. See Section 7.7 for more information about the size of real numbers.

7.7 Continuum function

The function which to every cardinal $\aleph_\alpha$ assigns the cardinal $2^{\aleph_\alpha}$ is called the *continuum function*. The behaviour of this function was and is one of the central themes of set theory. We know from Cantor's theorem one thing:

$$\text{For every } \alpha, \aleph_\alpha < 2^{\aleph_\alpha}. \tag{7.152}$$

Can we say more?

In 1900 David Hilbert, a distinguished German mathematician, listed the problem "what is the cardinal $2^{\aleph_0}$" 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 $\aleph_0$:

$$\text{Continuum hypothesis, } \mathsf{CH}\text{: } 2^{\aleph_0} = \aleph_1. \tag{7.153}$$

This can be generalised to:

$$\text{Generalised continuum hypothesis, } \mathsf{GCH}\text{: } (\forall\alpha)2^{\aleph_\alpha} = \aleph_{\alpha+1}. \tag{7.154}$$

Using König's theorem (7.21), 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 7.23 *Continuum function satisfies for very $\alpha, \beta \in \mathrm{ORD}$:*
(i) $\alpha < \beta \to 2^{\aleph_\alpha} \leq 2^{\aleph_\beta}$*;*
(ii) (Cantor's theorem) $\aleph_\alpha < 2^{\aleph_\alpha}$*;*
(iii) (consequence of König's theorem) $\aleph_\alpha < \mathrm{cf}(2^{\aleph_\alpha})$*.*

Note for every β, $\mathrm{cf}(\aleph_\beta) \leq \aleph_\beta$, and in particular $\mathrm{cf}(2^{\aleph_\alpha}) \leq 2^{\aleph_\alpha}$, and so (iii) implies (ii) above. So in fact, only two properties of the continuum function are captured in Theorem (7.23): (i) and (iii).

For more than 30 years mathematicians tried to prove more about the continuum function than included in Theorem 7.23, but they failed. Only in early 30's, Kurt Gödel managed to prove that if ZF is consistent, so is ZF + AC + GCH.[16] However, this just showed that it is *possible* that GCH holds.

In early 60's, Paul Cohen managed to show[17] 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 7.23 are *the only properties* one can show about the continuum function in ZFC for *regular cardinals*[18]. For instance the following is consistent with ZFC:
(i) $2^{\aleph_0} = \aleph_2$;
(ii) $2^{\aleph_1} = \aleph_2$;
(iii) $2^{\aleph_2} = \aleph_{117}$;

[16] In fact ZF proves that GCH implies AC.

[17] 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.

[18] Situation for singular cardinals is more complex. For instance a celebrated result of Shelah is that the following holds in ZFC: if $2^{\aleph_n} < \aleph_\omega$ for every $n < \omega$, then $2^{\aleph_\omega} < \aleph_{\omega_4}$.

(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).

7.7.1 Fixed finite and infinite gaps

Easton's result which we referred to in the previous section says that the continuum function on regular cardinals can be very arbitrary. The value 2^κ for a *singular* cardinal κ cannot be manipulated so easily, but some results can be shown: For instance is is consistent that for any fixed $0 < n < \omega$, [19]

$$2^{\aleph_\alpha} = \aleph_{\alpha+n},$$

for all α.

However, ZFC does prohibit a fixed infinite gap: i.e. there is no $\beta \geq \omega$ such that for all α, $2^{\aleph_\alpha} = \aleph_{\alpha+\beta}$. We show this results and use this opportunity to introduce some more cardinal arithmetics in Lemmas 7.24 and 7.25.

Lemma 7.24 *Suppose κ is a limit cardinal, then*

$$(2^{<\kappa})^{\mathrm{cf}(\kappa)} = 2^\kappa.$$

Proof. It suffices to find an injective function from $(2^{<\kappa})^{\mathrm{cf}(\kappa)}$ to 2^κ and conversely. $2^\kappa \leq (2^{<\kappa})^{\mathrm{cf}(\kappa)}$ is shown by arguing that every subset of κ can be uniquely coded as a sequence of its initial segments indexed by a fixed cofinal sequence in κ of order-type $\mathrm{cf}(\kappa)$. The converse follows easily: $(2^{<\kappa})^{\mathrm{cf}(\kappa)} \leq (2^\kappa)^\kappa = 2^\kappa$. □

