Set Theory – Lecture Notes

Version of March 19, 2024

Andreas Lietz

March 2024

These lecture notes are intended for the introductory Set Theory lecture at TU Wien in the summer semester of 2024. If you have any suggestions, remarks or find typos/errors, feel free to send me an email!

Contents

1 The Continuum Hypothesis **1**

2 Zermelo-Fraenkel Set Theory **6**
2.1 Extensionality 7
2.2 The empty set 7
2.3 Pairing 8
2.4 Union 8
2.5 Powerset 9
2.6 Infinity 9
2.7 Separation 9
2.8 Replacement 10
2.9 Foundation 11

3 Ordinals **12**
3.1 The structure of Ord 14
3.2 Induction and recursion 16
3.3 Ordinal arithmetic 20

4 Cardinals **23**
4.1 The structure of $(\mathrm{Card}, \leq)$ 24
4.2 The Axiom of Choice 27

1 The Continuum Hypothesis

5.3.24

The real number line is perhaps the best studied mathematical object there is. Set Theorists are particularly interested in the subsets of $\mathbb{R}$ and the first interesting thing to try is classifying sets of reals by their size. Of course we

can realize any finite size via the set $\{0, \ldots, n\}$ for $n \in \mathbb{N}$, as well as the size of $\mathbb{N}, \mathbb{R}$ themselves as obviously $\mathbb{N}, \mathbb{R} \subseteq \mathbb{R}$. The statement that this is a complete classification is known as the Continuum Hypothesis.

Definition 1.1 (Continuum Hypothesis)**.** The **Continuum Hypothesis** (CH) states that every infinite $X \subseteq \mathbb{R}$ is either countable, so in bijection with $\mathbb{N}$ or has the same size as $\mathbb{R}$, so is in bijection with $\mathbb{R}$.

Whether or not the continuum hypothesis is true was one of the most important mathematical questions of the 20th century, appearing as the first of the 23 questions posed by David Hilbert at the ICM in the year 1900.

The Austrian Kurt Gödel proved in the 30s that CH is at least not contradictory. It took another 30 years for Paul Cohen to show the dual statment: The negation of CH is not contradictory either, netting him a fields medal.

Proving Gödels result will be a central part of this lecture. Let us now begin with Cantor's early attempts at settling CH. His idea was to show that simple sets of reals cannot contradict CH and then push through to more and more complex sets of reals until finally CH is proven completely. While this project cannot be fully completed it was nonetheless a very fruitful strategy. Nowadays, Set Theorists have a good understanding of how complicated counterexamples to CH must be if they exist.

Theorem 1.2 (Cantor-Bendixson)**.** *Closed sets of reals are not counterexamples to* CH*, i.e. an uncountable closed set is in bijection with* $\mathbb{R}$*.*

We will show this by proving that any closed set of reals is the union of a perfect closed set P and a countable set A. Moreover, non-empty perfect closed sets are in bijection with $\mathbb{R}$.

Definition 1.3. A set $P \subseteq \mathbb{R}$ is **perfect** if for all $x \in P$, $x \in \overline{P \setminus \{x\}}$.

We will not try to give the most efficient proof, rather we want to illustrate some Set Theoretical ideas.

We will replace $\mathbb{R}$ by the interval $[0, 1]$ and represent closed sets $C \subseteq [0, 1]$ by binary trees. For 0-1-sequences $s, t \in \{0, 1\}^{\leq \mathbb{N}}$ write $s \leq t$ if s is an initial segment of t, i.e. if there is some $r \in \{0, 1\}^{< \mathbb{N}}$ so that $t = s^\frown r$.

Definition 1.4 (Binary Trees)**.** (*i*) A binary tree is a subset $T \subseteq \{0, 1\}^{< \mathbb{N}}$ of finite 0-1-sequences which is closed under initial segments, i.e. if $t \in T$ and $s \leq t$ then $s \in T$.

(*ii*) A branch through a binary tree T is a subset $b \subseteq T$ which is closed under initial segments and linearly ordered by $\leq$.

(*iii*) The set of cofinal branches through T is

$$[T] := \{b \subseteq T \mid b \text{ is an infinite branch}\}.$$

For $b \in [T]$, b^* is the unique infinite sequence in $\{0, 1\}^{\mathbb{N}}$ which all points in b are an initial segment of.

(*iv*) A binary tree T **represents** the set

$$[\![T]\!] := \{x \in [0,1] \mid \exists b \in [T]\ x = (0.b^*)_2\}$$

Here, $(0.a_1a_2a_3\dots)_2 = \sum_{n=1}^{\infty} a_1 \cdot 2^{-n}$ is the evaluation of a binary representation.

Proposition 1.5. *The following are equivalent for a set $D \subseteq [0,1]$:*

(*i*) *D is closed.*

(*ii*) *There is a binary tree T representing D, that is $D = [\![T]\!]$.*

Proof. $(i) \Rightarrow (ii)$: The set

$$T_D := \{t \in \{0,1\}^{<\mathbb{N}} \mid \exists b \in \{0,1\}^{\mathbb{N}}\ (0.b)_2 \in D \wedge t \leq b\}$$

is a binary tree with $[\![T_D]\!] = D$. "$\subseteq$" is obivous, while "$\supseteq$" holds as D is closed: If $x \in [\![T_D]\!]$ the there is $b \in [T_D]$ with $x = (0.b*)_2$. Find sequences $a_n \in \{0,1\}^{\mathbb{N}}$ with $(0.a_n)_2 \in D$ and $b \restriction n \leq a_n$ where

$$b \restriction n = b_1 \dots b_n$$

for $b^* = b_1b_2\dots$. It follows that $|(0.a_n)_2 - (0.a_m)_2| \leq 2^{-n}$ for $n \leq m$ so that

$$(0.b^*)_2 = \lim_{n\to\infty} (0.a_n)_2 \in D.$$

$(ii) \Rightarrow (i)$: We show that $[\![T]\!]$ is closed for all binary trees T. Suppose that $x_n \in [\![T]\!]$ for $n \in \mathbb{N}$ and $x_n \xrightarrow{n\to\infty} x$. As $x_n \in [\![T]\!]$, there is a sequence

$$a_1^n a_2^n \cdots \in \{0,1\}^{\mathbb{N}}$$

with all finite initial segments in T and $x_n = (0.a_1^n a_2^n \dots)_2$.

Claim 1.6. *There is a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ so that $(a_m^{n_k})_{k\in\mathbb{N}}$ is eventually constant for all $m \in \mathbb{N}$.*

Proof. Define sequences $(n_k^l)_{k\in\mathbb{N}}$ by induction on l. Let $n_k^0 = k$ for $k \in \mathbb{N}$ and now suppose that $(n_k^l)_{k\in\mathbb{N}}$ has been defined. $(a_l^{n_k^l})_{k\in\mathbb{N}}$ is a sequence which only takes one of two values, so we can then find a subsequence $(n_k^{l+1})_{k\in\mathbb{N}}$ on which it is constant.

Finally, the diagonal sequence $n_k = n_k^k$ does the job. □

(We have basically proven here that $\{0,1\}^{\mathbb{N}}$ is compact. The reader comfortable with this fact can ignore the claim above)

Let b_m be the eventual value of $((b_m)^{n_k})_{k\in\mathbb{N}}$. Then it is easy to see that

$$[\![T]\!] \ni (0.b_1b_2\dots)_2 = \lim_{k\to\infty} (0.a_1^{n_k} a_2^{n_k} \dots)_2 = \lim_{n\to\infty} x_n = x.$$

□

We can also describe perfect closed sets in terms of binary trees.

Definition 1.7. Suppose T is a binary tree.

(i) A node $t \in T$ **splits** if both $t^\frown 0$, $t^\frown 1$ are in T.

(ii) The tree T is **perfect** iff every $s \in T$ can be extended to some $s \leq t \in T$ which splits in T.

Proposition 1.8. *A closed set $D \subseteq [0,1]$ is perfect iff there is a perfect binary tree T representing D.*

Partial proof. We only show the easier direction as we have no use for the other implication anyway. Clearly $[\![\emptyset]\!] = \emptyset$ is perfect, so let T be a non-empty perfect tree and $x \in [\![T]\!]$, say $x = (0.a_1a_2\dots)_2$ and all finite initial segments of $a_1a_2\dots$ are in T. For each $k \in \mathbb{N}$, let a_{n_k} be the k-th splitting point along the branch b given by $a_1a_2\dots$, which must exist as T is perfect. Further, since T is perfect, we can extend $a_1\dots a_{n_k}^\frown(1 - a_{n_k+1})$ to an infinite branch b_k, so b and b_k differ first at their $n_k + 1$-th node. In particular,

$$|(0.b^*)_2 - (0.b_k^*)_2| \leq 2^{-k}$$

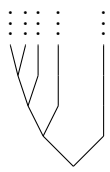
which shows $(0.b^*) = x \in \overline{[\![T]\!] \setminus \{x\}}$ □

Next we describe how we can reduce binary trees to perfect binary trees. The idea is to cut off isolated branches which do not split anymore.

