Set Theory – Lecture Notes

Version of May 29, 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.

(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,

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

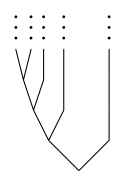
which shows $(0.b^*) = x \in \overline{[\![T]\!] \setminus \{x\}}$ $\square$

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.

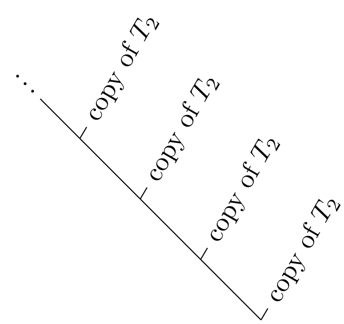
Definition 1.9. If T is a binary tree then the **derivative of** T is the binary tree

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

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 $b \mapsto (0.b^*)_2$. □

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

- $x \cup y \coloneqq \bigcup \{x, y\}$,
- $\bigcap x \coloneqq= \{z \mid \forall u (u \in x \rightarrow z \in u)\}$,
- $x \cap y \coloneqq \bigcap \{x, y\}$ 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 $\forall z \ (z \in x \rightarrow z \in y)$. We also let $\forall x \in y \varphi$ be sugar for $\forall x (x \in y \rightarrow \varphi)$, 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 $\forall x \forall y x + 1 = y + 1 \rightarrow x = y$ 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 $f : x \to y$ 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 $y = y'$. $\square$

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 $\forall x \forall y \forall y' \ (xFy \wedge xFy') \rightarrow y = y'$.

(*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 $F \colon x \rightarrow y$.

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

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

(*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

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

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.

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

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*

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

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. □

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.

(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* $0 := \emptyset$.
- $\{\emptyset\} = 0 + 1$ *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. □

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. Ord $\notin V$.

Proof. Suppose Ord $\in V$. Then Ord $\in$ Ord which contradicts Proposition 2.20. $\square$

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

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

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. $\square$

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, $\exists y \in x \forall z \in x\ \neg z R y$.

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. $\square$

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 $\beta + 1 \neq \alpha$. $\square$

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]

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

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. $\square$

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 $\forall \alpha \in \mathrm{Ord} \; \alpha \subseteq A \rightarrow \alpha \in \mathrm{Ord}$. Then $A = \mathrm{Ord}$.*

[2]Sometimes, 0 is included in Lim.

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. $\square$

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 $\alpha \in A$. $\square$

Now we get to recursive constructions.

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

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. $\square$

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 $G := \bigcup\{g \mid g$ is F-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

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

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

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

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

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

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

$$g := \{(x, F(g'))\}$$

witnesses $x \in \mathrm{dom}(G)$. □

$\square$

Remark 3.31. The resulting recursion G along F is unique in the sense that whenever $G' \colon V \to V$ also satisfies $\forall x G'(x) = F(G' \upharpoonright x)$ 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) $G(\alpha) = F(G \upharpoonright \alpha)$ *for limit ordinals α.*

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

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

$\square$

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 $x \in V_{\delta+1}$. $\square$

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* (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.

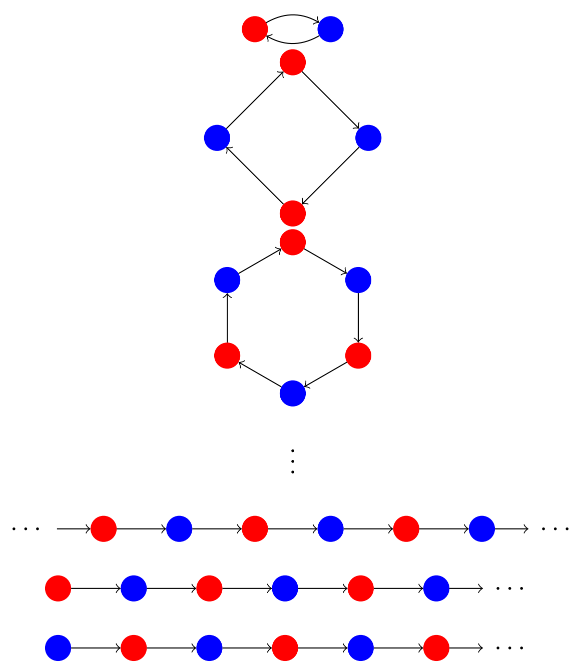
Consider the connected components of $\mathcal{G}$. These can be classified as follows: A connected component can be

(i) a cycle of even length,

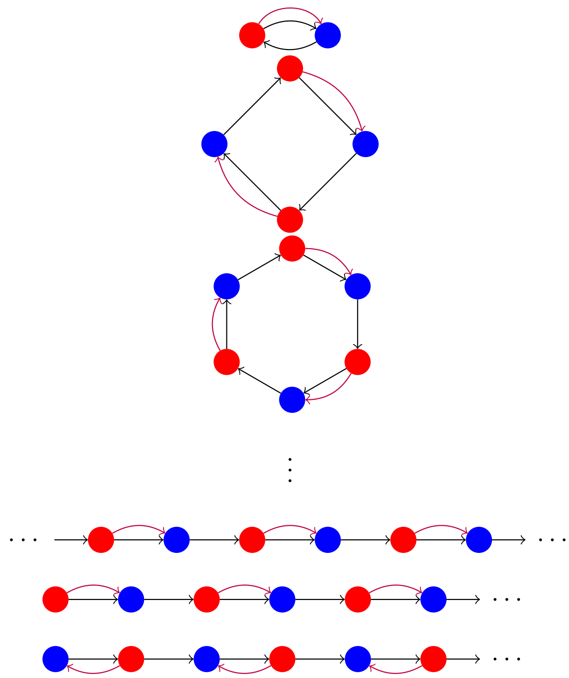
(ii) a chain infinite in both directions or

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

- **cofinal** if for all $x \in X$ there is $y \in Y$ with $x \triangleleft y$,
- **unbounded** if there is no $x \in X$ with $y \triangleleft x$ 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 $y := \mathrm{ran}(f)$ 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 proposition.

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

(i) *x is wellordered.*

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

(iii) *There is a wellordered y and an injection $f\colon x \hookrightarrow y$.*

(iv) *There is a wellordered y and a surjection $g\colon y \to x$.*

Proof. The equivalence of $(i) - (iii)$ is an easy consequence of Corollary 3.47 and trivially (i) implies (iv). On the other hand, (iv) implies (iii) by defining $f(a) = \min_\lhd g^{-1}(\{a\})$ for some wellorder $\lhd$ on y. $\square$

It follows that $\mathrm{WOCard} = \{|\alpha| \mid \alpha \in \mathrm{Ord}\}$. 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. □

Proposition 4.13. *The wellordered cardinals are exactly*

$$\{|n \mid n < \omega\} \cup \{\aleph_\alpha \mid \alpha \in \mathrm{Ord}\}$$

and these cardinals are all different.

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 4.14. 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 4.15 (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).

The Axiom of Choice is perhaps the most controversial axiom of ZFC. Some of its consequences seem obviously true, some other obviously false. Thus it is important to know that adding AC to ZF does not lead to any contradictions. A proof of this will be a cornerstone of this lecture. 20.3.24

We will now prove one of the more controversial consequences of AC, the Wellordering Theorem. Though undeniably, it is very useful.

Theorem 4.16 (Wellordering Theorem)**.** *If* AC *holds then every set is wellordered.*

Proof. Let $x \in V$. Our strategy is to build a wellorder on x recursively. At each step in the construction, we need to decide which element on x we would like to put on top of our wellorder next. There are usually many options left for this next element and none of them stand out particularly, so we will make use of a choice function that makes this decision for us uniformly.

Let $f\colon \mathcal{P}(x) \setminus \{\emptyset\} \to \mathcal{P}(x) \setminus \{\emptyset\}$ be the identity function. By AC, there is a choice function g for f, so $g(a) \in a$ for all non-empty $a \subseteq x$. The idea is that when we have built our wellorder partially and a are the remaining elements of x not yet on the wellorder then $g(a)$ should be the next point.

By the Recursion Theorem 3.27, there is a function

$$G\colon \mathrm{Ord} \to x \cup \{x\}$$

so that $G(\alpha) = g(x \setminus G[\alpha])$ if $x \nsubseteq G[\alpha]$ and $G(\alpha) = x$ otherwise. By Hartog's Lemma, there is some least α so that $G(\alpha) = x$ and hence $G \restriction \alpha\colon \alpha \to x$ is a bijection. We are done by Proposition 4.10. $\square$

Corollary 4.17. *The following are equivalent:*

(*i*) AC.

(*ii*) *Every set is wellordered.*

(*iii*) Card $=$ WOCard.

(*iv*) Card *is linearly ordered by* $\leq$.

Proof. (*i*)$\Rightarrow$(*ii*) is Theorem 4.16 and (*ii*)$\Rightarrow$(*iii*)$\Rightarrow$(*iv*) is trivial. To complete the proof, we will show (*iv*)$\Rightarrow$(*ii*) and (*ii*)$\Rightarrow$(*i*). So first assume that $\leq$ linearly orders Card and let x be a set. Then by Hartog's Lemma, there is some $\alpha \in \mathrm{Ord}$ so that $\alpha \not\preceq x$. But then we must have $x \preceq \alpha$. If $f\colon x \hookrightarrow \alpha$ is injective then $\mathrm{ran}(f)$ is clearly wellordered by $\in$, so x is wellordered as well.

Now assume every set is wellordered and we will prove AC. Suppose $f\colon x \to V \setminus \{\emptyset\}$ is a function. Then $\bigcup \mathrm{ran}(f) \in V$ by (Replacement) and (Union) and hence there is a wellorder $\lhd$ on $\bigcup \mathrm{ran}(f)$. We can now define a function $g\colon x \to V$ by mapping $a \in x$ to the $\lhd$-least element of a (note that $a \subseteq \bigcup \mathrm{ran}(f)$). Clearly, g is a choice function for f. $\square$

4.3 Cardinal Arithmetic

We now define a version of addition, multiplication and exponentiation for cardinals. For sets x, y, we let ${}^y x$ be the set of functions $f\colon x \to y$. Note that ${}^y x \in V$ as it is a subset of $\mathcal{P}(x \times y)$.

Definition 4.18. Suppose $\kappa, \lambda \in \mathrm{Card}$ and $x, y \in V$ are disjoint and of cardinality κ, λ respectively.

(*i*) $\kappa + \lambda := |x \cup y|$.

(*ii*) $\kappa \cdot \lambda := |x \times y|$.

(*iii*) $\kappa^\lambda := |{}^y x|$.

Note that in the above definition, the cardinals $\kappa+\lambda, \kappa\cdot\lambda, \kappa^\lambda$ do not depend on the choice of x and y. Moreover, it is always possible to find x of cardinality κ and y or cardinality λ so that $x \cap y = \emptyset$ as e.g. one could always replace x by $x \times \{y\}$. x and y being disjoint is only important in the definition of $\kappa + \lambda$.

Lemma 4.19. *Suppose* κ, λ, μ *are cardinals.*

(*i*) $+, \cdot$ *are associative and commutative on* Card.

(*ii*) *The following distributive law holds*

$$\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu.$$

(*iii*) $\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$.

(*iv*) $\left(\kappa^\lambda\right)^\mu = \kappa^{\lambda\cdot\mu}$.

Proof. (*i*) is easy. Now let x, y, z be of size κ, λ, μ respectively and pairwise disjoint. Then $x \times (y \cup z) = (x \times y) \cup (x \times z)$, so (*ii*) follows. Next,

$$x^{y \cup z} \approx x^y \times x^z$$

as witnessed by the map $f \to (f \restriction y, f \restriction z)$. (*iii*) follows. Finally,

$${}^z({}^y x) \approx {}^{y \times z} x$$

$\square$

4.4 Finite and countable cardinals

Definition 4.20. A set x is finite if $x \approx n$ for some $n \in \omega$. A cardinal κ is finite if any/all x of cardinality κ are finite.

By induction, it is straightforward to show that ω is closed under ordinal arithmetic, i.e. $n+m, n\cdot, n^m < \omega$ for $n, m < \omega$. This easily implies the following:

Proposition 4.21. *For $n, m < \omega$ we have $|n| + |m|, |n| \cdot |m|, |n|^{|m|}$ are all finite cardinals.*

We mention one more interesting fact (which we will not make further use of, so we will not give a proof).