Lemma 7.25 *Suppose κ is a singular cardinal and the continuum function is eventually constant below κ, i.e. there is $\mu < \kappa$ such that for all $\mu \leq \mu' < \kappa$, $2^\mu = 2^{\mu'} = \lambda$ for some λ. Then $2^\kappa = \lambda$.*

Proof. We have: $2^\kappa = (2^{<\kappa})^{\mathrm{cf}(\kappa)} = (2^\mu)^{\mathrm{cf}(\kappa)} = 2^\mu$, if we take μ large enough (i.e. greater than $\mathrm{cf}(\kappa)$). □

Theorem 7.26 *There is no $\beta \geq \omega$ such that for all α,*

$$2^{\aleph_\alpha} = \aleph_{\alpha+\beta}.$$

Proof. Suppose for contradiction that some such β exists. Let α be least such that

$$\alpha + \beta > \beta.$$

[19] The consistency of this statement depends on the consistency of fairly big large cardinals.

Notice that α is a limit ordinal less or equal to β: Since β is infinite, it can be written as $\beta^* + n$ for some $n < \omega$ and limit β^*; if $\alpha = \alpha^* + 1$, then by associativity $\alpha^* + 1 + \beta^* + n = \alpha^* + \beta^* + n = \alpha^* + \beta$ contradicting the minimality of α.

Consider the cardinal $\kappa = \aleph_{\alpha+\alpha}$; clearly, its cofinality is equal to $\mathrm{cf}(\alpha)$ and hence κ is a singular cardinal. By Lemma 7.24 and 7.25,

$$2^\kappa = \aleph_{\alpha+\beta}$$

since the continuum function is eventually constant below κ with value $\aleph_{\alpha+\beta}$: for all $\xi < \alpha$, $\xi + \beta = \beta$, and therefore $2^{\aleph_{\alpha+\xi}} = \aleph_{\alpha+\xi+\beta} = \aleph_{\alpha+\beta}$. However, we also have that

$$2^\kappa = \aleph_{\alpha+\alpha+\beta}.$$

This is a contradiction since $\aleph_{\alpha+\alpha+\beta}$ is strictly bigger than $\aleph_{\alpha+\beta}$ because $\alpha+(\alpha+\beta) > \alpha+\beta$ as $\alpha + \beta > \beta$ by our assumption. □

7.8 Cardinal exponentiation under GCH

We show that under GCH, we have a complete answer to the question what is κ^λ for $\lambda \geq \omega$.

First notice the following simple consequence of GCH. For κ, λ as above let $\mu = \max(\kappa, \lambda)$. Then

$$\kappa^\lambda \leq \mu^\mu = 2^\mu = \mu^+. \tag{7.155}$$

Theorem 7.27 *Assume GCH and $\lambda \geq \omega$, $\kappa > 0$.*

(i) If $\kappa \leq \lambda$, then $\kappa^\lambda = \lambda^+$.

(ii) If $\lambda < \kappa$ and $\mathrm{cf}(\kappa) \leq \lambda$, then $\kappa^\lambda = \kappa^+$.

(iii) If $\lambda < \kappa$ and $\lambda < \mathrm{cf}(\kappa)$, then $\kappa^\lambda = \kappa$.

Proof. Ad (i). By Lemma 7.17, $\kappa^\lambda = 2^\lambda$, which under GCH is equal to λ^+.

Ad (ii). We know by (7.155) that $\kappa^\lambda \leq \kappa^+$. By a corollary to König's lemma, we also know $\kappa < \kappa^{\mathrm{cf}(\kappa)}$. Together we have:

$$\kappa < \kappa^{\mathrm{cf}(\kappa)} \leq \kappa^\lambda \leq \kappa^+, \tag{7.156}$$

which implies $\kappa^\lambda = \kappa^+$.

Ad (iii). Any function in ${}^\lambda\kappa$ has its range bounded in κ because $\mathrm{cf}(\kappa) > \lambda$. It follows

$$ {}^\lambda\kappa = \bigcup_{\gamma<\kappa} {}^\lambda\gamma, \text{ and so} \tag{7.157}$$

$$\kappa^\lambda = \sum_{\gamma<\kappa} |\gamma|^\lambda. \tag{7.158}$$

Since $\max(|\gamma|, \lambda) < \kappa$ for every $\gamma < \kappa$, we have by (7.155), that $|\gamma|^\lambda \leq \kappa$. So $\kappa^\lambda = \kappa$ by (7.158). □

8 Infinite combinatorics

We close the lecture by introducing two notions which are very useful in many fields: trees and closed unbounded sets.