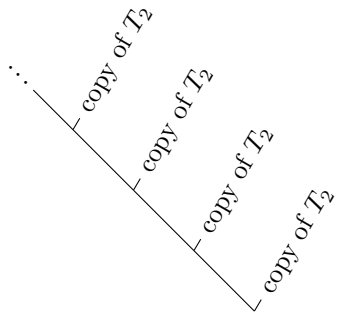
Definition 1.9. If T is a binary tree then the **derivative of** T is the binary tree

$$T' := \{t \in T \mid T \text{ splits above } t\}.$$

In some sense T' is closed to a perfect tree than T was. However T' certainly need not be perfect. Consider for example to following tree T_2:



Then T_2' is the leftmost branch of T_2 and not perfect. In fact $(T_2')' = \emptyset$. We can easily continue to produce a tree whose 3rd derivative is $\emptyset$, but not the 2nd, e.g. the tree T_3:



For a binary tree T, define inductively $T^{(0)} = T$ and $T^{(n+1)} = (T^{(n)})'$. So for every n there is a binary tree T with $T^{(n+1)} = \emptyset \neq T^{(n)}$. We set $T^\omega = \bigcap_{n<\omega} T_n$. It is still not guaranteed that T^ω is perfect. Does this mean we have to abandon ship and this construction is not helpful? No! We just have to continue this continue this construction transfinitely! To do so properly, we have to introduce ordinals. In the end we will have the following:

Lemma 1.10. *For every binary tree T, there is some countable ordinal α so that $T^{(\alpha)}$ is perfect.*

Note that a binary tree S is perfect iff $S' = S$, so the above happens only at the first α so that $T^{(\alpha+1)} = T^{(\alpha)}$.

Now, if $C \subseteq [0,1]$ is closed, let T_C be a binary tree representing C. Then let α be countable with $T_C^{(\alpha)}$ perfect. We set $P = [\![T_C^{(\alpha)}]\!]$, which is perfect, and $A = C \setminus P$. We have to show that A is countable.

Proposition 1.11. *If T is a binary tree then $[\![T]\!] \setminus [\![T']\!]$ is countable.*

Proof. We cut off at most countable many branches and each branch is responsible for the binary representation of at most one real number in $[\![T]\!]$ the branch does not split. □

Hence we can write

$$A = [\![T_C]\!] \setminus [\![T_C^{(\alpha)}]\!] = \bigcup_{\beta<\alpha} [\![T_C^{(\beta)}]\!] \setminus [\![T_C^{(\beta+1)}]\!]$$

which is a countable union of countable sets and hence countable.

To complete the proof of the Cantor-Bendixson Theorem, it remains to show that non-empty perfect closed sets are large.

Lemma 1.12. *If $P \subseteq [0,1]$ is nonempty and perfect closed then there is a bijection between P and $[0,1]$.*

We make use of a theorem we promise to prove at a later stage.

Theorem 1.13 (Cantor-Schröder-Bernstein)**.** *If there are injections $X \hookrightarrow Y$ and $Y \hookrightarrow X$ then there is a bijection between X and Y.*

Proof of Lemma 1.12. Let P be non-empty perfect closed. Clearly there is an injection $P \hookrightarrow \mathbb{R}$, e.g. the inclusion, so it remains to find an injection $[0,1] \hookrightarrow P$. Let T be a perfect tree representing P. We may arrange that every $x \in P$ is uniquely represented by a branch through T in the sense that if $b, c \in [T]$ are different then $(0.b^*)_2 \neq (0.c^*)_2$, the details are left to the reader. We first define an embedding $j\colon \{0,1\}^{<\mathbb{N}} \to T$ of the full binary tree into P by induction. We make sure that all nodes in $\mathrm{ran}(j)$ are splitting nodes of T. Map the empty sequence to the (unique) shortest splitting node of T (this exists as P is non-empty, so T is non-empty). Next, if $j(s)$ is defined, for $i = 0, 1$ let $j(s^\frown i)$ be the next splitting node of T above $j(s)^\frown i$. As j respects the initial segment relation $\leq$, j lifts to a map on the cofinal branches

$$j^+\colon [\{0,1\}^{\mathbb{N}}] \to [T]$$

via $j^+(b) = j[b]$, the pointwise image of b under j. As j is injective, so is j^+.

Putting everything together, we get an injection

$$[0,1] \hookrightarrow \{0,1\}^{\mathbb{N}} = [\{0,1\}^{<\mathbb{N}}] \xhookrightarrow{j^+} [T] \hookrightarrow [\![T]\!] = P$$

where the first arrow is choosing a binary representation and the last map is $b \mapsto (0.b^*)_2$. □

Tree constructions as above are immensely useful in Set Theory. When working with real numbers, the non-uniqueness of binary representation is sometimes somewhat annoying (as it is above as well). For that reason, the interval $[0,1]$ is usually replaced by the infinite binary sequences $\{0,1\}^{\mathbb{N}}$ and $\mathbb{R}$ is replaced by $\mathbb{N}^{\mathbb{N}}$. While the replacements are not homeomorphic to the originals, the differences are minor and can be neglected in almost all cases of interest.

2 Zermelo-Fraenkel Set Theory

6.3.24

So what is a set? Generally one can say that sets are collections x of other sets which are called the elements of x. If y is an element of x we write $y \in x$. Furthermore, two sets with the same elements are identical so a set is uniquely determined by its elements.

This is clearly not a satisfactory definition, among other problems, it is self-referential.

Cantor's original definition of a set reads:

"A set is a collection into a whole of definite distinct objects of our intuition or of our thought. The objects are called the elements (members) of the set."

However, it is impossible to give a correct naive definition of what a set is. Trying to do so leads to a host of paradoxes, the most prominent of which is **Russels's Paradox**: Let x be the set having as elements all the sets which are not elements of themselves, that is $y \in x$ iff $y \notin y$. The problem arises when one asks the question whether x is an element of itself. If $x \in x$, this means that

$x \notin x$. But if $x \notin x$ instead, we have to include x in x, so $x \in x$. Both scenarios end in contradiction!

Sometimes the only winning move is not to play. We will never give a definition of what a set is. We challenge the reader who is unsatisfied with this solution to give a rigorous definition of a natural number (without using sets, of course).

Instead, we formalize the properties that sets should have and define valid operations on sets which yield new sets. All of this will be collected in the theory ZF of **Zermelo-Fraenkel** Set Theory (we will add the axiom of choice at a later stage!). The Peano axioms do the same thing for natural number. The axioms of ZF are first order formulas in the language $\mathcal{L}_\epsilon$ consisting of a single binary relation $\in$. We also call first order formulas in the language $\mathcal{L}_\in$ **$\in$-formulas**.

2.1 Extensionality

Definition 2.1 (Extensionality)**.** The axiom of **extensionality** is

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

This axiom formalizes what we stated earlier: Sets are uniquely determined by their elements.

2.2 The empty set

Definition 2.2 (Empty)**.** The axiom of the **empty** set is

$$\exists x \ \forall z z \notin x.$$

This axiom is also known as Set Existence.

It will quickly get tedious to write all the axioms as bland $\in$-formulas. Instead we introduce syntactic sugar which makes our life a lot easier.

Definition 2.3. A **class term** is of the form

$$\{x \mid \varphi(x, v_0, \ldots, v_n)\}$$

for a variable x and a $\in$-formula φ with free variables among $x, v_0, \ldots, v_n$. We will often write only $\{x \mid \varphi\}$ instead.

So far a class term is only syntax without any inherent meaning. Nonetheless, we recommend to think of $\{x \mid \varphi\}$ as the collection of all sets x which satisfy φ. A **term** is either a variable or a class term.

Definition 2.4 (Class Term Sugar)**.** We introduce the following short hand notations:

- $y \in \{x \mid \varphi(x, v_0, \ldots, v_n)\}$ for $\varphi(y, v_0, \ldots, v_n)$.

- $y = \{x \mid \varphi\}$ for $\forall z\ z \in y \leftrightarrow z \in \{x \mid \varphi\}$.
- $\{x \mid \varphi\} \in y$ for $\exists z\ z = \{x \mid \varphi\} \wedge z \in y$.
- $\{x \mid \varphi\} = y$ for $y = \{x \mid \varphi\}$.

Definition 2.5. The term for the **empty set** is $\emptyset \coloneqq \{x \mid x \neq x\}$ and the term for the **universe of sets** is $V \coloneqq \{x \mid x = x\}$.

The empty set axiom can be formalized equivalently by $\exists x\ x = \emptyset$ or even simpler $\emptyset \in V$. These do not "desugar" to our original definition exactly, but they are trivially equivalent.

2.3 Pairing

For terms x, y the class term $\{x, y\}$ is defined as $\{z \mid z = x \vee z = y\}$.