Proposition 4.22. *If x is a finite set and $f\colon x \to x$ is either injective or surjective then f is bijective.*

This property is known as Dedekind-finiteness. Under the axiom of choice, Dedekind-finiteness is equivalent to finiteness.

Proposition 4.23. *If $\alpha \geq \omega$ is an ordinal then $\alpha + 1 \preceq \alpha$. In particular, α is not Dedekind-finite.*

Proof. Mapping α to 0, $n \in \omega$ to $n + 1$ and all other ordinals to themselves yields an injection $f\colon \alpha + 1 \hookrightarrow \alpha$. □

However, if AC fails there may be infinite Dedekind-finite sets.

Definition 4.24. A set x is **countable** or **enumerable** if $x \preceq \omega$. Otherwise, x is **uncountable**.

We have not yet officially adopted a notation for sequences: $\langle x_i \mid i \in I\rangle$ is just another notation for the function with domain I which maps $i \in I$ to x_i.

Proposition 4.25. *Suppose* AC *holds. A countable union of countable sets is countable. That is, if I is countable and $\langle x_i \mid i \in I\rangle$ is a sequence of countable sets then $\bigcup_{i\in I} x_i$ is countable.*

Proof. We may assume that $I = \omega$. We get a sequence $\langle g_n \mid n < \omega\rangle$ of surjections $g_n\colon \omega \to x_n$ by applying AC to the function $F\colon \omega \to V$ defined via

$$F(n) = \{g \mid g\colon \omega \to x_i \text{ is surjective}\}.$$

We get a surjection $f\colon \omega \times \omega \to \bigcup_{i\in I} x_i$ via

$$f((n, m)) = g_n(m).$$

It is easy to see that $\omega \times \omega \approx \omega$, for example we get an injection $\omega \times \omega \to \omega$ via $(n, m) \mapsto 2^n \cdot 3^m$. □

Once again, this is not provable in ZF alone. For example, it is consistent with ZF that the reals are a countable union of countable sets. In such a universe, much of analysis and measure theory completely falls apart.

4.5 Cardinal arithmetic under AC

Convention *We now switch to a different notion of cardinality: Let* $F_{\text{std}}\colon V \to V$ *be defined via*

$$F_{\text{std}}(x) = \begin{cases} \min\{\alpha \in \mathrm{Ord} \mid x \approx \alpha\} & \text{if } x \text{ is wellordered} \\ F_{\mathrm{CL}}(x) & \text{else.} \end{cases}$$

While the case split seems somewhat unnatural, this notion of cardinality is more useful to work with and this is the standard notion of cardinality used in practice. Note that none of what we proved depended on the specific choice of our prior notion of cardinality. Here are some side-effects of this switch:

- Wellordered cardinals are ordinals. In fact these are exactly those ordinals[3] α which do not inject into any smaller $\beta < \alpha$.
- We now have $\aleph_\alpha = \omega_\alpha$, so the distinction between them is only syntactical. We use the symbol $\aleph_\alpha$ if we think about it as a cardinal and ω_α if we think about ordinals.

Convention *From now on* ***we work in*** ZFC*. If we assume a different theory in a theorem/lemma/etc, we mark it with it, e.g. (*ZF*).*

In particular, we have Card = WOCard and thanks to our new notion of cardinality, all cardinals are ordinals (but not all ordinals are cardinals of course). The downside of this is that the symbols $+, \cdot$ are overloaded with ordinal and cardinal arithmetic. If it is not clear from context, we will from now on denote ordinal addition and multiplication by $+_{\mathrm{Ord}}, \cdot_{\mathrm{Ord}}$ respectively and reserve $+, \cdot$ for cardinal arithmetic. Ordinal and cardinal exponentiation can only be differentiated by context unfortunately.

These conventions make it easy to state, e.g. the following:

Lemma 4.26. *The cardinals are closed and unbounded in the ordinals, i.e.*

(i) if x is a set of cardinals then $\bigcup x$ is a cardinal and

(ii) for any $\alpha \in \mathrm{Ord}$ there is a cardinal $\kappa > \alpha$.

Proof. We already know that (ii) holds, so we show (i). Let $x \subseteq \mathrm{Card}$ be a set, so $\kappa := \sup x = \bigcup x \in \mathrm{Ord}$. We are done if we can show that $\kappa \not\preceq \alpha$ for all $\alpha < \kappa$. But if $\alpha < \kappa$ then there is $\lambda \in x$, $\lambda \leq \kappa$ with $\alpha < \lambda$. As λ is a cardinal, $\lambda \not\preceq \alpha$, so in particular $\kappa \not\preceq \alpha$. □

Lemma 4.27. *Let κ, λ be cardinals. Then $\kappa + \lambda = \kappa \cdot \lambda = \max\{\kappa, \lambda\}$.*

The proof of this is based on a famous wellorder.

[3] Such ordinals are sometimes called initial ordinals. In the context of ZFC, these are just the cardinals and referred to as such.

Definition 4.28. Gödel's wellordering of Ord^2 is defined by

$$\begin{aligned}(\alpha,\beta) <_G (\gamma,\delta) :\Leftrightarrow &\max\{\alpha,\beta\} < \max\{\gamma,\delta\} \\ &\vee(\max\{\alpha,\beta\} = \max\{\gamma,\delta\} \wedge \alpha < \gamma) \\ &\vee(\max\{\alpha,\beta\} = \max\{\gamma,\delta\} \wedge \alpha = \gamma \wedge \gamma < \delta)\end{aligned}$$

Lemma 4.29. $(\mathrm{Ord}^2, <_G)$ *is a wellordering of ordertype* Ord*. If* κ *is an infinite cardinal then the restriction* $(\kappa^2, <_G)$ *is of ordertype* κ *and an initial segment of* $(\mathrm{Ord}^2, <_G)$*.*

10.04.24

Proof. Clearly, $<_G$ is a set-like extensional linear order. It is not hard to see that $<_G$ is wellfounded as well: To find a $<_G$-minimal element of $x \subseteq \mathrm{Ord}^2$, first minimize the maximum of both coordinates, then the first coordinate and finally the second. By Mostowski' Collapse Lemma, let $C\colon \mathrm{Ord}^2 \to \mathrm{Ord}$ be the collapse. We have to show that if κ is an infinite cardinal then $C[\kappa \times \kappa] = \kappa$. As κ is infinite, $\kappa = \aleph_\alpha$ for some α and we will do an induction along α.

$\alpha = 0$: If $n, m < \omega$ then $(n+1) \times (m+1)$ is finite by Proposition 4.21, hence, as $\mathrm{pred}_{<_G}((n,m)) \subseteq (n+1) \times (m+1)$, we must have $C((n,m)) < \omega$. Clearly $C[\omega \times \omega]$ is infinite, so we we must have $C[\omega \times \omega] = \omega$.

$\alpha = \beta + 1$. This is basically the same argument, only higher up. Suppose $(\gamma,\delta) \in \aleph_\alpha \times \aleph_\alpha$. Then $\mathrm{pred}_{<_G}((\gamma,\delta)) \subseteq (\gamma+1) \times (\delta+1)$. As $\gamma, \delta < \aleph_\alpha$, both $\gamma+1, \delta+1$ are of size at most $\aleph_\beta$. By induction, a restriction of C witnesses that $\aleph_\beta \times \aleph_\beta \approx \aleph_\beta$. It follows that there is a surjection $f\colon \aleph_\beta \to \mathrm{pred}_{<_G}((\gamma,\delta))$. As $\aleph_\alpha$ is a cardinal, there is no injection $\aleph_\alpha \hookrightarrow \aleph_\beta$, so we must have $C((\gamma,\delta)) < \aleph_\alpha$. Since $\aleph_\alpha \times \aleph_\alpha$ has size at least $\aleph_\alpha$, we must have $C[\aleph_\alpha \times \aleph_\alpha] \geq \aleph_\alpha$, so we are done.

$\alpha \in \mathrm{Lim}$: We have

$$C[\aleph_\alpha \times \aleph_\alpha] = \bigcup_{\beta<\alpha} C[\aleph_\beta \times \aleph_\beta] = \bigcup_{\beta<\alpha} \aleph_\beta = \aleph_\alpha.$$

□

Theorem 4.30 (Hessenberg)**.** *If* κ, λ *are infinite cardinals then* $\kappa + \lambda = \kappa \cdot \lambda = \max\{\kappa, \lambda\}$.

Proof. As a consequence of Lemma 4.29, $\mu^2 = \mu \cdot \mu = \mu$ for every infinite cardinal as witnessed by the Mostowski collapse of $(\mu, <_G)$. Wlog suppose that $\kappa \leq \lambda$, so we have

$$\lambda \leq \kappa + \lambda \leq \kappa \cdot \lambda \leq \lambda^2 = \lambda,$$

so we actually have equalities across the board. □

Proving this theorem requires the full strength of the axiom of choice. Without AC, the maximum of two cardinals does not even make sense, as there may be two incompatible cardinals. For example, it is possible that $||$ and $\aleph_1$ are incompatible. On the other hand, a special case of Hessenberg's theorem gives back AC.

Theorem 4.31 (Tarski)**.** *The following are equivalent over* ZF*:*

(*i*) AC.

(*ii*) $\kappa^2 = \kappa$ *for every infinite cardinal* κ*.*

Corollary 4.32. *For any infinite cardinal* κ *and* $2 \leq \lambda \leq \kappa$ *we have*

$$2^\kappa = \lambda^\kappa = \kappa^\kappa = |\mathcal{P}(\kappa)|.$$

Proof. As ${}^\kappa 2 \subseteq {}^\kappa \lambda \subseteq {}^\kappa \kappa$, we have

$$2^\kappa \leq \lambda^\kappa \leq \kappa^\kappa.$$

Further,

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

where we apply Hessenberg's theorem in the last equality. Hence $2^\kappa = \lambda^\kappa = \kappa^\kappa$. Also $|\mathcal{P}(\kappa)| = 2^\kappa$ as taking a subset of κ to it's characteristic function constitutes a bijection between $\mathcal{P}(\kappa)$ and ${}^\kappa 2$. $\square$

Next, we introduce a central and ubiquitous Set-Theoretical concept, the cofinality.

Definition 4.33. Suppose $\alpha \in \mathrm{Lim}$.

(*i*) For any set X, a function $f\colon X \to \alpha$ is **cofinal** if $\sup \operatorname{ran}(f) = \alpha$.

(*ii*) The **cofinality** of α, $\operatorname{cof}(\alpha)$, is the least ordinal β so that there is a cofinal function $f\colon \beta \to \alpha$.

(*iii*) If $\operatorname{cof}(\alpha) = \alpha$, the ordinal α is **regular**, otherwise α is **singular**.

For example:

- ω is regular,
- the ordinal $\omega + \omega$ (in terms of ordinal arithmetic) is singular of cofinality ω as witnessed by $f(n) = \omega + n$.
- Similarly, the maps $n \mapsto \omega \cdot n$ and $n \mapsto \omega^n$ witness that $\omega \cdot \omega$ and ω^ω (again in terms of ordinal arithmetic) are singular of cofinality ω.
- The cardinal ω_1 is regular: if $\beta < \omega_1$ and $f\colon \beta \to \omega_1$ then $\bigcup_{i<\beta} f(i)$ is countable as a countable union over countable sets, so f cannot be cofinal.

We mentioned earlier that ZF does not prove that all countable union of countable sets are countable. Indeed, it is consistent with ZF that the cardinal ω_1 is singular.

Sometimes it is useful that we have cofinal functions $f\colon \operatorname{cof}(\alpha) \to \alpha$ with additional nice properties at hand.

Proposition 4.34. *Let $\alpha \in \mathrm{Lim}$. There is a strictly increasing continuous cofinal function $f\colon \operatorname{cof}(\alpha) \to \alpha$.*

Proof. Let $g\colon \operatorname{cof}(\alpha) \to \alpha$ be any cofinal function. We define $h\colon \operatorname{cof}(\alpha) \to \alpha$ via

$$h(\beta) = \sup g[\beta].$$

Note that $g \restriction \beta\colon \beta \to \alpha$ is not cofinal as $\beta < \operatorname{cof}(\alpha)$, hence the codomain of h is really α and h is certainly increasing and cofinal. We leave it to the reader to check that h is continuous. But h may fail to be strictly increasing.

To fix this, let $f\colon (\gamma, <) \xrightarrow{\sim} (\operatorname{ran}(g), <)$ be the anti-collapse map for some ordinal γ.

Claim 4.35. $\gamma = \operatorname{cof}(\alpha)$.