8.1 Trees

8.1.1 Basic definitions

Definition 8.1 *We say that $(T,<)$ is a* tree *if $(T,<)$ is a partial order such that for each $t \in T$, the set $\{s \in T \,|\, s < t\}$ is well-ordered by $<$. Let*

$$\mathrm{ht}(t,T) = \mathrm{ot}(\{s \in T \,|\, s < t\}),$$

where "ot" denotes the order-type of a given well-ordered set. We define $T_\alpha = \{t \in T \,|\, \mathrm{ht}(t,T) = \alpha\}$. We set $\mathrm{height}(T)$ *to be the least α such that $T_\alpha = \emptyset$. We further set $T \restriction \alpha = \bigcup_{\beta<\alpha} T_\beta$ (which makes $T \restriction \alpha$ a subtree of T of height α).*

Note that we do not require that a tree has a single node of height 0 (*root*).[20]

Examples. Recall that $2^{<\omega}$ denotes the set of all finite sequences of 0's and 1's. We say that $T \subseteq 2^{<\omega}$ is closed under initial segments if $s \in T$ and $t \subseteq s$ implies $t \in T$. For every $T \subseteq 2^{<\omega}$ closed under initial segments, $(T, \subseteq)$ is a tree: $\emptyset$ is the root of the tree and for every $n < \mathrm{height}(T)$, T_n is equal to the set of all sequences $s \in T$ of length n. Note that T is infinite if and only if T has height ω. In general, if α is an infinite ordinal, then $T \subseteq 2^{<\alpha}$ closed under initial segments is a tree with the ordering $\subseteq$.

Definition 8.2 *For a regular cardinal $\kappa \geq \omega$, T is called a* κ-tree *if T has height κ, and $|T_\alpha| < \kappa$ for each $\alpha < \kappa$.*

Note that a tree $(T, \subseteq)$ with $T \subseteq 2^{<\omega}$, T infinite, is an ω-tree.

If T is a tree and $B \subseteq T$, we say that B is a *branch* if it is a maximal (under inclusion) chain in T. The following is a basic observation concerning ω-trees, due to König.

Theorem 8.3 (König) *Every ω-tree T has an infinite branch.*

Proof. We construct a branch B by induction on levels. Since T is an ω-tree, $|T_0| < \omega$. It follows there is some $t_0 \in T_0$ such that $S(t_0) = \{s \in T \,|\, t_0 < s\}$ is infinite (this is true because T_0 is finite, T is infinite, and every element in T is above an element of T_0). The set $S(t_0) \cap T_1$ is finite – pick $t_1 \in S(t_0) \cap T_1$ such that $S(t_1) = \{s \in T \,|\, t_1 < s\}$ is infinite. Proceed in the same fashion and pick t_n for each $n < \omega$. Then $B = \{t_n \,|\, n < \omega\}$ is a branch in T. □

Examples. Note that there are trees of height ω which have no infinite branches: for instance consider a tree $T = \{\{i\} \times i \,|\, i < \omega\}$ where the ordering $<$ on T is as follows: $(m,n) < (k,l)$ if and only if $m = k$ and $n < l$. T has height ω but every branch is finite; this does not contradict König's theorem because T is not an ω-tree: already the level 0 has ω-many nodes.

[20] Very often, more "well-behaved" trees are considered: such as with a root or with the property that above every node there is a node which splits ($t \in T$ splits if there are $s \neq s'$, s and s' immediate successors of t in T); since we are stating just the simple facts, we will not go into details here.

8.1.2 Aronszajn trees

Definition 8.4 *Let κ be a regular cardinal. We say that a κ-tree $(T, <)$ is an* Aronszajn tree *if it has no branch of size κ.*

Remark 8.5 An Aronszajn κ-tree T is a typical example of an "incompact object": by definition, for each $\alpha < \kappa$, there is a branch B_α of height α in T – if T were to be "compact" (in the analogous sense as first-order logic is compact), then from the assumption that for each $\alpha < \kappa$, there exists a branch of height α, we should be able to conclude that there is a branch of height κ.

By König's theorem, there is no Aronszajn tree on ω. Is there an Aronszajn tree on ω_1? Yes, there is, as we will show in Theorem 8.8. Before the theorem, we will state some more properties of trees which are useful.