Definition 2.6 (Pairing)**.** The **pairing** axiom is

$$\forall x \forall y\ \{x, y\} \in V.$$

More generally, for terms $x_0, \ldots, x_n$, we let

$$\{x_0, \ldots, x_n\} = \{z \mid z = x_0 \vee \cdots \vee z = x_n\}.$$

Note that from pairing and extensionality, we can prove the existence and uniqueness of the singleton $\{x\}$ for all x.

2.4 Union

Next up, we define the union axiom. We want to be able to build the union $x \cup y$ or even a union $\bigcup_{i \in I} x_i$ from a sequence $(x_i)_{i \in I}$. There is a simple convenient operation which allows for this without having to talk about sequences.

Definition 2.7 (Union)**.** For a term x, define the class term

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

The **union** axiom is

$$\forall x\ \bigcup x \in V.$$

While we are at it, we define several more useful class terms.

Definition 2.8. Let x, y be terms. We define the class terms

- $x \cup y \coloneqq \bigcup \{x, y\}$,
- $\bigcap x \coloneqq= \{z \mid \forall u (u \in x \to z \in u)\}$,
- $x \cap y \coloneqq \bigcap \{x, y\}$ and
- $x \setminus y = \{z \mid z \in x \wedge z \notin y\}$.

2.5 Powerset

For terms x, y we let $x \subseteq y$ be syntactic sugar for $\forall z\ (z \in x \to z \in y)$. We also let $\forall x \in y \varphi$ be sugar for $\forall x(x \in y \to \varphi)$, so $x \subseteq y$ can equivalently be defined as $\forall z \in x\ z \in y$. Similarly, $\exists x \in y \varphi$ is short for $\exists x\ (x \in y \wedge \varphi)$.

Definition 2.9 (Power)**.** For a term x, let $\mathcal{P}(x)$ be the class term $\{y \mid y \subseteq x\}$. The **power set** axiom is

$$\forall x\ \mathcal{P}(x) \in V.$$

2.6 Infinity

We want to express the existence of an infinite set. However, we do not currently have a working definition of what a finite set is. Instead, we demand the existence of a set which is closed under an appropriate operation.

Definition 2.10. For a term x, $x + 1$ is the class term $x \cup \{x\}$.

Note that we can prove $\forall x\ x + 1 \in V$ from the axioms we introduced so far, as well as $\forall x \forall y x + 1 = y + 1 \to x = y$ and $\forall x x + 1 \neq \emptyset$.

Definition 2.11. The axiom of **infinity** is

$$\exists x(\emptyset \in x \wedge \forall y \in x\ y + 1 \in x).$$

Intuitively, if x witnesses the axiom of infinity then the $+1$-operation induces an injective function from x to x which is not surjective as $\emptyset \in x$. Thus x could not be finite in any reasonable sense.

2.7 Separation

So far, all we only introduced finitely many axioms. Our axiomatization of ZF will not (and indeed cannot) be finite. Schemes are collections of formulas which are the result of transforming first order formulas in a uniform way.

Definition 2.12 (Separation)**.** For a $\in$-formula φ, the class term $\{x \in y \mid \varphi\}$ is defined as $\{x \mid x \in y \wedge \varphi\}$. The **separation scheme** consists of

$$\forall y\ \{x \in y \mid \varphi\} \in V$$

for all $\in$-formulas φ.

The reader may also know the operation of separating out elements according to a concrete criterium from any programming language implementing functional programming concepts as the `filter` command.

In most (but not all) proof-calculi the formula $\exists x\ x = x$ is a tautology. In this case, or just in presence of (Infinity), the (Empty) axiom can be derived from the separation scheme and (Extensionality) as from any x, we can separate out $\{y \in x \mid y \neq y\}$.

2.8 Replacement

Next, we introduce another scheme which is more powerful then the separation. We want that if $f : x \to y$ is a function between sets x, y then the range of f is a set. To formalize this, we first have to define what a function is, for which we have to formalize relations, for which we have to formalize the following:

Definition 2.13 (Kuratowski Pair)**.** The **ordered pair** (x, y) is the class term $\{\{x\}, \{x, y\}\}$.

Proposition 2.14. *From (Extensionality) and (Pairing), it follows that*

$$\forall x \forall y \forall x' \forall y' \ (x, y) = (x', y') \to (x = x' \wedge y = y').$$

Proof. Suppose that $(x, y) = (x', y')$. If $x = y$ then $(x, y) = \{\{x\}\}$ has only one element, so (x', y') also only has one element and it follows that $x' = y'$ and $(x', y') = \{\{x'\}\}$. Thus $\{\{x\}\} = \{\{x'\}\}$ and hence $\{x\} = \{x'\}$ so that $x = x'$. A symmetric argument works in case $x' = y'$.

So suppose $x \neq y$ and $x' \neq y'$. Then (x, y) has a unique element which is a singleton, namely $\{\{x\}\}$ and (x', y') has contains a unique singleton, namely $\{\{x'\}\}$. Hence we have $\{\{x\}\} = \{\{x'\}\}$, so $x = x'$.

Now the other elements of (x, y), (x', y') must agree as well, hence $\{x, y\} = \{x', y'\} = \{x, y'\}$. So $y \in \{x, y'\}$ and as $x \neq y$, we have $y = y'$. □

We leave the proof to the reader. There are many ways to achieve this effect, the above definition of (x, y) due to Kuratowski is simply the most common one. Ordered pairs are often taught as primitive notions in introductory math lectures, yet there is no need at all to do so. The encoding of an ordered pair as a set is our first example of emulating higher level mathematical concepts using sets.

Definition 2.15 (More Sugar)**.** For a class terms $\{x \mid \varphi(x, v_0, \ldots, v_n)\}$ and $\{y \mid \psi\}$, we set

$$\{\{x \mid \varphi\} \mid \psi\} = \{z \mid \exists v_0 \ldots \exists v_n \ z = \{x \mid \varphi(x, v_0, \ldots, v_n)\} \wedge z \in \{y \mid \psi\}\}.$$

Definition 2.16 (Relations)**.** A **(binary) relation** is a class term of the form

$$\{(x, y) \mid \varphi(x, y, v_0, \ldots, v_n)\}.$$

Suppose R is a binary relation.

(*i*) xRy is syntactic sugar for $(x, y) \in R$.

(*ii*) The **domain** of R is $\mathrm{dom}(R) = \{x \mid \exists y \ xRy\}$.

(*iii*) The **range** of R is $\mathrm{ran}(R) = \{y \mid \exists x \ xRy\}$.

Definition 2.17 (Functions)**.** Suppose F is a binary relation.

(*i*) F is a **function** if $\forall x \forall y \forall y' \ (xFy \wedge xFy') \to y = y'$.

(*ii*) For terms x, y, F is a **function from** x **to** y if F is a function, $\mathrm{dom}(F) = x$ and $\mathrm{ran}(F) \subseteq y$. We abbreviate this by $F \colon x \to y$.

(*iii*) The **value** of F at x is

$$F(x) \coloneqq \{z \mid \forall y \; xFy \land z \in y\}.$$

(*iv*) The **pointwise image** of x under F is[1]

$$F[x] = \{F(a) \mid a \in x\}.$$

Outside of Set Theory, there is often not notational distinction between the value $F(x)$ and pointwise image $F[x]$ and both are denoted by $F(x)$. This would be poor practice in Set Theory, as we will often deal with functions F and sets x so that both $x \in \mathrm{dom}(F)$ and $x \subseteq \mathrm{dom}(F)$. It would then be ambiguous whether we intend to take the value or pointwise image.

Definition 2.18 (Replacement)**.** The replacement scheme consists of

$$\text{“}F \text{ is a function”} \to \forall x \; F[x] \in V$$

for every binary relation F.

Note that we cannot define the replacement scheme by all formulas $\forall x \; F[x] \in V$ for all functions F. This would not make sense as “F is a formula” is a first order formula which does not have any truth associated to it. In contrast, saying F is a binary relation is simply a syntactic qualification of F.

Many programming languages implement replacement via the `map` command.

2.9 Foundation

So far, the axioms we have defined cannot rule out the existence of sets x which satisfy, e.g., $x = \{x\}$. Such a set would be quite unsettling, so it should not exist.

Definition 2.19 (Foundation)**.** The foundation scheme consists of the $\in$-formula

$$A \neq \emptyset \to \exists x \in A \; A \cap x = \emptyset$$

for any class term A.

One useful consequence of foundation is the non-existence of $\in$-cycles.

Proposition 2.20. *From the (Foundation) scheme it follows that*

$$\neg(\exists x_0 \ldots \exists x_n \; x_0 \in x_1 \land \cdots \land x_{n-1} \in x_n \land x_n \in x_0)$$

for any $n \in \mathbb{N}$.

[1] In other sources, $F\text{“}x$ is a common alternative notation for $F[x]$.

The natural numbers above are the usual (meta-theoretic) natural numbers. We have not yet defined natural numbers in terms of sets.