Proof. The map $\pi\colon (\gamma, <) \to (\operatorname{cof}(\alpha, <))$, $\pi(\beta) = \min g^{-1}(\{h(\beta)\})$ is an embedding, so we must have $\gamma \leq \operatorname{cof}(\alpha)$. On the other hand $\operatorname{ran}(f) = \operatorname{ran}(h)$ and hence f is cofinal, so $\operatorname{cof}(\alpha) \geq \gamma$. □

It follows that $f\colon \operatorname{cof}(\alpha) \to \alpha$ is strictly increasing and continuous as h is continuous. □

Lemma 4.36. *For any limit ordinal α, the ordinal $\operatorname{cof}(\alpha)$ is a regular cardinal $\leq\alpha$.*

Proof. $\operatorname{cof}(\alpha) \leq \alpha$ is obvious. Let $f\colon \operatorname{cof}(\alpha) \to \alpha$ be cofinal and let us also assume that f is increasing.

Let $\kappa = |\operatorname{cof}(\alpha)| \leq \operatorname{cof}(\alpha)$ and let $g\colon \kappa \to \operatorname{cof}(\alpha)$ be a bijection. Then

$$f \circ g\colon \kappa \to \alpha$$

is cofinal as $\operatorname{ran}(f \circ g) = \operatorname{ran}(f)$ so that $\operatorname{cof}(\alpha) \leq \kappa$ and hence $\operatorname{cof}(\alpha) = \kappa$.

Next suppose that $\beta \leq \operatorname{cof}(\alpha)$ and $h\colon \beta \to \alpha$ is cofinal. As f is increasing, $f \circ h\colon \beta \to \alpha$ is increasing and hence $\operatorname{cof}(\alpha) \leq \beta$ so that $\operatorname{cof}(\alpha)$ is regular. □

Definition 4.37. The class Reg is the class of all regular cardinals. The class Sing is the class of all singular ordinals and $\mathrm{SingCard} = \mathrm{Sing} \cap \mathrm{Card}$.

Definition 4.38. A cardinal κ is a **successor cardinal** if $\kappa = \lambda^+$ for some other cardinal λ. Otherwise, κ is a **limit cardinal**.

A cardinal $\aleph_\alpha$ is a successor/limit cardinal iff α is a successor/limit ordinal and

Lemma 4.39. *All (infinite) successor cardinals are regular.*

Proof. The argument is basically the same as the one which showed that ω_1 is regular. Suppose $\kappa = \lambda^+$ and $f\colon \operatorname{cof}(\kappa) \to \kappa$ is cofinal. Note that $|f(i) \leq \lambda$ for every $i < \operatorname{cof}(\kappa)$ and hence by AC, there is a sequence $\langle g_i \mid i < \operatorname{cof}(\kappa)\rangle$ so that $g_i\colon \lambda \to f(i)$ is surjective (we may assume that $f(i) \neq 0$).

The map $F\colon \lambda \cdot \operatorname{cof}(\kappa) \to \kappa$ given by

$$F(\alpha, i) = g_i(\alpha)$$

if surjective and hence $\kappa \leq \lambda \cdot \operatorname{cof}(\kappa) = \max\{\lambda, \operatorname{cof}(\kappa)\}$ and $\operatorname{cof}(\kappa) = \kappa$ follows. □

- $\aleph_0, \aleph_1, \aleph_2, \aleph_3, \dots$ are all regular.
- The limit of this sequence is $\aleph_\omega$ and hence $\aleph_\omega$ is singular of cofinality ω.
- Then, $\aleph_{\omega+1}, \aleph_{\omega+2}, \dots$ are all regular again and the limit of this sequence is the singular cardinal $\aleph_{\omega+\omega}$, again of cofinality ω.
- Eventually, we reach $\aleph_{\omega_1}$ which is singular, but of cofinality ω_1 as witnessed by $\alpha \mapsto \aleph_\alpha$.

The theory ZFC is not strong enough to prove the existence of another regular limit cardinal beyond ω. We will get back to this when we deal with "large cardinals".

We introduce transfinite cardinal arithmetic now. For a sequence $\langle X_i \mid i \in I\rangle$ the the product $\bigtimes_{i\in I} X_i$ is the set of all functions $f\colon I \to V$ with $f(i) \in X_i$ for all $i \in I$.

Definition 4.40. Suppose $\langle \kappa_i \mid i \in I\rangle$ is a sequence of cardinals. We define

$$\sum_{i\in I} \kappa_i := |\bigcup_{i\in I} X_i|,\ \prod_{i\in I} \kappa_i := |\bigtimes_{i\in I} X_i|$$

where $\langle X_i \mid i \in I\rangle$ is any sequence of pairwise disjoint sets with $|X_i| = \kappa_i$ (e.g. $X_i = \kappa_i \times \{i\}$).

This makes sense in the absence of choice as well, though is of limited use then. The axiom of choice is, basically by definition, equivalent to $\prod_{i\in I} \kappa_i \neq 0$ for all sequences of non-zero cardinals $\langle \kappa_i \mid i \in I\rangle$.

Lemma 4.41 (König)**.** *Suppose $\langle \kappa_i \mid i \in I\rangle, \langle \lambda_i \mid i \in I\rangle$ are sequences or cardinals with $\kappa_i < \lambda_i$ for all $i \in I$. Then*

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

Proof. The inequality $\leq$ is easy to see, so suppose that $f\colon \bigcup_{i\in I} \kappa_i \times \{i\} \to \bigtimes_{i\in I} \lambda_i$ is any function. We will show that f is not surjective. For any $i \in I$, note that

$$\{f(\alpha, i)(i) \mid \alpha < \kappa_i\} \subsetneq \lambda_i$$

as $\kappa_i < \lambda_i$. Let $\xi_i < \lambda_i$ be the minimal ordinal not in this set. Then the function mapping $i \in I$ to ξ_i is in $\prod_{i\in I} \lambda_i$ and is not in the range of f. □

Corollary 4.42. *Let κ be any infinite cardinal.*

(i) $\kappa^{\operatorname{cof}(\kappa)} > \kappa$.

(ii) $\operatorname{cof}(2^\kappa) > \kappa$.

Note that *(i)* is an improvement to Cantor's theorem (if κ is singular) as Cantor's result only states $\kappa < 2^\kappa = \kappa^\kappa$.

Proof. *(i)*: Let $f\colon \operatorname{cof}(\kappa) \to \kappa$ be cofinal. Applying Lemma 4.41 with $\kappa_i = |f(i)|$ and $\lambda_i = \kappa$, we see that

$$\kappa \leq \sum_{i<\operatorname{cof}(\kappa)} |f(i)| < \prod_{i<\operatorname{cof}(\kappa)} \kappa = |{}^{\operatorname{cof}(\kappa)}\kappa| = \kappa^{\operatorname{cof}(\kappa)}.$$

(ii): Let $g\colon \operatorname{cof}(2^\kappa) \to 2^\kappa$ be cofinal. Then

$$2^\kappa \leq \sum_{i<\operatorname{cof}(2^\kappa} |g(i)| < \prod_{i<\operatorname{cof}(2^\kappa)} 2^\kappa = (2^\kappa)^{\operatorname{cof}(2^\kappa)} = 2^{\kappa\cdot\operatorname{cof}(2^\kappa)} = 2^{\max\{\kappa,\operatorname{cof}(2^\kappa)\}}.$$

Hence we must have $\kappa < 2^{\operatorname{cof}(\kappa)}$. □

We have now proven two crucial things about the continuum: We know $\omega_1 \leq 2^\omega$ and $\operatorname{cof}(2^\omega)$ is uncountable. This are "the only restraints" on the continuum that ZFC can prove (in a way which can be made precise). Any concrete cardinal which has these properties can consistently with ZFC be the continuum. For example, it is consistent with ZFC that $2^\omega = \dots$

- $\aleph_1$,
- $\aleph_2$,
- $\aleph_{42}$,
- $\aleph_{\omega+1}$,
- $\aleph_{\aleph_{\omega_1}}$

but not $2^\omega = \aleph_\omega$.

Let $F\colon \mathrm{Card} \to \mathrm{Card}$ denote the continuum function $F(\kappa) = 2^\kappa$. We know that F is

(i) (weakly) increasing, i.e. $\kappa \leq \lambda$ implies $F(\kappa) \leq F(\lambda)$ and

(ii) $\kappa \leq \operatorname{cof}(F(\kappa))$ for all infinite cardinals κ.

Of course these properties also hold for the restriction $F \restriction \mathrm{Reg}$ of F to the regular cardinals. Once again, these are "the only restraints" that ZFC can prove about $F \restriction \mathrm{Reg}$, in the sense that any concrete function with these properties can consistently be $F \restriction \mathrm{Reg}$. For example, it is consistent that

- $F(\kappa) = \kappa^+$ for all $\kappa \in \mathrm{Reg}$,

- $F(\kappa) = \kappa^{++}$ for all $\kappa \in \mathrm{Reg}$,
- $F(\kappa) = \aleph_{\kappa+37}$ for all $\kappa \in \mathrm{Reg}$, etc.

It turns out, however, that $F \restriction \mathrm{Sing}$ is much **much** more complicated and ZFC can prove many more restraints on $F \restriction \mathrm{Sing}$ than the two above. For example, the following is a famous result of Shelah.

Theorem 4.43 (Shelah)**.** *Suppose that $2^{<\aleph_\omega} = \aleph_\omega$. Then $2^{\aleph_\omega} < \aleph_{\omega_4}$.*

The role of the number 4 in the above inequality is one of the big open problems of Set Theory. The 4 should really be a 1 and if that tighter bound could be proven, we know that it would be optimal, but no one knows how to do that.

We will next prove a simple result about $F \restriction \mathrm{Sing}$ that does not hold for $F \restriction \mathrm{Reg}$.

Definition 4.44. Suppose κ, λ are cardinals and κ is infinite. Then

$$\lambda^{<\kappa} = \sup_{\alpha<\kappa} \lambda^{|\alpha|}.$$

Note that if $2 \leq \lambda$ and κ is infinite then $\lambda^{<\kappa}$ is the size of $\bigcup_{\alpha<\kappa} {}^{\alpha}\lambda$. It is clear that $\lambda^{<\kappa} \leq |\bigcup_{\alpha<\kappa} {}^{\alpha}\lambda|$ and on the other hand,

$$|\bigcup_{\alpha<\kappa} {}^{\alpha}\lambda| = \sum_{\alpha<\kappa} \lambda^{|\alpha|} \leq \sum_{\alpha<\kappa} \lambda^{<\kappa} = \kappa \cdot \lambda^{<\kappa} = \lambda^{<\kappa}.$$

The last equality holds as $2 \leq \lambda$ implies $\lambda^{|\alpha|} \geq \alpha$ for all α so that $\kappa \leq \lambda^{<\kappa}$.

Lemma 4.45. *Suppose κ is an infinite cardinal. Then $(2^{<\kappa})^{\mathrm{cof}(\kappa)} = 2^\kappa$.*

Proof. Note that ${}^{\mathrm{cof}(\kappa)}(\bigcup_{\alpha<\kappa} {}^{\alpha}2)$ has size precisely $(2^{<\kappa})^{\mathrm{cof}(\kappa)}$. Moreover, we can define an injection

$$f \colon {}^{\kappa}2 \to {}^{\mathrm{cof}(\kappa)}(\bigcup_{\alpha<\kappa} {}^{\alpha}2)$$

via

$$f(g) = \langle g \restriction h(i) \mid i < \mathrm{cof}(\kappa)\rangle$$

where h is any fixed cofinal function $h \colon \mathrm{cof}(\kappa) \to \kappa$. Hence

$$2^\kappa \leq (2^{<\kappa})^{\mathrm{cof}(\kappa)} \leq (2^\kappa)^{\mathrm{cof}(\kappa)} = 2^{\kappa\cdot\mathrm{cof}(\kappa)} = 2^\kappa.$$

□

Corollary 4.46. *Suppose κ is a singular cardinal and there is a cardinal $\lambda < \kappa$ so that $2^\lambda = 2^{<\kappa}$. Then $2^\kappa = 2^\lambda$.*

Proof. We calculate

$$2^\kappa = (2^{<\kappa})^{\mathrm{cof}(\kappa)} = (2^\lambda)^{\mathrm{cof}(\kappa)} = 2^{\lambda\cdot\mathrm{cof}(\kappa)} \leq 2^{<\kappa} = 2^\lambda.$$

Note that $\lambda \cdot \mathrm{cof}(\kappa) < \kappa$ as κ is singular. □

On the other hand, this fails badly for regular cardinals. For example $2^{<\omega_1} = 2^\omega$ but 2^{ω_1} can be any cardinal above 2^ω of cofinality at least ω_2.