Very often, a κ-tree T is isomorphic to a subtree of the full κ-ary tree $(\kappa^{<\kappa}, \subseteq)$. More precisely, whenever T is *normal* (indeed, normal here means representable as a subtree of $(\kappa^{<\kappa}, \subseteq)$). See Definition 8.6.

Definition 8.6 *A* normal κ-tree *is a tree T with the following properties:*

(i) $\mathrm{height}(T) = \kappa$*;*
(ii) $|T_0| = 1$*;*
(iii) $|T_\alpha| < \kappa$, for every $\alpha < \kappa$;
(iv) each node has ρ-many successors (exact number varies; $\rho < \kappa$);
(v) each $x \in T$ has some $y > x$ at each higher level of T;
(vi) if $\beta < \kappa$ is a limit ordinal, and $\mathrm{ht}(x, T) = \mathrm{ht}(y, T) = \beta$ and x, y have the same predecessors, then $x = y$.

Lemma 8.7 *Every normal tree T is isomorphic to a subtree $\bar{T}$ of the full κ-ary tree $({}^{<\kappa}\kappa, \subseteq)$, where $\bar{T}_\beta$ consists of sequences with domain β. In fact, only the items (i),(ii),(iii),(vi) of normality are required.*

Proof. We define by induction isomorphisms $i_\alpha : T \restriction \alpha \to \bar{T} \restriction \alpha$ and $i = \bigcup i_\alpha : T \to \bar{T}$. Set $\bar{T}_0 = \{\emptyset\}$; by (ii), $i_1(r) = \emptyset$ is an isomorphism between T_0 and $\bar{T}_0$, where r is the unique root of T. Suppose we have constructed $i_\beta : T \restriction \beta \to \bar{T} \restriction \beta$ for each $\beta < \alpha$ and we wish to construct i_α.

Assume first that α is limit. Set $i_\alpha = \bigcup_{\beta < \alpha} i_\beta$.

Suppose α is a successor of a limit cardinal: $\alpha = \alpha' + 1$ where α' is limit. Then define i_α by extending $i_{\alpha'}$ setting for each $x \in T_\alpha$

$$i_\alpha(x) = \{\langle \beta, i_{\alpha'}(y) \rangle \mid \beta < \alpha' \ \&\ y < x \ \&\ \mathrm{ht}(y, T) = \beta\}.$$

By (vi), i_α is 1-1. It is obviously also an isomorphism.

Assume now that α is a successor of a successor ordinal β. Since $|T_\beta| < \kappa$ by (iii), one can naturally extend i_β to i_α by including the level T_β using some 1-1 function from T_β into κ.

Set $\bar{T} = \bigcup\{\mathrm{rng}(i_\alpha) \mid \alpha < \kappa\}$. □

Theorem 8.8 *There is an Aronszajn tree T^*. We construct T^* as a subtree of $T = \{s \in {}^{<\omega_1}\omega \mid s \text{ is 1-1}\}$ with $\subseteq$ as the ordering. In particular, our tree will be normal according to Definition 8.6.*

Proof. Consider the subtree $T = \{s \in {}^{<\omega_1}\omega \mid s \text{ is 1-1}\}$ of the tree ${}^{<\omega_1}\omega$. T cannot have an ω_1-branch, because it would yield a 1-1 function from ω_1 to ω. However, T is not the required tree because it has uncountable levels, and so is not an ω_1-tree.

Let us define for s and t in ${}^{<\omega_1}\omega$ the following equivalence relation

$$s \approx t \leftrightarrow \text{dom}(s) = \text{dom}(t) \ \& \ \{\beta \in \text{dom}(s) \mid s(\beta) \neq t(\beta)\} \text{ is finite.}$$

We call a sequence $\langle s_\alpha \mid \alpha < \omega_1 \ \& \ \text{dom}(s_\alpha) = \alpha\rangle$ a *semi-branch* whenever
(i) $s_\alpha \restriction \beta \approx s_\beta$, for every $\beta \leq \alpha$.
(ii) $\omega \setminus \text{rng}(s_\alpha)$ is infinite for each $\alpha < \omega_1$.

A semi-branch satisfying (i) and (ii) makes it easy to define an Aronszajn tree:

$$T^* = \{s \in T \mid \exists \alpha \ s \approx s_\alpha\}.$$

It is immediate to verify that T^* is an Aronszajn tree, and a subtree of T.