Proof. Suppose $x_0 \in x_1, \ldots, x_{n-1} \in x_n$ and $x_n \in x_0$. We apply (Foundation) to the class term $A = \{x_0, \ldots, x_n\}$. Let $y \in A$ so that $y \cap A = \emptyset$. We must have $y = x_i$ for some $i \leq n$. If $i = 0$ then $x_n \in x_i \cap A$ and if $i \neq 0$ then $x_{i-1} \in x_i \cap A$, contradiction. □

Intuitively, a similar argument shows that there are no infinite descending $\in$-chains $x_0 \ni x_1 \ni x_2 \ni \ldots$, however we cannot formalize this yet.

The axioms of the foundation scheme are maybe the least intuitive axioms of the lot. While this scheme is not provable from the other axioms, it does not add any consistency strength to the other axioms: Any model of the other axioms contains a "well-founded core" which is a model of all axioms/schemes defined so far, including (Foundation).

Definition 2.21 (ZF)**.** The of **Zermelo-Fraenkel** Set Theory, denoted ZF, is the collection of the axioms (Extensionality), (Empty), (Pairing), (Union), (Power), (Infinity) as well as the schemes (Separation), (Replacement) and (Foundation).

This is not a minimal representation of ZF: as we observed earlier, (Empty) is provable from the other axioms. Furthermore, the whole (Separation) scheme can be proven from the other axioms.

Nonetheless, this is the most prominent presentation of ZF for a number of reasons. On one hand, it is convenient as (Separation) is an important concept in any case, but it also has to do with the historical context. Zermelo first introduced his theory of Zermelo Set Theory, which did not include the (Replacement) and (Foundation) schemes. Later, Fraenkel observed the importance of these schemes which where widely used implicitly anyways. This is how ZF was born.

From now on, we will work in ZF without further mention.

Remark 2.22. We will mostly drop the word *term*, class terms will simply be called classes. We will call a term x a set if $x \in V$.

3 Ordinals

Ordinals are the backbone of the mathematical universe. They extend the natural numbers to a much much (much!) longer linear order along which induction and recursive definitions still work.

Definition 3.1. Suppose x is a set or class.

(*i*) x is **transitive** if whenever $z \in y \in x$ then $z \in x$. Equivalently, x is transitive if $\bigcup x \subseteq x$.

(*ii*) If x is a set then x is an **ordinal** if x is transitive and x is strictly linearly ordered by $\in$.

(*iii*) Ord is the class $\{x \mid x \text{ is an ordinal}\}$.

Examples 3.2
- $\emptyset$ *is trivially an ordinal. We set* $0 \coloneqq \emptyset$.
- $\{\emptyset\} = 0 + 1$ *is an ordinal and we denote it by* 1.
- $\{\{\emptyset\}\}$ *is not transitive, but it is linearly ordered by* $\in$.
- $\{\emptyset, \{\emptyset\}\} = 1 + 1$ *is an ordinal which we will denote by* 2.
- $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ *is transitive, but not linearly ordered by* $\in$.

As a convention, ordinals are usually denoted by lowercase Greek letters $\alpha, \beta, \gamma, \dots$.

Lemma 3.3. *The class* Ord *is*

(*i*) *transitive and*

(*ii*) *strictly linearly ordered by* $\in$.

Proof. (*i*) : Suppose $\beta \in \alpha \in \text{Ord}$, we have to show that β is an ordinal.

Claim 3.4. β *is transitive.*

Proof. Suppose $\delta \in \gamma \in \beta$. By transitivity of α, $\gamma \in \beta \in \alpha$ implies $\gamma \in \alpha$ and now $\delta \in \gamma \in \alpha$ implies $\delta \in \alpha$ as well. Since α is strictly linearly ordered by $\in$, we have either the good case $\delta \in \beta$ or one of the bad cases $\delta = \beta$, $\beta \in \delta$.

However, both bad cases lead to $\in$-cycles: If $\beta = \delta$ then $\delta \in \gamma \in \delta$ and if $\beta \in \delta$ then $\delta \in \gamma \in \beta \in \delta$. This is impossible by Proposition 2.20. □

It is left to show that β is linearly ordered, but this is straightforward as this is true for α and $\beta \subseteq \alpha$ by transitivity of α.
(*ii*) Exercise! □

12.3.24

We have just proven that the class Ord satisfies all the requirements of being in Ord.

Corollary 3.5. $\text{Ord} \notin V$.

Proof. Suppose $\text{Ord} \in V$. Then $\text{Ord} \in \text{Ord}$ which contradicts Proposition 2.20. □

This is known as the **Burali-Forti Paradoxon**. It seems that this was the first time a class was proven to not be a set.

Definition 3.6. We say that a class A is a **proper class** if $A \notin V$. Otherwise A is a **set**.

Russel's paradoxon can be resolved by noting that V, as Ord is a proper class and not a set.

3.1 The structure of Ord

We already know that Ord is strictly linearly ordered by $\in$. Since this order is important, we reserve a symbol for it. But first, we introduce the Cartesian product.

Definition 3.7. For A, B, the **Cartesian product** of A and B is

$$A \times B \coloneqq \{(a, b) \mid a \in A \wedge b \in B\}.$$

We also define that square $A^2 = A \times A$.

Proposition 3.8. *For sets a, b, we have $a \times b \in V$.*

Proof. Exercise. □

Until now, the $\in$-relation is a logical symbol representing a binary relation, so it is pure syntax. It is often useful to interpret it as a class as well: We set $\in = \{(x, y) \mid x \in y\}$. So from now on, $\in$ will be overloaded with two different meanings. We trust the reader to figure out which one we mean.

Definition 3.9. We set $<=\in \cap \mathrm{Ord}^2$, so $(\mathrm{Ord}, <)$ is a strict linear order. We denote the corresponding (non-strict) linear order by $\leq$.

The linear order $(\mathrm{Ord}, <)$ has a further important property, namely it is wellfounded.

Definition 3.10. A linear order R is **wellfounded** iff for all non-empty sets $x \subseteq \mathrm{dom}(R)$, we have that x contains a R-minimal element. More precisely, $\exists y \in x \forall z \in x\ \neg zRy$.

A (strict) **wellorder** is a wellfounded (strict) linear order.

Note that $(\mathrm{Ord}, <)$ is a wellorder and $(V, \in)$ is wellfounded by (Foundation). Wellfoundedness (but not quite sufficient) property for inductive proofs and recursive constructions. We will get to that soon.

Lemma 3.11. *For all ordinals α, $\alpha + 1 \in \mathrm{Ord}$ and $\alpha + 1$ is the immediate successor of α in $(\mathrm{Ord}, <)$. This means that for all β, if $\beta < \alpha + 1$ then $\beta \leq \alpha$.*

Proof. Exercise. □

The following fact is easy to verify.

Proposition 3.12. *If X is a set of transitive sets then $\bigcup X$ is transitive.*

Next, we describe infima and suprema of ordinals. Note that if α, β are ordinals then $\min\{\alpha, \beta\} = \alpha \cap \beta$ and $\max\{\alpha, \beta\} = \alpha \cup \beta$.

Lemma 3.13. *Suppose X is a non-empty set of ordinals. Then*

$$\bigcup X = \sup X$$

and

$$\bigcap X = \inf X = \min X.$$

In particular, $\bigcup X, \bigcap X$ are ordinals.

Proof. Showing $\bigcap X = \inf X = \min X$ is easier: let $\alpha = \min X$ which we know exists by wellfoundedness of $\in$. Then $\alpha \subseteq \beta$ for all $\beta \in X$, so $\alpha = \bigcap X$.
Now let us first show that $\bigcup X \in \mathrm{Ord}$. First, $\bigcup X$ is transitive by Proposition 3.12. Next, as Ord is transitive, $\bigcup X \subseteq \mathrm{Ord}$ and is hence linearly ordered by $\in$ since Ord is. □

Not all ordinals are of the form $\alpha + 1$. Surely, 0 is not, but there are more interesting ordinals with this property. We now take a look at the smallest one.

Definition 3.14. We say that x is **inductive** if $0 \in x$ and $\forall y \in x\ y + 1 \in x$.

The axiom (Infinity) simply states that there is an inductive set.

Definition 3.15. We define $\omega = \bigcap \{x \mid x \text{ is inductive}\}$.

Lemma 3.16. *The ω is an inductive set.*

Proof. It is straightforward to see that ω is inductive. To show that $\omega \in V$, let x be an arbitrary inductive set by (Infinity). We then have

$$\omega = x \cap \omega = \{y \in x \mid y \in \omega\} \in V.$$

Here, the last class is guaranteed to be a set by (Separation). □

We will next show that ω is an ordinal. For this we need to know that proper classes are larger than sets.

Proposition 3.17. *If C is a proper class and x is a set then $C \setminus x$ is a proper class.*

Proof. If not then $C = (C \setminus x \cup x)$ is a union of two sets, so C is a set by (Pairing) and (Union), contradiction. □

Lemma 3.18. $\omega \in \mathrm{Ord}$.