We know prove a useful equality known as **Hausdorff's Formula**.

Lemma 4.47 (Hausdorff). *Suppose that κ, λ are infinite cardinals. Then*

$$(\kappa^+)^\lambda = \kappa^\lambda \cdot \kappa^+.$$

Proof. We split into two cases.

Case 1: $\kappa^+ \leq \lambda$. Then $(\kappa^+)^\lambda = 2^\lambda = \kappa^\lambda = \kappa^\lambda \cdot \kappa^+$.

Case 2: $\kappa^+ > \lambda$. Then, since κ^+ is regular,

$$(\kappa^+)^\lambda = |\bigcup_{\alpha<\kappa^+} {}^\lambda\alpha| = \bigcup_{\alpha<\kappa^+} |\alpha|^\lambda \leq \sum_{\alpha<\kappa^+} \kappa^\lambda = \kappa^\lambda \cdot \kappa^+.$$

□

10.04.24

5 Clubs and Stationary Sets

We now move to the part of Set Theory which is closest to Measure or Probability Theory. We will isolate a notion of "big" and "small" subsets of uncountable cardinals. Unfortunately, almost all of these cardinals are "too large" to support a useful measure, so we will only distinguish between sets of "full measure", sets of "measure zero" and sets of "positive measure".

5.1 Closed unbounded sets

Definition 5.1. Let α be a limit ordinal and $X \subseteq \alpha$ a subset.

(*i*) X is **unbounded** in α if $\sup X = \alpha$. (This is equivalent to $\forall \beta < \alpha \exists \gamma \in X\ \beta < \gamma$).

(*ii*) The set of **limit points**[4] **of** X below α is

$$\mathrm{Lim}(X) = \{\beta < \alpha \mid \sup(X \cap \beta) = \beta\}.$$

(*iii*) X is **closed** in α if X is a closed set in the the order-topology on $(\alpha, <)$. This is equivalent to

$$\forall \beta < \alpha\ (\sup(X \cap \beta) = \beta \to \beta \in X).$$

(*iv*) X is **club**[5] in α if X is both closed and unbounded in α.

[4] It is customary to suppress the ordinal α in the notation. This is an example of bad notation.

[5] It may be obvious to you, but I know of people for which it wasn't: club is short for **cl**osed **unb**ounded.

Usually, the ordinal α is clear from context, so we merely write "X is unbounded/closed" or "X is a club".

Trivial examples of clubs are the intervals $[\beta, \alpha) = \{\gamma < \alpha \mid \beta \leq \gamma\}$. Note that, e.g. $\omega+1$ is closed in $\omega+\omega$ but not unbounded. $\mathrm{Succ} \cap \omega+\omega$ is unbounded in $\omega+\omega$ but not closed and $\omega \subseteq \omega+\omega$ is neither.

Typical examples of clubs arise from the following.

Lemma 5.2. *Suppose κ is a regular uncountable cardinal and $f\colon \kappa \to \kappa$ is a function. Then*

$$C_f = \{\alpha < \kappa \mid f[\alpha] \subseteq \alpha\}$$

is a club in κ.

The elements of C_f are closed under f and we call them closure points of f.

Proof. This argument is a typical "catching up" argument found in Set Theory. We describe some kind of process which seems to run away, but we catch up "at infinity".

First, it is quite easy to see that C_f is closed. If $\alpha < \kappa$ is a limit of closure points and $\beta < \alpha$, we may find a closure point $\beta < \gamma < \alpha$ so that $f(\beta) \in f[\gamma] \subseteq \gamma$ and hence $f(\beta) < \alpha$.

Now let us show that C_f is unbounded. Let $\alpha_0 < \kappa$. If α_n is defined, we set

$$\alpha_{n+1} = \max(\alpha_n, \sup f[\alpha_n]) + 1.$$

As κ is regular, $f \upharpoonright \alpha_n$ cannot be cofinal in κ and hence $\sup f[\alpha_n] < \kappa$. We have thus described a strictly increasing sequence $(\alpha_n)_{n<\omega}$ of ordinals below κ. Since κ is uncountable and regular, $\alpha_\omega \coloneqq \sup_{n<\omega} \alpha_n < \kappa$.

Claim 5.3. $\alpha_\omega \in C_f$.

Proof. If $\beta < \alpha_\omega$ then $\beta < \alpha_n$ for some $n < \omega$. But then $f(\beta) \leq \sup f[\alpha_n] < \alpha_{n+1} \leq \alpha_\omega$. □

As $\alpha_\omega > \alpha_0$, this shows that C_f is indeed unbounded. □

Note that we only use that κ is regular to see that $f \upharpoonright \alpha$ is not cofinal for any $\beta < \kappa$. The rest of the argument only needs that κ has uncountable cofinality. So, for example, we get:

Lemma 5.4. *Suppose $\alpha \in \mathrm{Lim}$ and $\mathrm{cof}(\alpha)$ is uncountable. If $f\colon \alpha \to \alpha$ is increasing then C_f is a club.*

Another useful example of clubs, which is really just a special case of Lemma 5.4, are sets of limit points of unbounded sets.

Corollary 5.5. *Suppose $\alpha \in \mathrm{Lim}$ has uncountable cofinality and $X \subseteq \alpha$ is unbounded. Then $\mathrm{Lim}(X)$ is a club in α.*

Proof. Define $f\colon \alpha \to \alpha$ via $f(\beta) = \min X \setminus \beta$. It is not hard to see that $\mathrm{Lim}(X) = C_f$. □

Intersections along big sets should be big and we will show this next.

Lemma 5.6. *Suppose α is a limit ordinal of uncountable cofinality and $\langle C_\beta \mid \beta < \gamma\rangle$ is a sequence of clubs in α of length $\gamma < \operatorname{cof}(\alpha)$. Then $\bigcap_{\beta<\gamma} C_\beta$ is a club.*

Proof. The interection of closed sets is obviously closed, so we only show unboundedness. For each $\beta < \alpha$, let $f_\beta \colon \alpha \to \alpha$ be defined via

$$f_\beta(\delta) = \min C_\beta \setminus \delta.$$

Next, let $f_* \alpha \to \alpha$ be the supremum of the f_beta's, i.e. $f_*(\delta) = \sup_{\beta<\gamma} f_\beta(\delta)$. Since $\gamma < \operatorname{cof}(\alpha)$, we have $f_*(\delta) < \alpha$.

The function f_* is clearly increasing so that C_{f_*} is a club in α and $C_{f_*} \subseteq C_{f_\beta}$ for all $\beta < \gamma$. Moreover, since all C_β are closed, we have $C_{f_\beta} \subseteq C_\beta$ and hence $C_{f_*} \subseteq C_\beta$. It follows that

$$C_{f_*} \subseteq \bigcap_{\beta<\gamma} C_\beta$$

and since C_{f_*} is unbounded, we are done. □

5.2 Stationary sets and Fodor's lemma

In Measure Theory, a countable intersection of sets of full measure is still a set of full measure and any superset of a full measure set is of full measure. A set has positive measure iff it meets every set of full measure. We can thus draw the following analogies.

Contains a club	meets every club	disjoint from some club
Full measure	positive measure	measure zero

Definition 5.7. Suppose α is a limit ordinal of uncountable cofinality. A set $X \subseteq \alpha$ is **stationary** (in α) iff $X \cap C \neq \emptyset$ for every club set C. Otherwise, X is **nonstationary**.

For limit ordinals of countable cofinality, there are always two clubs disjoint from one another so the above analogy breaks down.

On regular cardinals, we can slightly improve Lemma 5.6.

Definition 5.8. Suppose α is a limit ordinal and $\langle X_\beta \mid \beta < \alpha$ is a sequence of subsets of α of length α. The set

$$\triangle_{\beta<\alpha} X_\beta \coloneqq \{\delta < \alpha \mid \delta \in \bigcap_{\beta<\delta} X_\delta\}$$

is the **diagonal intersection** along $\langle X_\beta \mid \beta < \alpha\rangle$.

Theorem 5.9. *Suppose κ is a regular uncountable cardinal and $\langle C_\alpha \mid \alpha < \kappa\rangle$ is a sequence of clubs on κ. Then $\triangle_{\alpha<\kappa} C_\alpha$ is a club.*

Proof. First, let us see that $\triangle_{\alpha<\kappa}C_\alpha$ is closed. If $\beta < \kappa$ is a limit point of $\triangle_{\alpha<\kappa}C_\alpha$ and $\alpha < \beta$ then $[\alpha+1,\beta) \cap \triangle_{\alpha<\kappa}C_\alpha \subseteq C_\alpha \cap \beta$ so β is a limit point of C_α and hence $\beta \in C_\alpha$ as C_α is closed.

Next, we show that $\triangle_{\alpha<\kappa}C_\alpha$ is unbounded. The proof is almost the same as the proof of Lemma 5.6. For each $\alpha < \kappa$, let $f_\alpha \colon \kappa \to \kappa$ be defined via $f_\alpha(\beta) = \min C_\alpha \setminus \beta$. Then, set

$$f_*(\beta) = \sup_{\alpha<\beta} f_\alpha(\beta)$$

for $\beta < \kappa$ and note that this results in a function $f_* \colon \kappa \to \kappa$ since κ is regular. As before, we see that $C_{f_*} \subseteq \triangle_{\alpha<\kappa}C_\alpha$ and as C_{f_*} is unbounded, $\triangle_{\alpha<\kappa}C_\alpha$ is, too. $\square$

This slight improvement leads to a new tool in our toolbox known as **Fodor's Lemma**.

Definition 5.10. For an ordinal α, function $f \colon \alpha \to \alpha$ is **regressive** if $f(\beta) < \beta$ for all $0 < \beta < \alpha$.

Lemma 5.11 (Fodor)**.** *Suppose κ is a regular uncountable cardinal, $S \subseteq \kappa$ is stationary and $f \colon \kappa \to \kappa$ is regressive. Then there is some stationary $T \subseteq S$ so that $f \restriction T$ is constant.*

Proof. Suppose not. For each $\alpha < \kappa$, let C_β be a club disjoint from $f^{-1}(\alpha)$. Hence $\triangle_{\alpha<\kappa}C_\alpha$ is a club and as S is stationary, there is some $\beta \in S \cap \triangle_{\alpha<\kappa}C_\alpha$. As f is regressive, $\alpha := f(\beta) < \beta$ and hence $\beta \in C_\alpha$. But C_α was supposed to be disjoint from $f^{-1}(\alpha)$, contradiction. $\square$

5.3 Solovay's splitting theorem

In any reasonable measure space, any set of positive measure can be split into two (or even countably many) disjoint sets of positive measure. As a warm-up, let us give an example of disjoint stationary sets.

Definition 5.12. Suppose λ is regular and $\lambda < \kappa$ is a limit ordinal. The set

$$E^\kappa_\lambda = \{\alpha < \kappa \mid \operatorname{cof}(\alpha) = \lambda\}$$

is the set of cofinality λ ordinals below κ.

Proposition 5.13. *If λ is a regular cardinal and $\lambda < \kappa$ is has cofinality $> \lambda$ then E^κ_λ is stationary in κ.*

Proof. Exercise. $\square$

So for example $E^{\omega_2}_\omega$ and $E^{\omega_2}_{\omega_1}$ are two disjoint stationary subsets of ω_2. Of course, this does not help with finding disjoint stationary subsets of ω_1 and this is indeed a little bit tricky to do. In fact, we need to make use of the axiom of choice again: It is consistent with ZF that ω_1 is regular, yet every subset of ω_1 either contains or is disjoint from some club.

The result that nonetheless in ZFC, any stationary set can be split into many disjoint stationary sets is known as **Solovay's Splitting Lemma**.

Theorem 5.14 (Solovay)**.** *Suppose κ is regular uncountable and $S \subseteq \kappa$ is stationary. Then there is a sequence $\langle S_i \mid i < \kappa\rangle$ of κ-many pairwise disjoint stationary sets so that*

$$S = \bigcup_{i<\kappa} S_i.$$

16.04.24

The proof splits into two cases and in the more difficult case, the stationary set concentrates on regular cardinals. Let us introduce some tools to deal with this case.