To finish the prove of the theorem, it suffices to construct a semi-branch $\langle s_\alpha \mid \alpha < \omega_1\rangle$ satisfying (i) and (ii) above. The construction is by induction on $\alpha < \omega_1$. For $\alpha + 1$, define $s_{\alpha+1} = s_\alpha \cup \{\langle \alpha, n\rangle\}$, where n is any natural number in $\omega \setminus \text{rng}(s_\alpha)$ (this is possible by (ii)).

At a limit stage γ, first fix an increasing sequence $\langle \alpha_n \mid n < \omega\rangle$ with limit γ. Define $t \in T_\gamma$ as the union $t = \bigcup_n t_n$, where each t_n is in T_{α_n} and $t_n \approx s_{\alpha_n}$ (which implies $t_n \restriction \beta \approx s_\beta$ for each $\beta \leq \alpha_n$). The sequence $\langle t_n \mid n < \omega\rangle$ is defined by induction. First set $t_0 = s_{\alpha_0}$. To construct t_{n+1} when we have already constructed t_n, consider first t^*_{n+1} defined as $t_n \cup (s_{\alpha_{n+1}} \setminus s_{\alpha_n})$. The domain of t^*_{n+1} is equal to α_{n+1} and by the induction assumption on t_n and the properties of $\langle s_\alpha \mid \alpha < \gamma\rangle$,

$$t^*_{n+1} \approx s_{\alpha_{n+1}}. \tag{8.159}$$

However, while $s_{\alpha_{n+1}}$ is 1-1, t^*_{n+1} may not be 1-1 because of the finite disagreement (8.159). Define t_{n+1} by making finitely many changes to t^*_{n+1} to ensure:
(i) $t_{n+1} \approx t^*_{n+1}$.
(ii) t_{n+1} is 1-1.

This can be done because $\omega \setminus \text{rng}(s_{\alpha_{n+1}})$ is infinite, and so there is plenty of room to make t_{n+1} 1-1. It follows $t_{n+1} \in T_{\alpha_{n+1}}$.

Finally, the range of t may have used up all of ω, so we define s_γ by setting $s_\gamma(\alpha_n) = t(\alpha_{2n})$, thus leaving $t(\alpha_{2n+1})$'s outside the range of s_γ. Note that still $s_\gamma \restriction \beta \approx s_\beta$ for every $\beta < \gamma$, because α_n's are bounded below each $\beta < \gamma$, and so $s_\gamma \restriction \beta \approx t_n \restriction \beta \approx s_\beta$, for any n such that $\beta < \alpha_n$. □

What about Aronszajn trees at ω_2 or ω_3? Here things get more complicated: it is consistent that there are Aronszajn trees at ω_2 and ω_3 (for instance GCH implies this), but under some large-cardinal assumptions, it is also consistent that there are no Aronszajn trees at ω_2 and ω_3.

8.1.3 Souslin trees

The notion of an Aronszajn tree can be strengthened as follows:

Definition 8.9 *An Aronszajn tree T is called a* Souslin tree *if all antichains in T are at most countable.*

Note that the tree we have constructed in Theorem 8.8 is not Souslin:

Lemma 8.10 *Suppose that T is an Aronszajn tree and a subtree of the tree $\{s \in {}^{<\omega_1}\omega \mid s \text{ is 1-1}\}$. Then T is not Souslin.*

Proof. Notice that for each $n \in \omega$, $A_n = \{s \in T \mid \exists \alpha \operatorname{dom}(s) = \alpha + 1 \;\&\; s(\alpha) = n\}$ is an antichain. To see this, let $s \neq t$ be in A_n, and assume for contradiction that $s \subsetneq t$. Then $s(\alpha) = n$ and $t(\alpha) = t(\alpha') = n$, where $\operatorname{dom}(s) = \alpha + 1$ and $\operatorname{dom}(t) = \alpha' + 1$. This contradicts that t is 1-1. Now, many A_n's can be just countable, but by the pigeon hole principle, there must be some n such that A_n is uncountable (the set $A = \{s \in T \mid \exists \alpha \operatorname{dom}(s) = \alpha + 1\}$ is uncountable and $A = \bigcup_n A_n$). □

An ω_1-Souslin tree exists under some additional set-theoretical assumptions (such as $V = L$ or $\diamondsuit$).

8.2 Filter of closed unbounded sets

8.2.1 Closed and unbounded sets

Let κ be a regular uncountable cardinal.

To motivate the notional of closed unbounded set, consider the following example. Let $f : \kappa \to \kappa$ be a function. Let us say that $\alpha < \kappa$ is a *closure point* of f if for all $\beta < \alpha$, $f(\beta) < \alpha$. Let us denote $\mathrm{CL}(f)$ the set of closure points.