Proof. Let $\alpha = \min(\mathrm{Ord} \setminus \omega)$. This minimum exists as $(\mathrm{Ord}, <)$ is a wellorder and since $\mathrm{Ord} \setminus \omega \neq \emptyset$ by Proposition 3.17. We are done if we can show that $\alpha = \omega$. By (Extensionality), it suffices to show both $\alpha \subseteq \omega$ and $\omega \subseteq \alpha$.
"$\alpha \subseteq \omega$" : This is trivial as $\alpha \subseteq \mathrm{Ord}$ by transitivity of Ord and the choice of α.
"$\omega \subseteq \alpha$" : It suffices to show that α is inductive as ω is the smallest inductive set. First, $0 \leq \alpha$ and since $0 \in \omega$, $0 \neq \alpha$, hence $0 \in \alpha$. Second, if $\beta \in \alpha$ then $\beta + 1$ is the immediate successor of β, hence $\beta + 1 \leq \alpha$. But $\beta + 1 \in \omega$ since ω is inductive and $\beta \in \omega$ so that $\beta + 1 \neq \alpha$. □

As ω is inductive, ω is not of the form $\alpha + 1$ for any ordinal (or set) α.

Definition 3.19. The class of **successor ordinals** is

$$\mathrm{Succ} \coloneqq \{\alpha \in \mathrm{Ord} \mid \exists \beta\ \alpha = \beta + 1\}.$$

The class of **limit ordinals** is[2]

$$\mathrm{Lim} \coloneqq \mathrm{Ord} \setminus (\mathrm{Succ} \cup \{0\}).$$

[2]Sometimes, 0 is included in Lim.

Both Succ and Lim are proper classes, as we will seen soon.

Remark 3.20. If we put the topology given by $<$ (i.e. basic open sets are open intervals in $<$) then an ordinal α is a limit ordinal iff $\alpha \in \overline{\mathrm{Ord} \setminus \{\alpha\}}$. This fails for 0, so clearly 0 is not and should not be considered a limit ordinal.

We have that $\omega = \min \mathrm{Lim}$. We know that $\mathrm{Lim} \neq \emptyset$ as $\omega \in \mathrm{Lim}$, so $\min \mathrm{Lim}$ exists and is easily seen to be inductive hence it must be ω.

3.2 Induction and recursion

We know prove that several inductions work as intended.

Lemma 3.21 (Induction along ω)**.** *Suppose $A \subseteq \omega$ so that $0 \in A$ and $\forall n \in A\ n+1 \in A$. Then $A = \omega$*

Proof. This is trivial as this we assume A is inductive. □

Much more interestingly, we can reason inductively along all ordinals. This is known as **transfinite induction**.

Lemma 3.22 (Induction along Ord, version 1)**.** *Suppose $A \subseteq \mathrm{Ord}$ and $\forall \alpha \in \mathrm{Ord}\ \alpha \subseteq A \to \alpha \in \mathrm{Ord}$. Then $A = \mathrm{Ord}$.*

Proof. Suppose $A \neq \mathrm{Ord}$. Then let $\alpha \in \mathrm{Ord} \setminus A$ be $\in$-minimal by (Foundation). But then $\alpha \subseteq A$ which implies $\alpha \in A$ by assumption on A, contradiction. □

Basically the same argument shows:

Lemma 3.23 (Induction along V)**.** *Suppose $A \subseteq V$ and $\forall x\ x \subseteq A \to x \in A$. Then $A = V$.*

In practice, transfinite inductions along ordinals often split into a successor case and limit case. Because of this, it is useful to formulate a second version of transfinite induction.

Lemma 3.24 (Induction along Ord, version 2)**.** *Suppose $A \subseteq \mathrm{Ord}$ satisfies*

(i) $0 \in A$,

(ii) $\forall \alpha \in A\ \alpha + 1 \in A$ *and*

(iii) $\forall \alpha \in \mathrm{Lim}(\forall \beta < \alpha \beta \in A \to \alpha \in A)$.

Then $A = \mathrm{Ord}$.

Proof. By the first version, it suffices to show $\alpha \subseteq \mathrm{Ord} \to \alpha \in \mathrm{Ord}$ for all ordinals α. This is trivial if $\alpha = 0$. If $\alpha = \beta + 1$ then $\alpha \subseteq A$ implies $\beta \in A$ so $\alpha = \beta + 1 \in A$. Finally, if $\alpha \in \mathrm{Lim}$ and $\alpha \subseteq A$ then clearly $\forall \beta < \alpha\ \beta \in A$ so $\alpha \in A$. □

Now we get to recursive constructions.

Definition 3.25. Suppose F is a function. For any x, the **restriction of F to x** is $F \restriction x \coloneqq F \cap (x \times V)$.

We will make use of the following intuitively true fact.

Proposition 3.26. *If F is a function and x is a set then $F \restriction x$ is a set.*

Proof. Exercise. □

Theorem 3.27 (The Recursion Theorem)**.** *For any function $F\colon V \to V$, there is a function $G\colon V \to V$ which is defined by recursion along F, that is*

$$\forall x\ G(x) = F(G \restriction x).$$

Remark 3.28. We take some time to explain how to understand this theorem more precisely. Usually, if we prove a theorem/lemma/etc, we show that $\mathrm{ZF} \vdash \varphi$ for some single sentence φ. The Recursion Theorem is a "Meta Theorem" which means that we prove many theorems at once which are parametrized in some way. This parametrization is somewhat hidden in the Recursion Theorem: it really says that for any $\in$-formula φ, we can uniformly turn φ into another $\in$-formula ψ (read: we can write a computer program which does it) and we prove

$$\mathrm{ZF} \vdash (F\colon V \to V) \to [(G\colon V \to V) \wedge (\forall x\ G(x) = F(G \restriction x))]$$

where $F = \{(x, y) \mid \varphi\}$ and $G = \{(x, y) \mid \psi\}$.

Proof. The strategy of our proof will be to approximate G by smaller set-sized functions. Let us say that a function $g\colon a \to b$ is **F-recursive** if

- a is a transitive set and
- for all $x \in a$, $g(x) = F(g \restriction x)$ (note that $x \subseteq \mathrm{dom}(g)$.

We will show that $G \coloneqq \bigcup\{g \mid g \text{ is } F\text{-recursive}\}$ works. To do so, we have to prove that G is a function and that $\mathrm{dom}(G) = V$.

Claim 3.29. *If g, g' are F-recursive and $x \in \mathrm{dom}(g) \cap \mathrm{dom}(g')$ then $g(x) = g'(x)$.*

Proof. Suppose not. Let $x \in \mathrm{dom}(g) \cap \mathrm{dom}(g')$ be $\in$-minimal with $g(x) \neq g'(x)$. But then by choice of x we have

$$g(x) = F(g \restriction x) = F(g' \restriction x) = g'(x),$$

contradiction. □

With a moment of reflection, one concludes that G is indeed a function. We are done if we prove:

Claim 3.30. $\mathrm{dom}(G) = V$.

Proof. By induction along V, it suffices to show $x \subseteq \mathrm{dom}(G)$ implies $x \in \mathrm{dom}(G)$. So if $x \subseteq \mathrm{dom}(G)$, we know that if $y \in x$ then there is a F-recursive function g with $y \in \mathrm{dom}(g)$.

(We did not define the axiom of choice yet, but if we would assume it, it would guarantee the existence of a function mapping $y \in x$ to such a g, we will make do without the axiom of choice by describing an explicit such g for any $y \in x$.)

For $y \in x$, let

$$g_y \coloneqq \bigcap \{g \mid g \text{ is } F\text{-recursive with } y \in \mathrm{dom}(g)\}.$$

Using the agreement of two F-recursive functions on their common domain, it is easy to show that g_y is F-recursive with $y \in \mathrm{dom}(g_y)$. Hence the class

$$H \coloneqq \{(y, g_y) \mid y \in x\}$$

is a well-defined function and by (Replacement),

$$g' \coloneqq \bigcup H[x] \in V.$$

It is once again easy to see that g' is F-recursive. Finally, the function

$$g \coloneqq \{(x, F(g'))\}$$

witnesses $x \in \mathrm{dom}(G)$. □

□

Remark 3.31. The resulting recursion G along F is unique in the sense that whenever $G' \colon V \to V$ also satisfies $\forall x G'(x) = F(G' \restriction x)$ then $G = G'$ in the "syntax sugar" sense, equivalently $\forall x\ G(x) = G'(x)$. However, the exact syntactic class term G is not unique!

13.3.24

As for induction, it is convenient to formulate a variant of recursion along the ordinals.