Definition 5.15. Suppose κ is a limit ordinal of uncountable cofinality and $S \subseteq \kappa$ is stationary. The **trace** of S is

$$\mathrm{Tr}(S) = \{\alpha < \kappa \mid \mathrm{cof}(\alpha) > \omega \wedge S \cap \alpha \text{ is stationary in } \alpha\}.$$

Lemma 5.16. *Suppose that κ is a regular uncountable cardinal and $S \subseteq \kappa$ is stationary. Then $S \setminus \mathrm{Tr}(S)$ is stationary in κ.*

Proof. Suppose not. Then there is a club $C \subseteq \kappa$ disjoint from $S \setminus \mathrm{Tr}(S)$. The set $\mathrm{Lim}(C)$ is another club in κ and hence meets S, so let $\alpha = \min \mathrm{Lim}(C) \cap S$.

As $\alpha \in \mathrm{Lim}(C)$, it is easy to see that $C \cap \alpha \subseteq \alpha$ is a club. It follows that $\mathrm{Lim}(C \cap \alpha) = \mathrm{Lim}(C) \cap \alpha \subseteq \alpha$ is a club as well. But, by choice of α, $\mathrm{Lim}(C) \cap \alpha \cap S = \emptyset$, contradiction. □

Proof of Theorem 5.14. Case 1: $S' = \{\alpha \in S \mid \alpha \text{ is singular}\}$ is stationary. Then $\mathrm{cof}\colon S' \to \kappa$ is a regressive function and by Fodor's lemma, there is some $S'' \subseteq S'$ stationary and $\lambda < \kappa$ so that $\mathrm{cof}(\alpha) = \lambda$ for all $\alpha \in S''$. Using the axiom of choice, we find a sequence $\langle (c^\alpha_\xi)_{\xi<\lambda} \mid \alpha \in S''\rangle$ so that $(c^\alpha_\xi)_{\xi<\lambda}$ is a increasing sequence cofinal in α.

Claim 5.17. *There is some $\xi < \kappa$ so that*

$$\{\alpha \in S'' \mid c^\alpha_\xi \geq \beta\}$$

is stationary for all $\beta < \kappa$.

Proof. Suppose not. Then for each $\xi < \kappa$ there is some $\beta_\xi < \kappa$ and a club $C_\xi \subseteq \kappa$ so that $c^\alpha_\xi < \beta_\xi$ for all $\alpha \in C_\xi \cap S''$. By Lemma 5.6,

$$C_* = \bigcup_{\xi<\lambda} C_\xi$$

is a club. Let $\beta_* = \sup_{\xi<\lambda} \beta_\xi$ and note that β_* since κ is regular. The club C_* has unbounded intersection with S'', so take some $\alpha \in S'' \cap C_*$ with $\alpha > \beta_*$. But then $c^\alpha_\xi < \beta_\xi$ for all $\xi < \lambda$ so that $\sup_{\xi<\lambda} c^\alpha_\xi \leq \beta_* < \alpha$, contradiction. □

Let ξ be as in the claim. For $\beta < \kappa$ let $S''_\beta = \{\alpha \in S'' \mid c^\alpha_\xi = \beta\}$ and note that $S''_\beta \cap S''_\gamma = \emptyset$ for different $\beta < \gamma$. It remains to show that

$$X = \{\beta < \kappa \mid S''_\beta \text{ is stationary}\}$$

has size κ as then $\langle S''_\beta \mid \beta \in X\rangle$ is a sequence of disjoint stationary subsets of S (these S''_β may not union of to the full S, but we can put the remaining ordinals in one the sets).

As κ is regular, it is enough to show that $X \subseteq \kappa$ is unbounded. If $\beta < \kappa$ then $S''_{\geq\beta} := \{\alpha \in S'' \mid c^\alpha_\xi \geq \beta\}$ is stationary by choice of ξ. The map

$$f\colon S''_{\geq\beta} \to \kappa,\ \alpha \mapsto c^\alpha_\xi$$

is clearly regressive and hence constant on some stationary subsets of $S''_{\geq\beta}$ with value γ. But then $\gamma \in X$ and $\gamma \geq \beta$.

Case 2: $S' = \{\alpha \in S \mid \alpha \text{ is regular}\}$ is stationary. By Lemma 5.16, $S'' = S' \setminus \mathrm{Tr}(S')$ is stationary. Hence, using AC, we may find a sequence

$$\langle (c^\alpha_\xi)_{\xi<\alpha} \mid \alpha \in S''\rangle$$

so that $(c^\alpha_\xi)_{\xi<\alpha}$ is the increasing enumeration of a club $D_\alpha \subseteq \alpha$ disjoint from S'. Equivalently, $(c^\alpha_\xi)_{\xi<\alpha}$ is strictly increasing, continuous and cofinal in α with values in $\alpha \setminus S'$.

Claim 5.18. *There is a $\xi < \kappa$ so that*

$$\{\alpha < \kappa \mid \xi < \alpha \wedge c^\alpha_\xi \geq \beta\}$$

is stationary for every $\beta < \kappa$.

Proof. Suppose not. Then for every $\xi < \kappa$ there is some $\beta_\xi < \kappa$ and a club $C_\xi \subseteq \kappa$ so that $c^\alpha_\xi < \beta_\xi$ for every $\alpha \in C_\xi \cap S''$. By Theorem 5.9,

$$C_* := \triangle_{\xi<\kappa} C_\xi$$

is a club. Let $f\colon \kappa \to \kappa$ be function $\xi \mapsto \beta_\xi$.

We also know that $C_f = \{\alpha < \kappa \mid f[\alpha] \subseteq \alpha\}$ is a club.

As S'' is stationary, take $\alpha \in S'' \cap C_f$ and let $\beta \in S'' \cap C_*$ with $\alpha < \beta$. Then for all $\xi < \alpha$, $c^\beta_\xi < \beta_\xi < \alpha$ so that

$$c^\beta_\alpha = \sup_{\xi<\alpha} c^\beta_\xi \leq \alpha.$$

On the other hand, $\alpha \leq c^\beta_\alpha$ since $(c^\beta_\xi)_{\xi<\beta}$ is strictly increasing. But then $c^\beta_\alpha = \alpha \in S''$, contradiction. □

We can now proceed exactly as in Case 1: take ξ as in the Claim and let $S''_\beta = \{\alpha \in S'' \mid \xi < \alpha \wedge c^\alpha_\xi = \beta\}$ for $\beta < \kappa$. As before, $X = \{\beta < \kappa \mid S''_\beta$ is stationary$\}$ has size κ by an application of Fodor's lemma and hence $\langle S''_\beta \mid \beta \in X\rangle$ works.

These two cases are exhaustive as

$$S = \{\alpha \in S \mid \alpha \text{ is singular}\} \cup \{\alpha \in S \mid \alpha \text{ is regular}\} \cup (S \cap \mathrm{Succ})$$

and $S \cap \mathrm{Succ}$ is nonstationary as it is disjoint from the club $S \cap \mathrm{Lim}$. $\square$

The structure of the stationary subsets of an uncountable regular cardinal κ is very interesting and complicated, even for $\kappa = \omega_1$. The theory ZFC leaves a lot of this structure undecided. We give an example.

Definition 5.19. Let κ be a regular uncountable cardinal. An **antichain of stationary subsets of** κ is a set $\mathcal{A}$ of stationary subsets of κ so that $S \cap T$ is nonstationary for $S \neq T$ both in $\mathcal{A}$.

How big can antichains of stationary subsets of ω_1 get? By Solovay's splitting theorem, there are such antichains of size ω_1. It turns out that ZFC does not decide whether there are such antichains of size ω_2 and the absence of such large antichains has some very interesting consequences.

5.4 Silver's theorem

We already remarked that the continuum function $\kappa \mapsto 2^\kappa$ is much more complicated on the singular cardinals. We will now give a proof of Silver's theorem, the first non-trivial result about cardinal arithmetic on singular cardinals which fails for regular cardinals.

Silver's theorem is about the first cardinal at which GCH fails (if it even exists), i.e. the least infinite cardinal κ such that $2^\kappa > \kappa^+$. Consistently, this cardinal can be

- $\aleph_1$ (if CH fails), $\aleph_2$, $\aleph_{42}$ and any other $\aleph_n$ for finite n,
- $\aleph_\omega$ (but this is much much more difficult to arrange)
- $\aleph_{\omega+1}$ or any $\aleph_{\omega+n}$ for finite n
- $\aleph_{\omega+\omega}$ (which is once again quite difficult), etc.

However, it **cannot** be $\aleph_{\omega_1}$.

Theorem 5.20 (Silver)**.** *Suppose κ is a singular cardinal of uncountable cofinality. If* GCH *holds below κ, i.e. $2^\lambda = \lambda^+$ for every infinite cardinal $\lambda < \kappa$, then $2^\kappa = \kappa^+$.*

Proof. Fix a strictly increasing continuous sequence $(c_\alpha)_{\alpha<\kappa}$ cofinal in κ. We may assume that every c_α is infinite and even a cardinal. We will identify every subset $X \subseteq \kappa$ with a function f_X in $\bigtimes_{\alpha<\mathrm{cof}(\kappa)} c_\alpha^+$. The idea of how to do that is similar to the one in the proof of Lemma 4.45. We map a subset $X \subseteq \kappa$ to

$$X \mapsto \langle X \cap c_\alpha \mid \alpha < \mathrm{cof}(\kappa)\rangle \in \bigtimes_{\alpha<\mathrm{cof}(\kappa)} \mathcal{P}(c_\alpha) \cong \bigtimes_{\alpha<\mathrm{cof}(\kappa)} c_\alpha^+.$$

where the last bijection comes from identifying each $\mathcal{P}(c_\alpha)$ with c_α^+ which is possible as GCH holds below κ. The function f_X is then the image of

$$\langle X \cap c_\alpha \mid \alpha < \mathrm{cof}(\kappa)\rangle$$

under this bijection.

The set $\mathcal{F} = \{f_X \mid X \subseteq \kappa\}$ has a special property: it is **almost disjoint**. This means that if $X \neq Y$ then for some $\beta < \mathrm{cof}(\kappa)$, $f_X(\alpha) \neq f_Y(\alpha)$ for all $\beta \leq \alpha < \mathrm{cof}(\kappa)$.

Claim 5.21. *Suppose $\mathcal{G} \subseteq \bigtimes_{\alpha<\mathrm{cof}(\kappa)} A_\alpha$ is a almost disjoint set and $S \coloneqq \{\alpha < \mathrm{cof}(\kappa) \mid |A_\alpha| \leq c_\alpha\}$ is stationary. Then $|\mathcal{G}| \leq \kappa$.*

Proof. Wlog we may assume that $A_\alpha = c_\alpha$ for $\alpha \in S$. For every $g \in \mathcal{G}$, the function

$$h_g \colon S \cap \mathrm{Lim} \to \mathrm{cof}(\kappa)\ \alpha \mapsto \min\{\beta < \alpha \mid g(\alpha) < c_\beta\}$$

is regressive. By Fodor's Lemma, we can choose a stationary $S_g \subseteq S$ and $\beta_g < \mathrm{cof}(\kappa)$ so that $h_g \restriction S_g$ is constant with value β_g. If $g \restriction S_g = g' \restriction S_{g'}$ for $g, g' \in \mathcal{G}$ then $g = g'$ since $\mathcal{G}$ is almost disjoint and $S_g = S_{g'}$ is unbounded in κ.

Crucially, the range of $g \restriction S_g$ is bounded in κ by c_{β_g}. The bounded functions $S_g \to \kappa$ have size at most

$$|\bigcup_{\alpha<\kappa} {}^{S_g}\alpha| = \sum_{\alpha<\kappa} |\alpha|^{|S_g|} = \sum_{\alpha<\kappa} |\alpha|^{\mathrm{cof}(\kappa)} \leq \sum_{\alpha<\kappa} 2^{|\alpha|} = \sum_{\alpha<\kappa} \alpha^+ \leq \sum_{\alpha<\kappa} \kappa = \kappa \cdot \kappa = \kappa$$

where we once again use GCH below κ (and that successor cardinals are regular).