Claim 8.11 *(i) The set $\mathrm{CL}(f)$ is* unbounded *in κ, that is for every $\alpha < \kappa$ there is some β such that $\beta \in \mathrm{CL}(f)$ and $\alpha \leq \beta$.*

(ii) The set $\mathrm{CL}(f)$ is closed *in κ, that is if $\alpha < \kappa$ is a limit ordinal, and $\mathrm{CL}(f) \cap \alpha$ is unbounded, then $\alpha \in \mathrm{CL}(f)$.*

Proof. Ad (i). The proof is a special case of the Skolem hull argument for the construction of a substructure of κ which is closed under f and contains as a subset a given $\alpha \in \kappa$ (note that by transitivity of κ, $\alpha \subseteq \kappa$). Let $\alpha \in \kappa$ be given. By induction of length ω construct $\beta \supseteq \alpha$, $\beta \in \kappa$, which is closed under f. Set $\alpha = \alpha_0$, and if n is already constructed, $\alpha_{n+1} = \max(\alpha_n, \sup\{f(\gamma) + 1 \mid \gamma \in \alpha_n\})$. Set $\beta = \sup\{\alpha_n \mid n \in \omega\}$; by regularity of κ, $\beta \in \kappa$. It follows that $\beta \geq \alpha$ is a closure point of f.

Ad (ii). Trivial. □

The two properties of $\mathrm{CL}(f)$ identified above lead to the concept of a closed unbounded set. We say that $X \subseteq \kappa$ is *club* if it is unbounded and closed in κ.

Lemma 8.12 *If C and D are clubs in κ, then $C \cap D$ is a club in κ*

Proof. We first show that $C \cap D$ is closed. This is clear: if α is a limit ordinal and $C \cap \alpha$ and $D \cap \alpha$ are both unbounded in α, then by closedness of C, D, $\alpha \in C \cap D$.

The key of the proof is to show the unboundedness. Let $\gamma < \kappa$ be given, we wish to find some $\delta \geq \gamma$ such that $\delta \in C \cap D$. Let us construct by recursion a sequence $\langle c_i \,|\, i < \omega \rangle$ of elements of C and $\langle d_i \,|\, i < \omega \rangle$ of elements of D as follows. Choose $c_0 \in C$ and $d_0 \in D$ so that $\gamma < c_0 < d_0$. In general, in the step $n+1$, choose $c_{n+1} \in C$ and $d_{n+1} \in D$ so that $\ldots c_n < d_n < c_{n+1} < d_{n+1}$. Let us denote $\delta_1 = \sup\{c_i \,|\, i < \omega\}$ and $\delta_2 = \sup\{d_i \,|\, i < \omega\}$. First note that $\delta_1 = \delta_2$; let us denote this ordinal δ. Note that δ is a limit ordinal of countable cofinality. By closedness of C and D, $\delta \in C \cap D$. □

Lemma 8.13 *If $\{C_i \,|\, i < \mu\}$ is a set of clubs in κ for some $\mu < \kappa$, then $\bigcap_{i<\mu} C_i$ is a club in κ.*

Proof. This is a simple generalisation of Lemma 8.12, using the regularity of κ. Here are some details for the unboundedness: given $\gamma < \kappa$, construct by transfinite recursion increasing sequences $\langle c_i^n \,|\, n < \omega \rangle$ for every $i < \mu$ such that:

(i) $c_0^0 \geq \gamma$;

(ii) for a fixed $i < \mu$, c_i^n is in C_i for every $n < \omega$;

(iii) for a fixed $n < \omega$, $j < i < \mu$ implies that $c_j^n < c_i^n$;

(iv) for $m < n$ and every $i, j < \mu$, $c_i^m < c_j^n$.

By regularity of κ, $X = \bigcup_{n<\omega}\{c_i^n \,|\, i < \mu\}$ is bounded in κ. We claim that $\delta = \sup X$ is in $\bigcap_{i<\mu} C_i$ (and by definition is larger than γ). This follows from the fact that δ is the supremum of every sequence $\langle c_i^n \,|\, n < \omega \rangle$ for $i < \mu$. □

Exercise. Let C be a club and let $\mathrm{Lim}(C)$ be the set of limit points of C, where $\alpha \in C$ is a limit point of C if $C \cap \alpha$ is unbounded in α. Show that $\mathrm{Lim}(C)$ is a club (which is strictly smaller than C).