Corollary 3.32 (Recursion along Ord)**.** *Suppose $F_0 \in V$ and $F_{\mathrm{Succ}}, F_{\mathrm{Lim}} \colon V \to V$ are functions. Then there is a function $G \colon \mathrm{Ord} \to V$ such that*

(*i*) $G(0) = F_0$,

(*ii*) $G(\alpha + 1) = F(G(\alpha))$ *and*

(*iii*) $G(\alpha) = F(G \restriction \alpha)$ *for limit ordinals α.*

Proof. Apply the Recursion Theorem 3.27 to the function F defined by

$$F(x) = \begin{cases} F_0 & \text{if } x = \emptyset \\ F_{\mathrm{Succ}}(x(\max \mathrm{dom}(x)) & \text{if } x \text{ is a function with } \mathrm{dom}(x) \in \mathrm{Succ} \\ F_{\mathrm{Lim}}(x) & \text{if } x \text{ is a function with } \mathrm{dom}(x) \in \mathrm{Lim} \\ \emptyset & \text{else.} \end{cases}$$

□

We will now apply the recursion theorem and make some important definitions. If $F\colon X \to V$ is a function then $\bigcup_{x\in X} F(x)$ is shorthand for $\bigcup\{F(x) \mid x \in X\}$.

Definition 3.33. The **Von-Neumann rank inital segments** are defined by

- $V_0 = \emptyset$
- $V_{\alpha+1} = \mathcal{P}(V_\alpha)$ and
- $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$ for $\alpha \in \mathrm{Lim}$.

Remark 3.34. To make this definition precise, we hand some input to the recursion theorem in the background: we let $F_0 = \emptyset$, let F_{Succ} be the powerset operation $x \mapsto \mathcal{P}(x)$ and define F_{Lim} via

$$F_{\mathrm{Lim}}(x) = \begin{cases} \bigcup \mathrm{ran}(x) & \text{if } x \text{ is a function} \\ \emptyset & \text{else.} \end{cases}$$

We get back a function G and set $V_\alpha = G(\alpha)$ for an ordinal α. In the future, we will hide such details.

Lemma 3.35. *Suppose α, β are ordinals.*

(*i*) *V_α is transitive.*

(*ii*) *If $\alpha \leq \beta$ then $V_\alpha \subseteq V_\beta$.*

(*iii*) *If $\alpha < \beta$ then $V_\alpha \in V_\beta$.*

(*iv*) *$V = \bigcup_{\alpha\in\mathrm{Ord}} V_\alpha$.*

Proof. (*i*): We prove this by induction on α.

$\alpha = 0$: is trivial.

$\alpha = \beta + 1$: Suppose $y \in x \in V_\alpha = \mathcal{P}(V_\beta)$. Then $y \in x \subseteq V_\beta$, so $y \in V_\beta$. By induction, V_β is transitive so $y \subseteq V_\beta$ an hence $y \in V_\alpha$.

$\alpha \in \mathrm{Lim}$: $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$ is transitive by induction and Proposition 3.12.

(*ii*): By induction on β.

$\beta = \alpha$: trivial.

$\beta = \gamma + 1$: We have $V_\alpha \subseteq V_\gamma$ and hence $V_\alpha \in \mathcal{P}(V_\gamma) = V_\beta$. As V_β is transitive by (*i*), $V_\alpha \subseteq V_\beta$.

$\beta \in \mathrm{Lim}$: trivial.

(*iii*): Clearly $V_\alpha \in \mathcal{P}(V_\alpha) = V_{\alpha+1}$. If $\alpha < \beta$ then $\alpha + 1 \leq \beta$ so that by (*ii*), $V_{\alpha+1} \subseteq V_\beta$ and hence $V_\alpha \in V_\beta$.
(*iv*): We show $\bigcup_{\alpha \in \mathrm{Ord}} V_\alpha = V$ by induction. Suppose x is a set and $x \subseteq \bigcup_{\alpha \in \mathrm{Ord}} V_\alpha$. Define a function $F \colon x \to \mathrm{Ord}$ by

$$F(y) = \min\{\alpha \in \mathrm{Ord} \mid y \in V_\alpha\}.$$

By (Replacement), $F[X]$ is a set and let $\delta = \sup F[X]$. Then for all $y \in x$ there is some $\gamma \leq \alpha$ with $y \in V_\gamma$ so that $y \in V_\delta$ by (*ii*). It follows that $x \subseteq V_\delta$ and consequently $x \in V_{\delta+1}$. □

Part (*iv*) of the Lemma above motivates the following definition.

Definition 3.36. The **rank of a set** x is

$$\mathrm{rk}(x) = \min\{\alpha \in \mathrm{Ord} \mid x \in V_{\alpha+1}\}$$

For example, $\mathrm{rk}(V_\alpha) = \alpha$: by (*iii*) above, $\mathrm{rk}(V_\alpha) \leq \alpha$. But if $\beta \leq \alpha$ then $V_\alpha \notin V_\beta$ as otherwise $V_\alpha \in V_\alpha$ by (*ii*) above. An induction shows that $V_\alpha \cap \mathrm{Ord} = \alpha$ so that $\mathrm{rk}(\alpha) = \alpha$ for all ordinals α.

3.3 Ordinal arithmetic

Ordinals admit natural addition, multiplication and exponentiation operations which restrict to the "usual ones" on ω. We define them via the recursion theorem.

Definition 3.37. For an ordinal α, we define $\alpha + \beta,\ \alpha \cdot \beta,\ \alpha^\beta$ for all ordinals β by recursion. Ordinal addition is defined via:

$\alpha + 0 = \alpha$,

$\alpha + \beta = (\alpha + \beta) + 1$ and

$\alpha + \beta = \sup_{\gamma < \beta} \alpha + \gamma$ for $\beta \in \mathrm{Lim}$.

Ordinal multiplication is defined via

$\alpha \cdot 0 = 0$,

$\alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha$ and

$\alpha \cdot \beta = \sup_{\gamma < \beta} \alpha \cdot \gamma$ for $\beta \in \mathrm{Lim}$.

Ordinal exponentiation is defined via:

$\alpha^0 = 1$,

$\alpha^{\beta+1} = (\alpha^\beta) \cdot \alpha$ and

$\alpha^\beta = \sup_{\gamma < \beta} \alpha^\gamma$ for $\beta \in \mathrm{Lim}$.

Ordinal addition, multiplication and exponentiation follow (mostly) the rules one would expect.

Lemma 3.38. (*i*) $+, \cdot$ *are associative.*

(*ii*) $+, \cdot$ *are* ***not*** *commutative. Nonetheless* $+, \cdot$ *restricted to natural numbers are commutative.*

(*iii*) *The following distributive law holds: If* α, β, γ *are ordinals then*

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

(*iv*) *If* $\alpha \leq \beta$ *then there is a unique* γ *so that* $\alpha + \gamma = \beta$.

Proof. Exercise. □

Remark 3.39. We have essentially shown that if $\mathcal{M}$ is a model of ZF then $(\omega, 0, 1, +, \cdot)$ as calculated in $\mathcal{M}$ is a model of Peano Arithmetic (PA). It is worth noting that not every model of Peano Arithmetic is of this form. There are sentences φ in the language of arithmetic which are not provable in PA, yet hold in every model as above. For the reader who has been exposed to Gödel's incompleteness theorems, it is perhaps not a shock that the sentence attesting the consistency of PA is one of those. Though there are more natural such sentences φ, for example Goodstein's theorem.

We have now enough tools at our disposal to encode essentially all of mathematics into Set Theory. We can define

- $(\mathbb{Z}, +, \cdot)$ from $(\mathbb{N}, +, \cdot)$ by defining appropriate operations on $\mathbb{N} \times 2$ (where $(n, 0)$ is supposed to be the integer n and $(n, 1)$ is supposed to code $-(n+1)$),
- $(\mathbb{Q}, +, \cdot)$ by addition and multiplications on the equivalence classes of an appropriate equivalence relation $\sim$ on $\mathbb{Z} \times \mathbb{N}$ (where $[(i, n)]_\sim$ is supposed to code the fraction $\frac{i}{n}$),
- $(\mathbb{R}, +, \cdot)$ via Dedekind cuts from $\mathbb{Q}$,
- $(\mathbb{C}, +, \cdot)$ by defining multiplication appropriately on the vector space $\mathbb{R}^2$, etc.

We will refrain from doing so in detail and encourage the interested reader to seek more information elsewhere.

So far, we have done induction recursion along $\in$. We will now explain how this can be generalized to other relation. Occasionally, this will come in handy.

Definition 3.40. A binary relation R on X is **set-like** if for all $x \in X$ we have

$$\mathrm{pred}_R(x) \coloneqq \{y \mid yRx\}$$

is a set.

Theorem 3.41 (The General Recursion Theorem)**.** *Suppose that R is a binary set-like wellfounded relation and $F\colon V \to V$ is a function. Then there is a function $G\colon V \to V$ satisfying*

$$G(x) = F(G \restriction \mathrm{pred}_R(x)).$$

The proof is almost exactly the same as for Theorem 3.27. We leave it to the reader to formalize induction along a binary wellfounded set-like relation. The above theorem cannot be generalized any further: If the recursion theorem holds for a binary relation R then R is wellfounded and set-like (though, admittedly, the proof that R must be set-like relies on the exact definition of function application $F(x)$).