We can now calculate

$$\begin{aligned} |\mathcal{G}| &\leq |\bigcup_{T \subseteq \mathrm{cof}(\kappa)} \{g \restriction T \mid g \in \mathcal{G} \wedge S_g = T\}| \leq \sum_{T \subseteq \mathrm{cof}(\kappa)} |\bigcup_{\alpha<\kappa} {}^{T}\alpha| \\ &\leq \sum_{T \subseteq \mathrm{cof}(\kappa)} \kappa = |\mathcal{P}(\mathrm{cof}(\kappa))| \cdot \kappa = \mathrm{cof}(\kappa)^+ \cdot \kappa = \kappa. \end{aligned}$$

Here, we use that κ is singular and that GCH holds below κ. □

For $X, Y \subseteq \kappa$, let us write $f_X \leq f_Y$ in case $\{\alpha < \mathrm{cof}(\kappa) \mid f_X(\alpha) \leq f_Y(\alpha)\}$ is stationary. Note that either $f_X \leq f_Y$ or $f_Y \leq f_X$ or both. Let us define

$$\mathcal{F}_X = \{f_Y \in \mathcal{F} \mid f_Y \leq f_X\}.$$

Claim 5.22. $|\mathcal{F}_X| \leq \kappa$.

Proof. We have that

$$\mathcal{F}_X = \bigcup_{S\subseteq\kappa \text{ stationary}} \underbrace{\{f_Y \in \mathcal{F} \mid \{\alpha < \operatorname{cof}(\kappa) \mid f_Y(\alpha) \leq f_X(\alpha)\} = S\}}_{=:\mathcal{F}_{X,S}}$$

and note that it follows from our first claim that $\mathcal{F}_{X,S}$ is of size $\leq \kappa$. Hence we can calculate

$$|\mathcal{F}_X| \leq \bigcup_{S\subseteq\kappa \text{ stationary}} |F_{X,S}| \leq |\mathcal{P}(\operatorname{cof}(\kappa))| \cdot \kappa = \operatorname{cof}(\kappa)^+ \cdot \kappa = \kappa.$$

□

Now define a sequence $\langle X_i \mid i < \delta\rangle$ recursively so that

$$f_{X_i} \notin \bigcup_{j<i} \mathcal{F}_{X_j}$$

for all $i < \delta$ for as long as possible. We must have that $\delta \leq \kappa^+$: otherwise X_{κ^+} is defined and it follows that $f_{X_i} \in \mathcal{F}_{X_{\kappa^+}}$ for all $i < \kappa^+$. But $\mathcal{F}_{X_{\kappa^+}}$ has size at most κ, contradiction.

Finally, we have

$$2^\kappa = |\mathcal{F}| = |\bigcup_{i<\delta} \mathcal{F}_{X_i}| \leq \sum_{i<\delta} |\mathcal{F}_{X_i}| \leq \sum_{i<\delta} \kappa = |\delta| \cdot \kappa \leq \kappa^+ \cdot \kappa = \kappa^+.$$

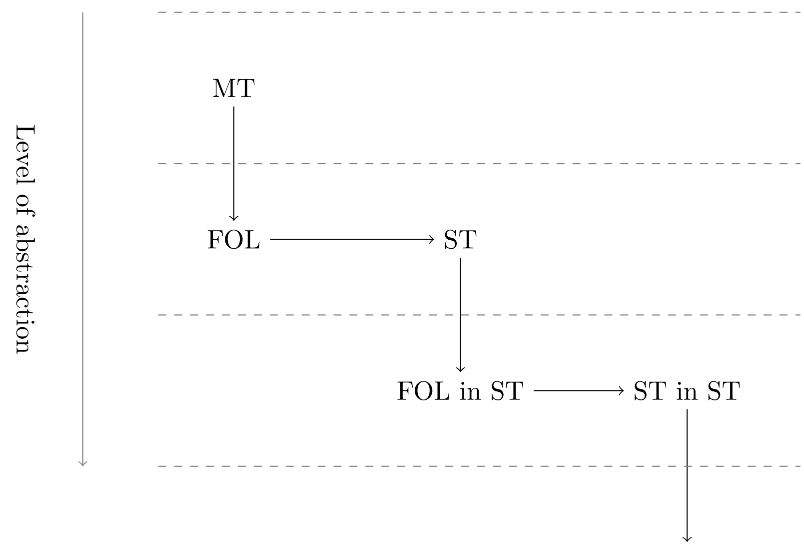
□

Silver's theorem is purely about cardinal arithmetic and superficially appears to not have any connection to clubs or stationary set. Nonetheless, the notion of stationary sets and Fodor's lemma featured prominently in the argument. This is typical for set theory, stationary sets and Fodor's lemma can be extremely powerful in many circumstances. Always look out for such potential applications.

6 First Order Logic in Set Theory

The road we took to end up with a formalization of Set Theory is as follows:

- As every bit of mathematic does, we start in an informal "naive" framework of mathematics (one has to start somewhere!). This is referred to as the **Metatheory**.
- Then first order logic is formalized inside the metatheory (which we have taken for granted).
- Set Theory is then the study of specific first order theories, in our case ZF and ZFC. (This is where these lecture notes start).



A depiction of the path of abstraction. MT is short for metatheroy, FOL for first order logic and ST for Set Theory.

Now something interesting happens: As we already mentioned, Set Theory can be used as a formal framework for the whole rest of mathematics as a rigorous substitute for the metatheory. This includes first order logic!

One can map "concrete objects" such as natural numbers down along levels of abstraction. This is known as **Gödelization** and is typically denoted by $x \mapsto \ulcorner x \urcorner$. For example, there is the number 0 of the metatheory, which you have seen and used in any other mathematics course. In terms of Set Theory, we have defined a zero as the empty set, so $\ulcorner 0 \urcorner$ is the (term for the) empty set, and $\ulcorner 1 \urcorner$ is the (term for the) set $\{\ulcorner 0 \urcorner\}$. This can be continued along the natural numbers.

If $\mathcal{M} = (M, \in^{\mathcal{M}})$ is a concrete model of ZFC in our metatheory, then these terms evaluate to concrete members of M. For example, $\ulcorner 0 \urcorner^{\mathcal{M}}$ is the unique x in M so that

$$\mathcal{M} \models x = \emptyset.$$

Thus we get a map $\mathbb{N}^{\mathrm{MT}} \to \omega^{\mathcal{M}}$ which maps any natural number n of the metatheory to its version $\ulcorner n \urcorner^{\mathcal{M}}$ in $\mathcal{M}$.

Usually, it is not possible in a reasonable way do go back up. For example the map $n \mapsto \ulcorner n \urcorner^{\mathcal{M}}$ may not be surjective! Using the completeness theorem of first order logic, a model of ZFC with nonstandard natural numbers can be constructed in the same way as a nonstandard model of Peano arithmetic.

Let us now shortly describe the step from the second level to the third, i.e. how to formalize first order logic inside of Set Theory. A language is defined as in the metatheory: an arbitrary set $\mathcal{L}$, each element of which is called a symbol and is designated either as a "relation" or a "function" together with a function arity: $\mathcal{L} \to \omega$ which assigns symbols their arity.

A first order $\mathcal{L}$-structure is then a tuple $\mathcal{N} = (N, (s^{\mathcal{N}})_{s\in\mathcal{L}})$ where $s^{\mathcal{N}}$ is a subset of $N^{\mathrm{arity}(s)}$ if s is a relation symbol or a function $N^{\mathrm{arity}(s)} \to N$ if s is a function symbol.

The first order $\mathcal{L}$-formulas are also defined as usual, but with a specific encoding as sets (which is usually swept under the rug if working in the metatheory). For simplicity, let us work with the $\in$-language. We could then define

- $v_i = v_j \coloneqq (0, 0, i, j)$,
- $v_i \in v_j \coloneqq (0, 1, i, j)$,
- $\neg\varphi \coloneqq (1, \varphi)$,
- $\varphi \wedge \psi \coloneqq (2, \varphi, \psi)$,
- $\exists v_i \varphi \coloneqq (3, \varphi, i)$,

by recursion, where $i, j < \omega$. This leads to a set $\mathrm{Fml}_\in$ of all $\in$-formulas. As usual, each such formula has an associated finite set free variables which code as the set of $i < \omega$ such that "v_i appears free in φ".

Once again, if $\mathcal{M}$ is a model of ZFC in the metatheory, then the map

$$\mathrm{Fml}_\in^{\mathrm{MT}} \to \mathrm{Fml}_\in^{\mathcal{M}}$$

which sends a metatheory $\in$-formula φ to its Gödelization $\ulcorner\varphi\urcorner^{\mathcal{M}}$ may not be surjective.

Working inside Set Theory again, if $\mathcal{N} = (N, E)$ is a $\in$-structure, we can define a partial function

$$\mathrm{Sat}_{\mathcal{N}} \colon N^{<\omega} \times \mathrm{Fml}_\in \to 2$$

by recursion so that $(a, \varphi) \in \mathrm{dom}(\mathrm{Sat})$ if $i \in \mathrm{dom}(a)$ for all $i \in \mathrm{free}(\varphi)$ and in this case

- $\mathrm{Sat}_{\mathcal{N}}(a, v_i = v_j) = 1$ iff $a_i = a_j$,
- $\mathrm{Sat}_{\mathcal{N}}(a, v_i \in v_j) = 1$ iff $(a_i, a_j) \in E$,
- $\mathrm{Sat}_{\mathcal{N}}(a, \neg\varphi) = 1$ iff $\mathrm{Sat}(a, \varphi) = 0$,
- $\mathrm{Sat}_{\mathcal{N}}(a, \varphi \wedge \psi) = 1$ iff $\mathrm{Sat}_{\mathcal{N}}(a, \varphi) \cdot \mathrm{Sat}(a, \psi) = 1$ and
- $\mathrm{Sat}_{\mathcal{N}}(a, \exists v_i \varphi) = 1$ iff $\exists x \in N$ $\mathrm{Sat}_{\mathcal{N}}(a^i_x, \varphi) = 1$ where $a^i_x(i) = x$ and $a^i_x \restriction \omega \setminus \{i\} = a \restriction \omega \setminus \{i\}$.

Technically, this is a recursion along the relation $\varphi \prec \psi$ iff φ is either the second or third coordinate of the tuple ψ.

We then write $\mathcal{N} \models \varphi(x_0, \dots, x_n)$ for $\mathrm{Sat}_{\mathcal{N}}(a, \varphi) = 1$ whenever $a \in N^{<\omega}$ so that $a(n_i) = x_i$ where n_i is the i-th element of $\mathrm{free}(\varphi)$.

Of course, this can be done for arbitrary languages as well.

If this $\mathcal{N}$ is an element of a metatheory model $\mathcal{M}$ of ZFC then the metatheory and $\mathcal{M}$ agree about the satisfaction relation. To be precise, the model $\mathcal{N}$ corresponds to a metatheory first order structure

$$\mathcal{N}^{\mathrm{MT}} := (\{x \in M \mid \mathcal{M} \models x \in N\}, \{(x,y) \mid \mathcal{M} \models (x,y) \in E\})$$

where the terms are evaluated "in the metatheory". For any $x_0, \ldots, x_n \in^{\mathcal{M}} N$, and any metatheory first order $\in$-formula φ we have

$$\mathcal{N}^{\mathrm{MT}} \models^{\mathrm{MT}} \varphi(x_0, \ldots, x_n) \Leftrightarrow \mathcal{M} \models^{\mathrm{MT}} (\mathcal{N} \models \ulcorner\varphi\urcorner(x_0, \ldots, x_n)).$$

Now that everything is set up, we can import all theorems from first order logic as theorems of first order logic inside of Set Theory, such as the Löwenheim-Skolem theorems and Gödel's completeness and incompleteness theorems. We don't have to prove anything here again, any argument valid in the metatheory is valid in ZFC (but not necessarily in ZF, e.g. the proof of Gödel's completeness theorem makes use of the axiom of choice!).

We can also import the axioms of ZFC as well as the whole theory ZFC. For any metatheory axiom φ of ZFC (defined in this lecture notes) the Gödelization $\ulcorner\varphi\urcorner$ is an axiom of the Gödelized $\ulcorner\mathrm{ZFC}\urcorner$[6].

Gödel's second incompleteness theorem implies that ZFC cannot prove that $\ulcorner\mathrm{ZFC}\urcorner$ is consistent, so there may not be any $\in$-model $\mathcal{N}$ such that $\mathcal{N} \models \ulcorner\mathrm{ZFC}\urcorner$. However, we will see that we can get arbitrarily close to that.

For readability, we now **stop making Gödelizations explicit**. We confuse e.g. ZFC and $\ulcorner\mathrm{ZFC}\urcorner$.

Definition 6.1 (The Levy-Hierarchy)**.** We define complexity classes of $\in$-formulas.