Lemma 8.12 allows us to define the *closed unbounded filter* generated by the club sets. Let us denote this filter as $\mathrm{Club}(\kappa)$:

$$\mathrm{Club}(\kappa) = \{X \subseteq \kappa \,|\, \text{there is a club } C \text{ such that } C \subseteq X\}.$$

We say that a filter F is *κ-complete* for a regular cardinal κ if for every family $\{X_i \,|\, i < \lambda\}$ of elements of F, where $\lambda < \kappa$, the intersection $\bigcap\{X_i \,|\, i < \lambda\}$ is in F.

Corollary 8.14 *The filter* $\mathrm{Club}(\kappa)$ *is κ-complete.*

Proof. Follows from Corollary 8.13. □

Note. Under AC, $\mathrm{Club}(\kappa)$ is never an ultrafilter (see in Theorem 8.18). The existence of an ω_1-complete non-principal ultrafilter on any regular κ is a very strong assumption which postulates the existence of the so called *measurable* cardinal.

8.2.2 Stationary and non-stationary sets

Let us denote as $\mathrm{NS}(\kappa)$ the dual ideal to $\mathrm{Club}(\kappa)$:

$$\mathrm{NS}(\kappa) = \{X \subseteq \kappa \,|\, \kappa \setminus X \in \mathrm{Club}(\kappa)\}.$$

We call the ideal $\mathrm{NS}(\kappa)$ the *non-stationary ideal* on κ. By Corollary 8.14, the non-stationary ideal $\mathrm{NS}(\kappa)$ is κ-complete.[21] We say that $X \subseteq \kappa$ is *stationary* if $X \notin \mathrm{NS}(\kappa)$.

Lemma 8.15 *$X \subseteq \kappa$ is stationary iff $X \cap C \neq \emptyset$ for every club C.*

Proof. If X is stationary, then $\kappa \setminus X$ is not in $\mathrm{Club}(\kappa)$. This means that there is no C so that $C \subseteq \kappa \setminus X$, or equivalently for any club C, $C \nsubseteq \kappa \setminus X$, which is the same as $C \cap X \neq \emptyset$.

For the converse, just run the argument in the opposite direction. □

The club filter $\mathrm{Club}(\kappa)$ satisfies another important property, that of *normality*. Let X_i for $i < \kappa$ be subsets of κ. Let us define the diagonal intersection

$$\triangle_{i<\kappa} X_i = \{\xi < \kappa \,|\, \xi \in \bigcap_{\zeta<\xi} X_\zeta\}.$$

Lemma 8.16 *The filter* $\mathrm{Club}(\kappa)$ *is* normal*, that is it is closed under the diagonal intersections of length κ: If for every $i < \kappa$, X_i is an element of* $\mathrm{Club}(\kappa)$*, then*

$$\triangle_{i<\kappa} X_i \in \mathrm{Club}(\kappa).$$

Proof. Let $\{C_i \,|\, i < \kappa\}$ be clubs such that $C_i \subseteq X_i$. It suffices to show that $D = \triangle_{i<\kappa} C_i$ is closed unbounded.

We first show that D is closed. Let α be a limit ordinal and $D \cap \alpha$ unbounded, we wish to show $\alpha \in D$. This is equivalent to demanding that for all $\beta < \alpha$, $\alpha \in C_\beta$. Fix such $\beta < \alpha$. Then for all γ, $\beta < \gamma < \alpha$, $\gamma \in D$ implies $\gamma \in C_\beta$; it follows $D \cap \alpha$ is unbounded in C_β, and hence $\alpha \in C_\beta$ as desired.

We now show that D is unbounded. Let $\alpha < \kappa$ be given, we wish to show there exists $\beta \geq \alpha$, $\beta \in D$. Set $\alpha_0 = \alpha$ and $A_0 = \emptyset$. Assume α_n and A_n are already constructed, we show how to construct α_{n+1} and A_{n+1}. Choose an increasing sequence $\langle a_\beta \,|\, \beta < \alpha_n\rangle$ such that $a_\beta \in C_\beta$ and $a_\beta > \alpha_n$ for each $\beta < \alpha_n$. Set $A_{n+1} = \{a_\beta \,|\, \beta < \alpha_n\}$ and $\alpha_{n+1} = \sup\, A_{n+1}$. Finally set $\beta = \sup\, \{\alpha_n \,|\, n < \omega\}$. In order to verify $\beta \in D$, we need to check that $\beta \in C_\gamma$ for each $\gamma < \beta$. Notice that for every $\gamma < \beta$ there exists $n < \omega$ such that $\gamma < \alpha_n$; it follows that for each $m \geq n$, there is some $a \in A_m \cap C_\gamma$. Hence $\beta \cap C_\gamma$ is unbounded and so $\beta \in C_\gamma$ as required. □