Definition 3.42. A binary relation R on X is a **extensional** iff for all $x, y \in X$ we have $x = y \leftrightarrow \mathrm{pred}_R(x) = \mathrm{pred}_R(y)$.

The $\in$-relation is wellfounded, set-like and extensional. We will see that $\in$ is essentially the only relation with these properties: all other ones are (isomorphic to) restrictions of $\in$, even to transitive sets.

Proposition 3.43. *Suppose X, Y are transitive and $\pi\colon (X, \in) \to (Y, \in)$ is an isomorphism. Then $X = Y$ and $\pi = \mathrm{id}_X$.*

Proof. We show $\pi(x) = x$ by induction on $x \in X$. Suppose $\pi(y) = y$ for all $y \in X$. If $\pi(x) \neq x$, then there is some $z \in \pi(x) \setminus \pi[x]$. As Y is transitive, $z \in Y$ and hence there must be some $y \in X$ with $\pi(y) = z$. But as π is an isomorphism, we have $y \in x$, contradiction.

So $\pi = \mathrm{id}_X$ and since π is surjective, $Y = X$. □

Lemma 3.44 (Mostowki's Collapse Lemma)**.** *Suppose that R is a wellfounded set-like extensional binary relation on X. Then there is a unique transitive Y so that*

$$(X, R) \cong (Y, \in).$$

Moreover, the isomorphism is unique.

Proof. By the General Recursion Theorem, there is a function

$$G\colon X \to V$$

which satisfies $G(x) = G[\mathrm{pred}_R(x)]$ for all $x \in X$. Simply plug in any function $F\colon V \to V$ which takes functions $f \in V$ to their range $\mathrm{ran}(f)$. Let $Y = \mathrm{ran}(G)$.

Claim 3.45. *Y is transitive.*

Proof. Suppose $b \in a \in Y$. We can find $x \in X$ so that $a = G(x)$. By definition of G, there is yRx with $b = G(y)$, in particular $b \in Y$. □

Claim 3.46. *G is an isomorphism.*

Proof. Clearly G is surjective. Let us show that G is injective. Suppose not and let x be R-minimal such that for some $x' \neq x$, $G(x) \neq G(x')$. Such an x exists as R is wellfounded. But then whenever yRx then $G(y) = G(y')$ implies $y = y'$. This implies

$$G(x) = G[\mathrm{pred}_R(x)] = G[\mathrm{pred}_R(x')] = G(x'),$$

contradiction. □

It remains to show uniqueness of Y and the isomorphism G. If one of those fails then, by composing two such isomorphisms, we get a nontrivial isomorphism between $\pi\colon (Y, \in) \to (Y', \in)$ with Y, Y' transitive. This contradicts Proposition 3.43. □

As an immediate consequence, we can classify all wellorders.

Corollary 3.47. *For any wellordered set $(x, \prec)$, there is a unique ordinal α with*

$$(x, \prec) \cong (\alpha, <).$$

Moreover, the isomorphism is unique.

Definition 3.48. If $(x, \prec)$ is a wellorder on a set x then the **ordertype** of $(x, \prec)$ (or just of $\prec$) is the unique ordinal α with $(x, \prec) \cong (\alpha, \prec)$. We write $\mathrm{otp}((x, \prec)) = \alpha$, or just $\mathrm{otp}(\prec) = \alpha$.

4 Cardinals

19.3.24

In some sense, Ordinals measure length. Specifically the length of wellorders. We know introduce cardinals which measure "size".

Definition 4.1. For sets $x, y \in V$, we write $x \preceq y$ iff there is an injection $f\colon x \hookrightarrow y$.

We write $x \approx y$ iff there is a bijection $g\colon x \leftrightarrow y$.

Clearly, $x \approx y$ implies $x \preceq y$ and $\approx$ is an equivalence relation. The idea is that if $x \preceq y$ then y is at least as large as x and if $x \approx y$ then x, y have the same size. "Cardinality" is simply the word for size in this context. As cardinals should be the abstract possible measurments of size, it is reasonable to define cardinals as equivalence classes $[x]_\approx$. The problem with this is that $[x]_\approx$ is a proper class whenever $x \neq \emptyset$ (why? Otherwise, we can find an $\in$-cycle starting and ending with $[x]_\approx$ by considering $x \times \{[x]_\approx\} \approx x$). However, we would like to have a class of all cardinals. We seek other solutions for this problem.

Definition 4.2. A **notion of cardinality** is a function $F\colon V \to V$ so that

$$\forall x \forall y\ x \approx y \leftrightarrow F(x) = F(y).$$

Cardinals (w.r.t. F) are elements of $\mathrm{ran}(F)$. The **class of all cardinals** is

$$\mathrm{Card} = \{|x| \mid x \in V\}.$$

We usually write $|x|$ instead of $F(x)$ and say that x **is of cardinality** $F(x)$.

A notion of cardinality is a uniform way of encoding the equivalence classes $[x]_\approx$ as sets. One way to do this is to pick a class of representatives for the equivalence relation $\approx$, but unfortunately such a class does not necessarily exist. A better way is to employ "Scott's trick".

Definition 4.3. We define $F_{\mathrm{CL}}(x) = [x]_\approx \cap V_\alpha$ where α is the least ordinal β so that $[x]_\approx \cap V_\beta \neq \emptyset$.

It is straightforward to show that F_{CL} is a notion of cardinality. It does not really matter which notion of cardinality we make use of, this is simply the standard one in a "choice-less" context (hence the CL subscript). When we adopt the axiom of choice later, we switch to a more convenient notion of cardinality.

Note that the $\preceq$ relation factors through the equivalence relation $\approx$ and hence induces a relation $\leq$ on Card.

4.1 The structure of $(\mathrm{Card}, \leq)$

We hold our promise from earlier and prove the Cantor-Schröder-Bernstein theorem. We note that it is elementary to state, has a simple proof, yet is non-trivial (of course these are all a matter of opinion). Because of this, every mathematician should see the proof at least once in their career.

Theorem 4.4 (Cantor-Schröder-Bernstein)**.** *The relation* $(\mathrm{Card}, \leq)$ *is antisymmetric.*

Proof. Let x, y be sets such that $x \preceq y$ and $y \preceq x$. We have to show that $x \approx y$. We will do a proof by picture. Let $f\colon x \to y$, $g\colon y \to x$ be two injections. We may assume w.l.o.g. that $x \cap y = \emptyset$. Now, consider the directed graph $\mathcal{G}$ on $x \cup y$ which has an edge from a to b iff either $a \in x$ and $f(a) = b$ or $a \in y$ and $g(a) = b$. Note that

- any $a \in x \cup y$ has exactly one outgoing edge,
- any $a \in x \cup y$ has at most one incoming edge and
- $\mathcal{G}$ is bipartite.

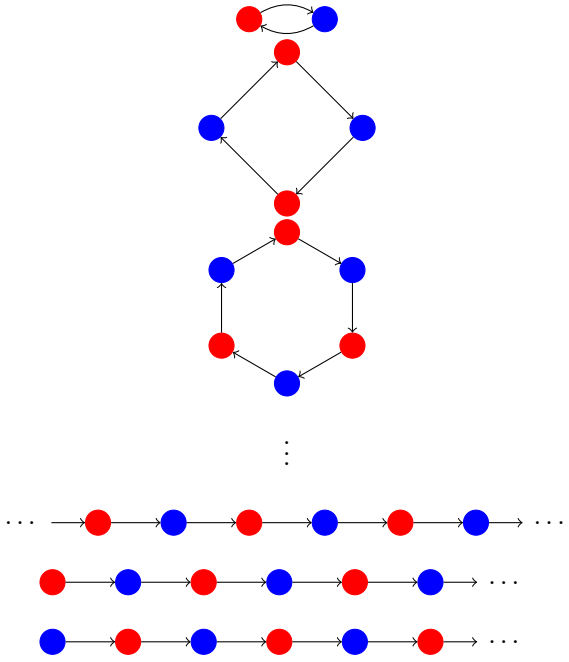
Consider the connected components of $\mathcal{G}$. These can be classified as follows: A connected component can be

(i) a cycle of even length,

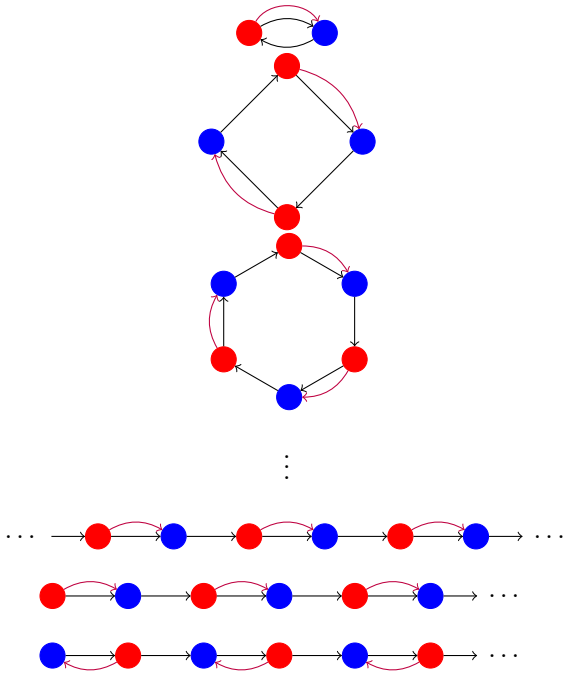
(ii) a chain infinite in both directions or

(iii) an infinite chain with a starting point either in x or y.