(*i*) The set of $\Sigma_0 = \Delta_0 = \Pi_0$-formulas is the smallest class of $\in$-formulas containing the atomic formulas and closed under $\neg, \wedge$ (hence $\vee, \rightarrow$) as well as bounded quantification $\exists x \in y$ (hence $\forall x \in y$).

(*ii*) The Σ_{n+1}-formulas are those of the form $\exists x_0 \ldots \exists x_n \varphi$ where φ is a Π_n-formula.

(*iii*) The Π_{n+1}-formulas are those of the form $\forall x_0 \ldots \forall x_n \varphi$ where φ is a Σ_n-formula.

Definition 6.2. Suppose $\mathcal{N}_0 = (N_0, \ldots)$ is a substructure of $\mathcal{N}_1 = (N_1, \ldots)$. A first order formula φ in the language of the $\mathcal{N}_i$ is

(*i*) downwards absolute between $\mathcal{N}_0, \mathcal{N}_1$ if for all $x_0, \ldots, x_n \in N_0$, $\mathcal{N}_1 \models \varphi(x_0, \ldots, x_n)$ implies $\mathcal{N}_0 \models \varphi(x_0, \ldots, x_n)$,

(*ii*) upwards absolute between $\mathcal{N}_0, \mathcal{N}_1$ if for all $x_0, \ldots, x_n \in N_0$, $\mathcal{N}_0 \models \varphi(x_0, \ldots, x_n)$ implies $\mathcal{N}_1 \models \varphi(x_0, \ldots, x_n)$,

[6] Since schemes are part of the axiomsystem ZFC, we may have that not every formula in $\ulcorner\mathrm{ZFC}\urcorner^{\mathcal{M}}$ is of the form $\ulcorner\varphi\urcorner^{\mathcal{M}}$.

(iii) absolute between $\mathcal{N}_0, \mathcal{N}_1$ if it is both downwards and upwards absolute.

Proposition 6.3. *Suppose T is any transitive set and $\varphi(v_0, \dots, v_n)$ is any Δ_0-formula (in the metatheory). Then φ is absolute between V and T, i.e.*

$$\forall x_0 \in T \dots \forall x_n \in T\ \varphi(x_0, \dots, x_n) \leftrightarrow (T, \in) \models \varphi(x_0, \dots, x_n).$$

Proof. This can be seen by an induction on the complexity of φ. The only non-trivial case is the one of bounded quantification. But if $\exists x \in y\ \varphi$ holds then any witness $x \in y \in T$ of this is itself in T by transitivity of T. So $(T, \in) \models \exists x \in y\ \varphi$. $\square$

It follows easily that Π_1-formulas are downwards absolute from V to a transitive set.

Corollary 6.4. *If T is a transitive set then*

(i) $(T, \in) \models$ *(Extensionality),*

(ii) $(T, \in) \models$ *(Pairing) iff $\{x, y\} \in T$ whenever $x, y \in T$,*

(iii) $(T, \in) \models$ *(Union) iff $\bigcup x \in T$ whenever $x \in T$,*

(iv) $(T, \in) \models$ *(Power) iff $\mathcal{P}(x) \cap T \in T$ for all $x \in T$,*

(v) *if $\omega \in T$ then $(T, \in) \models$ (Infinity) and*

(vi) $(T, \in) \models$ *(Foundation).*

Proof. (Extensionality) is a Π_1-formula and the formulas $z = \{x, y\}$, $z = \bigcup x$ and "z is inductive" are all Δ_0, so $(i) - (iii), (v)$ follow. For (iv), note that $x \subseteq y$ is a Δ_0-formula and hence if $z \coloneqq \mathcal{P}(x) \cap T \in T$ then $(T, \in) \models z = \mathcal{P}(x)$.

For (vi), observe that for any $\in$-term A (at the level of Set Theory, not the metatheory), $A' = \{x \in T \mid (T, \in) \models x \in A\}$ is a term of the metatheory. So if $x \in A'$ is such that $x \cap A' = \emptyset$ then $x \in T$ and $(T, \in) \models x \in A \land x \cap A = \emptyset$. $\square$

Lemma 6.5. *If $\alpha \in \mathrm{Lim}$, $\alpha > \omega$ then $V_\alpha \models \mathrm{ZFC} -$ (Replacement).*

Proof. Exercise. $\square$

6.1 The Reflection Theorem

We will now show that every first order property of V "reflects" down to some transitive set, in fact some V_α. This is known both as the **Reflection Theorem** and the **Reflection Principle**. It is a very valuable mathematical tool on the one side, but is also of philosophical interest. We warn the reader interested in the philosophy of Set Theory that these lecture notes here are going to disappoint in this aspect. We hope that the mathematical content is not disappointing.

Theorem 6.6 (Montague)**.** *Suppose $\varphi_0, \dots, \varphi_n$ are $\in$-formulas. Then there is an ordinal α so that $\varphi_0, \dots, \varphi_n$ are absolute between V and V_α, i.e.*

$$\forall x_0 \in V_\alpha \dots \forall x_{k_i} \in V_\alpha \ (\varphi_i(x_0, \dots, x_{k_i}) \leftrightarrow (V_\alpha, \in) \models \varphi_i(x_0, \dots, x_{k_i}))$$

holds for all $i \leq n$, where k_i is the number of free variables of φ_i.

This is once again a “metatheorem” in the sense that it is a single theorem for any instance of (meta-theoretical) $\in$-formulas $\varphi_0, \dots, \varphi_n$. The displayed statement is really one $\in$-formula in the sense that $(V_\alpha, \in) \models \varphi_i(\dots)$ is the satisfaction relation is the one formalized within Set Theory. In particular, the formula φ_i here is really the Gödelized $\ulcorner \varphi_i \urcorner$. We promise that this is the last time we mention Gödelization explicitly.

We will actually prove something slightly stronger and more general than the result above.

Definition 6.7. A sequence $\langle H_\alpha \mid \alpha \in \mathrm{Ord} \rangle$ is a **continuous cumulative hierarchy** or simply a **hierarchy** if

(*i*) $H_\alpha \subseteq H_\beta$ for $\alpha \leq \beta$ and

(*ii*) $H_\alpha = \bigcup_{\beta < \alpha} H_\beta$ for $\alpha \in \mathrm{Lim}$.

Such a hierarchy either stabilizes or grows into a proper class. The satisfaction relation can only be defined, inside of set theory, for set sized structures, but not for proper classes. We could interpret the proper class as a structure in the metatheory, but opt for the following nicer alternative instead. Both approaches have the same outcome.

Definition 6.8. For a class M and a first order formula φ (of the metatheory), we define φ^M by induction.

- $\varphi^M = \varphi$ if φ is atomic.
- $(\neg \varphi)^M = \neg \varphi^M$,
- $(\varphi \wedge \psi)^M = \varphi^M \wedge \psi^M$ and
- $(\exists x \ \varphi)^M = \exists x \in M \ \varphi^M$.

We also write $M \models \varphi$ or $(M, \in) \models \varphi$ for φ^M.

Lemma 6.9. *Suppose $\langle H_\alpha \mid \alpha \in \mathrm{Ord} \rangle$ is a hierarchy and φ is a $\in$-formula. Let $H_* = \bigcup_{\alpha \in \mathrm{Ord}} H_\alpha$. Then there is a proper class (term) C_φ so that*

(*i*) *C_φ is club in* Ord*, i.e. closed and unbounded in* Ord*,*

(*ii*) *for all $\alpha \in C_\varphi$, φ is absolute between H_α and H_*, i.e.*

$$\forall x_0 \in H_\alpha \dots \forall x_n \in H_\alpha ((H_\alpha, \in) \models \varphi(x_0, \dots, x_n) \leftrightarrow (H_*, \in) \models \varphi(x_0, \dots, x_n)).$$

More precisely, it is possible to write a computer program with output the term C_φ on input the formula φ.

Proof. We argue by induction on the complexity of φ.

φ is atomic. Then set $C_\varphi = \mathrm{Ord}$.

$\varphi = \neg\psi$. Set $C_\varphi = C_\psi$.

$\varphi = \psi \wedge \theta$. Set $C_\varphi = C_\psi \cap C_\theta$. Note that the proof of Lemma 5.6 shows that clubs of ordinals intersect in a club as well. Simply use (Replacement) instead of regularity and uncountability.

$\varphi = \exists x\ \psi$. Define the function

$$f_\varphi \colon \mathrm{Ord} \to \mathrm{Ord}$$

via

$$\begin{aligned} f_\varphi(\alpha) = \min\{\beta \in \mathrm{Ord}\ | \forall y_0 \in H_\alpha \ldots y_n \in H_\alpha (H_* \models \exists x\ \psi(x, y_0, \ldots, y_n) \\ \to \exists x \in H_\beta\ H_* \models \psi(x, y_0, \ldots, y_n))\} \end{aligned}$$

An application of (Replacement) shows that f_φ is a well-defined function. The set of closure points $C_{f_\varphi} = \{\alpha \in \mathrm{Ord} \mid f[\alpha] \subseteq \alpha\}$ is a club in Ord. We now set $C_\varphi = C_\psi \cap C_{f_\varphi} \cap \mathrm{Lim}$. As an intersection of three clubs, this is a club itself. Now suppose $\alpha \in C_\varphi$ and $y_0, \ldots, y_n \in H_\alpha$. Since α is a limit ordinal and $H_\alpha = \bigcup_{\beta<\alpha} H_\beta$, there is some $\beta < \alpha$ so that $y_0, \ldots, y_n \in H_\beta$. If $(H_\alpha, \in) \models \exists x\ \psi(x, y_0, \ldots, y_n)$ then it is easy to see that the same is true for H_* as $\alpha \in C_\psi$. On the other hand, if $(H_*, \in) \models \exists x\ \psi(x, y_0, \ldots, y_n)$, then there is an $x \in H_{f_\varphi}(\beta)$ with

$$(H_*, \in) \models \psi(x, y_0, \ldots, y_n).$$

As $\alpha \in C_{f_\varphi}$, $f_\varphi(\beta) < \alpha$ and hence $x \in H_\alpha$. As $\alpha \in C_\psi$, it follows that $(H_\alpha, \in) \models \varphi(x, y_0, \ldots, y_n)$ as well.

□

We remark that the class $\{\alpha \in \mathrm{Ord} \mid \varphi \text{ is absolute between } H_\alpha \text{ and } H_*\}$ is not closed itself in general.

The Reflection Theorem is an immediate consequence of applying Lemma 6.9 to the Von-Neumann-hierarchy and intersecting the relevant finitely many clubs.

Corollary 6.10. *The theory* ZFC *is not finitely axiomatizable.*

Proof. Suppose $\varphi_0, \ldots, \varphi_n$ are any finitely many $\in$-formulas that ZFC proves to hold true. By the Reflection Theorem, there is some α so that $\varphi_0, \ldots, \varphi_n$ are absolute between V and V_α so that V_α is a model of $\varphi_0 \wedge \cdots \wedge \varphi_n$ and V realizes this to be true. Hence ZFC proves the consistency of the theory $T \coloneqq \{\varphi_0, \ldots, \varphi_n\}$. By Gödel's second Incompleteness Theorem, the theory T cannot prove all of ZFC. □

In fact, it is possible to show that for any (meta) natural number n, there is an ordinal α so that all Σ_n-formulas are absolute between V_α and V. Similarly, the Σ_n-fragment of ZFC is strictly weaker than full ZFC.

We remark that there is a very small n so that the Σ_n-fragment of ZFC proves all prominent theorems of mathematics. Likely, $n = 4$ should be more than enough. 23.04.24

7 Gödel's Constructible Universe

In this section, **we work in** ZF. Our goal is to build a transitive proper class L so that $L \models \mathrm{ZFC} + \mathrm{GCH}$. Hence we will prove the consistency of $\mathrm{ZFC} + \mathrm{GCH}$ from the consistency of ZF. In particular, neither the axiom of choice nor the Continuum Hypothesis or even GCH can introduce any contradictions which are already present in ZF. This is great as the Axiom of Choice is not as broadly believed to be true among mathematicians as the axioms of ZF are.

Since ZF does not prove the existence of a set-sized model of ZF, L will be a proper class and contain all ordinals. In fact, L will be the smallest transitive class model of ZF containing all ordinals. We will define L in levels, similarly to the Von-Neumann hierarchy, the only difference is that we only put in new sets which "have to be there". In practice, this means that the new sets can be "constructed" in a very absolute and hence robust way from the previous sets.