Note that in general, we cannot hope that any proper filter F on κ is κ^+-complete – for every such F there is a family X_α, $\alpha < \kappa$, of elements in F such that $\bigcap_{\alpha<\kappa} X_\alpha = \emptyset$. It follows that the diagonal intersection is in some sense the best we can get.

*Exercise**. Any normal filter F on κ is also κ-complete.

Intuitively, if set X is stationary, it means that it is not small in the sense of the club filter. "Stationarity" is therefore a measure of "largeness" for subsets of regular cardinals of uncountable cofinality. It has no analogue in case of ω, because ω has no limit points.

[21] Often, we say σ-complete instead of ω_1-complete.

Remark 8.17 The club filter $\mathrm{Club}(\kappa)$ properly extends the *Fréchet filter* $F(\kappa)$ on κ, where $X \in F(\kappa) \leftrightarrow X \setminus \kappa$ is bounded in κ. A typical subset of κ for which $F(\kappa)$ makes no decision, but $\mathrm{Club}(\kappa)$ does, is the set A of all limit ordinals in κ – A nor its complement $\kappa \setminus A$ are in $F(\kappa)$, but $A \in \mathrm{Club}(\kappa)$.

We said above that the club filter $\mathrm{Club}(\kappa)$ is not an ultrafilter. This follows from the following important theorem:

Theorem 8.18 (Solovay) *If κ is regular uncountable, then every stationary subset of κ is a disjoint union of κ-many stationary sets.*

This easily implies that $\mathrm{Club}(\kappa)$ is not an ultrafilter: if it were, then every stationary set must be in it, but this contradicts the above theorem, which claims that there are two (in fact κ-many) stationary sets which have empty intersection.

We end the discussion of stationary sets by stating a very useful Fodor's lemma.

We say that a function $f : \kappa \to \kappa$ is *regressive* if $f(\alpha) < \alpha$ for every $\alpha > 0$.

Theorem 8.19 (Fodor's lemma) *If $f : \kappa \to \kappa$ is regressive, then there is a stationary set $S \subseteq \kappa$ on which f is constant.*

More generally, if $f : T \to \kappa$ is regressive, where $T \subseteq \kappa$ is stationary, then there a stationary set $S \subseteq T$ on which f is constant.

Proof. We will just show the case for $f : \kappa \to \kappa$, although the generalisation to the second part featuring T is easy.

Assume for contradiction that for each $\alpha < \kappa$, the set $f^{-1}{}''\{\alpha\}$ is non-stationary, and fix for each α a club C_α such that

$$f^{-1}{}''\{\alpha\} \cap C_\alpha = \emptyset. \tag{8.160}$$

By diagonal intersection, the set $\triangle_\alpha C_\alpha$ is a club. However, any $\xi \in \triangle_\alpha C_\alpha$ contradicts the fact that f is regressive: $\xi \in \bigcap_{\zeta<\xi} C_\zeta$ implies by (8.160) that $\xi \notin f^{-1}{}''\{\zeta\} \leftrightarrow f(\xi) \neq \zeta$ for every $\zeta < \xi$. Thus $f(\xi) \geq \xi$, which contradicts the fact that f is regressive. □

*Exercise**. Show that Fodor's lemma implies Lemma 8.16.

If κ is regular and $\mu < \kappa$, then $E^\mu_\kappa = \{\alpha < \kappa \,|\, \mathrm{cf}(\alpha) = \mu\}$ is a very useful example of a stationary set.

References

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

9 Further reading

Recommended books in the order of relevance.

- Bohuslav Balcar, Petr Štěpánek, *Teorie množin*, Academia 2000.
- Kenneth Kunen, *Set Theory: An Introduction to Independence Proofs*, Elsevier 2004.
- Thomas Jech, *Set Theory*, Springer 2003.
- Akihiro Kanamori, *The Higher Infinite*, Springer 2003.
- Jean van Heijenoort, *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931 (Source Books in the History of the Sciences)*, Harvard University Press, 1879.