Set Theory – Lecture Notes

Version of March 6, 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

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 $\{0, \dots, n\}$ 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 $b^* = b_1 b_2 \dots$. 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

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

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

$$[\![T]\!] \ni (0.b_1 b_2 \dots)_2 = \lim_{k\to\infty} (0.a_1^{n_k} a_2^{n_k} \dots)_2 = \lim_{n\to\infty} x_n = x.$$

□

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

Definition 1.7. Suppose T is a binary tree.

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

2. 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,

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

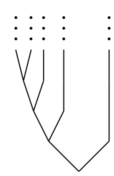
which shows $(0.b^*) = x \in \overline{[\![T]\!] \setminus \{x\}}$ $\square$

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.

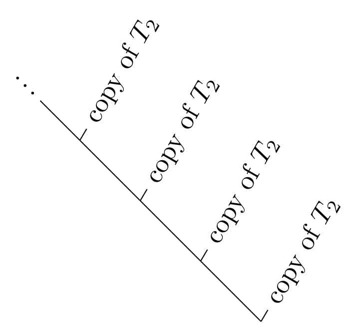
Definition 1.9. If T is a binary tree then the **derivative of** T is the binary tree

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

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 $\operatorname{ran}(j)$ 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 $b \mapsto (0.b^*)_2$. $\square$

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 $\forall z \ (z \in x \to z \in y)$.

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 $\forall x \forall y x+1 = y+1 \rightarrow x = y$ 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 $f : x \rightarrow y$ 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*

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

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 $\{x \mid \varphi(x, v_0, \dots, v_n)\}$ and $\{y \mid \psi\}$, we set

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

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

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

Suppose R is a binary relation.

1. xRy is syntactic sugar for $(x, y) \in R$.
2. The **domain** of R is $\mathrm{dom}(R) = \{x \mid \exists y \ xRy\}$.
3. The **range** of R is $\mathrm{ran}(R) = \{y \mid \exists x \ xRy\}$.

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

1. F is a **function** if $\forall x \forall y \forall y' \ (xFy \land xFy') \to y = y'$.
2. 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 $F \colon x \to y$.
3. The **value** of F at x is

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

4. 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

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

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

$$A \neq \emptyset \rightarrow \exists x \in A\ A \cap x = \emptyset$$

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*

$$\neg(\exists x_0 \dots \exists x_n\ x_0 \in x_1 \wedge \cdots \wedge x_{n-1} \in x_n \wedge x_n \in x_0)$$

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

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

Proof. Suppose $x_0 \in x_1, \dots, x_{n-1} \in x_n$ and $x_n \in x_0$. We apply (Foundation) to the class term $A = \{x_0, \dots, x_n\}$. 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. $\square$

[1] In other sources, F"x is a common alternative notation for $F[x]$.

Intuitively, a similar argument shows that there are no infinite descending $\in$-chains $x_0 \ni x_1 \ni x_2 \ni \dots$, 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.

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

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

3. Ord is the class $\{x \mid x \text{ is an ordinal}\}$.

Examples 3.2

- $\emptyset$ *is trivially an ordinal. We set* $0 := \emptyset$.

- $\{\emptyset\} = 0 + 1$ *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 $\beta \in \alpha \in \mathrm{Ord}$, 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. $\square$

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! $\square$