Coloring points in x blue and points in y red, these look as follows:



We now define a function $h\colon x \to y$ by adding purple arrows which determine to which blue node the red node at the base of the arrow maps to.



The function h is obviously a bijection, so we are done. $\square$

Corollary 4.5. $(\mathrm{Card}, \leq)$ *is a partial order.*

Theorem 4.6 (Cantor's Theorem)**.** *For any x, we have $\mathcal{P}(x) \not\preceq x$. In particular $x \prec \mathcal{P}(x)$, so there is no maximal cardinal.*

Proof. Clearly $x \preceq \mathcal{P}(x)$ since $a \mapsto \{a\}$ is injective. Assume toward a contradiction that $\mathcal{P}(x) \preceq x$ then $x \approx \mathcal{P}(x)$ by Theorem 4.4, say $f\colon x \to \mathcal{P}(x)$ is bijective. Consider the subset

$$y \coloneqq \{a \in x \mid a \notin f(a)\}.$$

But if $f(a) = y$ then

$$a \in y \Leftrightarrow a \notin f(a) \Leftrightarrow a \notin y,$$

contradiction. $\square$

If $(X, \lhd)$ is a partial order then a class $Y \subseteq X$ is

- **cofinal** if for all $x \in X$ there is $y \in Y$ with $x \lhd y$,
- **unbounded** if there is no $x \in X$ with $y \lhd x$ for all $y \in Y$.

Lemma 4.7. (*i*) *The class $\{|V_\alpha| \mid \alpha \in \mathrm{Ord}\}$ is cofinal in* $(\mathrm{Card}, \leq)$.

(*ii*) *The class $\{|\alpha| \mid \alpha \in \mathrm{Ord}\}$ is unbounded in* $(\mathrm{Card}, \leq)$.

Part (*ii*) above is known as Hartog's Lemma.

Proof. (*i*): For $x \in V$, we can find $\alpha \in \mathrm{Ord}$ so that $x \in V_\alpha$. As V_α is transitive, $x \subseteq V_\alpha$, so the inclusion witnesses $x \preceq V_\alpha$.
(*ii*): Once again, let $x \in V$. We have to find an $\alpha \in \mathrm{Ord}$ so that $\alpha \not\preceq x$. Let

$$\mathrm{pwo}(x) = \{\lhd \mid \lhd \text{ is a wellorder on some } y \subseteq x\}$$

be the class of all partial wellorders on x. Note that

$$\mathrm{pwo}(x) \subseteq \mathcal{P}(x \times x)$$

so that $\mathrm{pwo}(x)$ is a set by Proposition 3.8, (Power) and (Separation). By Corollary 3.47, we can define the function $f\colon \mathrm{pwo}(x) \to \mathrm{Ord}$ by $f(\lhd) = \mathrm{otp}(\lhd)$. By (Replacement), $\mathrm{ran}(f)$ is a set and we let $\alpha = \sup \mathrm{ran}(f) + 1$.

We are done if we can show $\alpha \not\preceq x$. So assume otherwise and let $f\colon \alpha \hookrightarrow x$ be an injection. Then we can transport the canonical wellorder of α onto $y \coloneqq \mathrm{ran}(f)$ via $a \lhd b$ iff $f^{-1}(a) \in f^{-1}(b)$. Hence $\lhd \in \mathrm{pwo}(x)$ and $\mathrm{otp}(\lhd) = \alpha$, contradiction. $\square$

Hartog's Lemma motivates the next definition.

Definition 4.8. For a set x, let $x^+ = \min\{\alpha \in \mathrm{Ord} \mid \alpha \not\preceq x\}$.

Special importance among the cardinals is given to the cardinalities of ordinals.

Definition 4.9. A set x is **wellordered** if there is a $\lhd$ so that $(x, \lhd)$ is a wellorder.
A cardinal κ is **wellordered** if any/all sets x of cardinality κ are wellordered.
WOCard is the class of wellordered cardinals.

The connection between ordinals and wellordered sets is given by the following easy consequence of Corollary 3.47.

Proposition 4.10. *The following are equivalent for any set x:*

(*i*) *x is wellordered.*

(*ii*) *There is an ordinal α with $x \approx \alpha$.*

It follows that WOCard $= \{|\alpha| \mid \alpha \in \mathrm{Ord}\}$. It is convenient to order the infinite wellordered cardinals increasingly.

Definition 4.11. Define $\aleph\colon \mathrm{Ord} \to \mathrm{WOCard}$ recursively by

- $\aleph(0) = \omega$ and
- $\aleph(\alpha) = |\beta|$ where $\beta = \min\{\gamma \in \mathrm{Ord} \mid \gamma \geq \omega \wedge |\beta| \notin \bigcup \aleph[\beta]\}$ for $\alpha > 0$.

We usually write $\aleph_\alpha$ instead of $\aleph(\alpha)$. We also define

$$\omega_\alpha = \min\{\beta \in \mathrm{Ord} \mid |\beta| = \aleph_\alpha\}.$$

Note that WOCard is a proper class by Hartog's Lemma so the above recursion makes sense.

Proposition 4.12. *For $\alpha \in \mathrm{Ord}$, $\omega_{\alpha+1} = \omega_\alpha^+$ and for $\gamma \in \mathrm{Lim}$, $\omega_\gamma = \sup_{\beta<\gamma} \omega_\gamma$.*

Proof. Exercise. □

The axiom system ZF does not prove much more about the structure of $(\mathrm{Card}, \leq)$ than we did above. It is consistent with ZF that $(\mathrm{Card}, \leq)$ is not a linear order, has infinite decreasing sequences and many other things.

4.2 The Axiom of Choice

We now introduce the Axiom of Choice and show that the cardinals are much better behaved assuming it.

Definition 4.13. The **Axiom of Choice** (AC) is the sentence

$$\forall x \forall f \left((f\colon x \to V \setminus \{\emptyset\}) \to \exists g \ (g\colon x \to V) \wedge \forall y \in x \ g(y) \in f(y) \right).$$

The system ZFC (**Zermelo-Fraenkel with Choice**) is ZF + AC.

If $f\colon x \to V \setminus \{\emptyset\}$ then a function $g\colon x \to V$ is called a **choice function for** f if $\forall y \in x\ g(y) \in f(y)$. With this terminology, the Axiom of Choice asserts that any such function f on a set x admits a choice function.

The system ZF proves only a tiny fragment of the Axiom of Choice.

Lemma 4.14 (Finite Choice)**.** *For any $n \in \omega$, any function $f\colon n \to V \setminus \{\emptyset\}$ admits a choice function.*

Proof. By induction on $n \in \omega$. The base case $n = 0$ is trivial as the only function $f\colon \emptyset \to V$ is the empty function $f = \emptyset$, which is its own choice function.
Now assume $f\colon n+1 \to V \setminus \{\emptyset\}$ is a function. By induction, we can find a choice function g' for $f \restriction n$. As $f(n) \neq \emptyset$, we can pick some $a \in f(n)$. Finally, $g = g' \cup \{(n, a)\}$ is a choice function for f. □

Naively, one might think that it may be possible to continue this induction. The next step would be to try and prove **Countable Choice** (AC_ω), the statement that any $f\colon \omega \to V \setminus \{\emptyset\}$ admits a choice function. The naive proof attempt runs as follows: Suppose f is as above. Then for each $n < \omega$, there is a choice function g_n for $f \restriction n+1$ and then $g\colon \omega \to V$ defined by $g(n) = g_n(n)$ is a choice function for f. This does not work! The problem is that the existence of a single g_n for each n is not enough. We need a function $G\colon \omega \to V$ so that $G(n)$ is a choice function for $f \restriction n+1$ to make the argument work. But to find G, we would want to apply AC_ω to the function $F\colon \omega \to V \setminus \{\emptyset\}$ defined by

$$F(n) = \{h\colon n+1 \to V \mid h \text{ is a choice function for } f \restriction n+1\},$$

however we are trying to prove AC_ω in the first place!

This problem cannot be avoided with a more sophisticated proof. Indeed, AC_ω is not provable in ZF (unless ZF is inconsistent).