Proposition 7.1. *Suppose X is a non-empty transitive sets closed under pairing and $M \in X$ is a first order structure in the language $\mathcal{L} \in X$. For every first order formula φ in the language of M and $a_0, \dots, a_n \in M$, we have $M \models \varphi(a_0, \dots, a_n)$ iff*

$$(X, \in) \models \varphi(a_0, \dots, a_n).$$

Proof. First note that every first order formula in the language $\mathcal{L}$ is an element of X since X is closed under pairing. The result now follows by induction along the complexity of φ. Note that the relevant functions coming up in the recursive definition of the satisfaction relation are all hereditarily finite objects "over $\mathcal{L}$" and belong to X. $\square$

It follows that any two transitive sets containing a common first order structure (and its language) agree about the satisfaction relation over that model.

Definition 7.2. Suppose X is any non-empty set. The **definable powerset of** X is

$$\mathrm{Def}(X) := \{A \subseteq X \mid \exists \varphi \in \mathrm{Fml}_\in \exists a_0, \dots a_n \in X\ A = \{x \in X \mid X \models \varphi(x, a_0, \dots, a_n)\}\}.$$

We also set[7] $\mathrm{Def}(\emptyset) = \mathcal{P}(\emptyset)$.

[7] We do this since technically the universe of any first order structure is non-empty by definition.

It follows from Proposition 7.1 that the definable powerset of a set X is extremely robust: any two transitive models of (Pairing) and (Separation) compute the exact same definable powerset of X.

Definition 7.3. The **L-hierarchy** is defined by recursion as follows:

(*i*) $L_0 = \emptyset$.

(*ii*) $L_{\alpha+1} = \mathrm{Def}(L_\alpha)$.

(*iii*) $L_\alpha = \bigcup_{\beta<\alpha} L_\beta$ if $\alpha \in \mathrm{Lim}$.

Gödels constructible universe is

$$L \coloneqq \bigcup_{\alpha \in \mathrm{Ord}} L_\alpha.$$

The L-hierarchy behaves somewhat similarly as the V-hierarchy.

Lemma 7.4. *Let α be an ordinal.*

(*i*) *L_α is transitive,*

(*ii*) *$L_\alpha \in L_{\alpha+1}$ and $L_\alpha \subseteq L_{\alpha+1}$,*

(*iii*) *$L_\alpha \cap \mathrm{Ord} = \alpha$.*

(*iv*) *If $x, y \in L_\alpha$ then $\{x, y\} \in L_{\alpha+1}$,*

(*v*) *if $x \in L_\alpha$ then $\bigcup x \in L_\alpha$.*

7.1 L is a model of ZFC

Lemma 7.5. *L is a model of* ZF*.*

Proof. The class L is a model of (Extensionality),(Pairing), (Union), (Infinity) and (Foundation) by Corollary 6.4 and Lemma 7.4. Let us show (Power), so we have to show $\mathcal{P}(x) \cap L \in L$ for $x \in L$. If $x \in L$ then define

$$f \colon \mathcal{P}(x) \cap L \to \mathrm{Ord}, a \mapsto \min\{\alpha \in \mathrm{Ord} \mid a \in L_\alpha\}.$$

By (Replacement), $\beta \coloneqq \sup \mathrm{ran}(f) \in \mathrm{Ord}$ and $\mathcal{P}(x) \cap L \subseteq L_{\beta+1}$. But then

$$\mathcal{P}(x) \cap L = \{a \in L_{\beta+1} \mid L_{\beta+1} \models a \subseteq x\}$$

and hence $\mathcal{P}(x) \cap L \in \mathrm{Def}(L_{\beta+1}) = L_{\beta+2} \subseteq L$.

Note that (Separation) follows from (Replacement) (given the other axioms), so we will show (Replacement). Suppose that $a, p \in L$ and

$$F = \{x \mid \varphi(x, p)\}$$

is a class term so that $L \models$ "F is a function on a". By the Reflection Theorem 6.6, there is some $\alpha \in \mathrm{Ord}$ so that $a, p \in L_\alpha$ and φ as well as the formula "F is a function on v" are absolute between L and L_α. But then

$$F[a]^L = \{c \in L_\alpha \mid L_\alpha \models \exists b \in a\ F(b) = c\} \in \mathrm{Def}(L_\alpha) = L_{\alpha+1} \subseteq L.$$

□

So far, working in ZF, we have constructed another model of ZF. Hardly impressive. The first bit of magic happens if we can show that the axiom of choice holds in L, or equivalently, that every set in L has a wellorder in L. Something better is even true, L has a **global wellorder**. This just means that there is a term $\prec_L$, when evaluated in L, gives a wellorder of the whole class L.

First, we remark that L can compute itself.

Definition 7.6. Suppose W is any transitive model of ZF containing all ordinals. Then $L_\alpha^W = L_\alpha$ for all ordinals α, in particular $L^W = L$.

Proof. This is true by induction on α. The case $\alpha = 0$ and the limit step is trivial. The successor step easily follows from Proposition 7.1. □

The upshot is that L can compute the L-hierarchy itself and it coincides with the "true" L-hierarchy (if $L \neq V$, this is not necessarily true for other hierarchies e.g. the V-hierarchy).

Definition 7.7. Fix a easily definable enumeration $\langle \varphi_n \mid n < \omega \rangle$ of all $\in$-formulas, in particular so that this enumeration is in L. We define an order $<_\alpha$ on L_α by recursion as follows:

(*i*) $<_0 = \emptyset$.

(*ii*) For $x, y \in L_{\alpha+1}$ we set $x <_{\alpha+1} y$ iff one of the following holds:

- $x, y \in L_\alpha$ and $x <_\alpha y$.
- $x \in L_\alpha$ and $y \notin L_\alpha$.
- $x, y \in L_{\alpha+1} \setminus L_\alpha$. Let $n, m < \omega$ be least so that x, y are definable over L_α by φ_n, φ_m respectively. Then either $n < m$ or if $a_0, \ldots, a_k$, $b_0, \ldots b_k$ are lexicographically[8] $<_\alpha$-least in L_α so that

  $$x = \{c \in L_\alpha \mid L_\alpha \models \varphi_n(c, a_0, \ldots, a_k)\}, y = \{c \in L_\alpha \mid L_\alpha \models \varphi_n(b_0, \ldots, b_k)\}$$

  then $a_0, \ldots, a_k$ is lexicographically $<_\alpha$ strictly less than $b_0, \ldots, b_k$.

(*iii*) For α a limit ordinal, $<_\alpha = \bigcup_{\beta<\alpha} <_\beta$.

The **canonical wellorder on** L **is** $<_L := \bigcup_{\alpha \in \mathrm{Ord}} <_\alpha$.

Proposition 7.8. *The order $<_L$ is a wellorder on L.*

[8] For any wellorder $\prec$ on a set A, the induced lexicographic wellorder on A^{k+1} is given by $(p_0, \ldots, p_k) \prec_{\mathrm{lex}} (q_0, \ldots, q_k)$ if the least i with $p_i \neq q_i$ exists and $p_i \prec q_i$ holds.

Proof. By a moment of reflection. □

Theorem 7.9. *L is a model of* ZFC.

Proof. It only remains that the Axiom of Choice holds in L. Let $x \in L$ be any set. By induction on α, one easily sees that the orders $<_\alpha$ are absolute between V and L, i.e. $(<_\alpha)^L = <_\alpha$. It follows that $(<_L)^L = <_L$ is definable over L and hence $\prec = <_L \cap (x \times x) \in L$ by $(\text{Separation})^L$. This is a wellorder on x. □

7.2 Condensation and GCH in L

We now develop some methods to show the Generalized Continuum Hypothesis in L. Recall the notion of an elementary substructure from first order logic.

Definition 7.10. Suppose M, N are first order structures in the same language and $M \subseteq N$ is a substructure. Then we write $M \prec N$ if M is an **elementary substructure of** N, i.e. all first order formulas in the language of M, N are absolute between M and N.

We aim to prove the following theorem known as the **Condensation Lemma** due to Gödel.

Lemma 7.11. *Suppose α is a limit ordinal and $(X, \in) \prec (L_\alpha, \in)$. If $(M, \in)$ is the transitive collapse of $(X, \in)$ then $M = L_\beta$ for some $\beta \leq \alpha$.*

Proof. We will show that there is a $\in$-formula φ so that whenever Y is a transitive set, $(Y, \in) \models \varphi$ iff $Y = L_\beta$ for some $\beta \in \text{Lim}$. The result then follows immediately. The formula φ is the conjunction of the following formulas:

(*i*) (Pairing),

(*ii*) $\forall \gamma \in \text{Ord}\ \exists x\ x = \gamma + 1$,

(*iii*) "L_γ exists for all ordinals γ", i.e. for any $\gamma \in \text{Ord}$, there is a sequence $\langle L'_\delta \mid \delta \leq \gamma \rangle$ so that $L'_0 = \emptyset$, $L'_{\delta+1} = \text{Def}(L'_\delta)$ and $L'_\delta = \bigcup_{\xi < \delta} L'_\xi$ for $\delta \in \text{Lim}$,

(*iv*) $V = \bigcup_{\alpha \in \text{Ord}} L_\alpha$.

It is not difficult to see that for an ordinal γ, there is some finite n so that $\langle L_\delta \mid \delta \leq \gamma \rangle \in L_{\gamma+n}$ (the exact n depends on the choice of implementation of an ordered pair, we encourage the reader to calculate the n for our official implementation). It follows that $L_\beta \models \varphi$ for any limit ordinal β.

Conversely, suppose Y is a transitive set so that $Y \models \varphi$. Clearly $Y \cap \text{Ord}$ is a limit ordinal β and we will show that $Y = L_\beta$. For any $\gamma < \beta$, let $\langle L'_\delta \mid \delta \leq \gamma \rangle$ witness that $Y \models$ "L_γ exists". By induction, using Proposition 7.1 in the successor step, one easily sees that $L'_\delta = L_\delta$ for any $\delta \leq \gamma$. Hence Y computes the "true" L-hierarchy up to β. As $Y \models \varphi$, we hence have $Y = \bigcup_{\gamma < \beta} L_\gamma = L_\beta$. □

Definition 7.12. The formula $V = L$ is the formula φ defined in the proof above.

We also remind the reader of the Löwenheim-Skolem downwards theorem that we have generously imported from first order logic. In our context, we can formulate it as follows.

Theorem 7.13. *Assume* ZFC*. Then for any non-empty X and $A \subseteq X$, there is some $A \subseteq Y \subseteq X$ so that $(Y, \in) \prec (X, \in)$ and $|Y| \leq |A| + \aleph_0$.*

The following simple observation will be crucial.

Proposition 7.14. $|L_\alpha| = |\alpha|$ *for infinite* $\alpha \in \mathrm{Ord}$.

Proof. First note that $L_n = V_n$ for $n < \omega$ and hence $L_\omega = V_\omega$ so that $|L_\omega| = \omega$. We have proven the base case of our induction along α. The limit step is easy, so we focus on the successor step. Any $x \in L_{\alpha+1}$ is defined over L_α via some $\in$-formula and a finite sequence of parameters from L_α. Hence we have $|L_{\alpha+1}| \leq |\mathrm{Fml}_\in| \cdot |L_\alpha^{<\omega}|$. For any infinite wellordered set X, we have $|X^{<\omega}| = \sum_{n<\omega} |X^n| = \sum_{n<\omega} |X|^n = \sum_{n<\omega} |X| = \aleph_0 \cdot |X| = |X|$, where we use Hessenberg's theorem to deduce $|X|^n = |X|$ and the last equality (Hessenberg's theorem for wellordered cardinalities does not rely on the axiom of choice).

We already know that L_α is wellordered and hence $|L_\alpha^{<\omega}| = |L_\alpha| = |\alpha|$ holds by induction. It follows that $|\alpha| \leq |L_{\alpha+1}| \leq \aleph_0 \cdot |\alpha| = |\alpha|$. $\square$

Theorem 7.15. *L is a model of* GCH*.*

Proof. Let us work in L and let κ be an infinite cardinal.

Claim 7.16. $\mathcal{P}(\kappa) \subseteq L_{\kappa^+}$.

Proof. Suppose $X \subseteq \kappa$. There is some α so that $X \in L_\alpha$ and by Löwenheim-Skolem, there is some $Y \prec L_\alpha$ so that $\kappa + 1 \cup \{X\} \subseteq Y$ and $|Y| = \kappa$. Let $\pi\colon M \to Y$ be the anti-collapse map and by the Condensation Lemma, there is some $\beta \leq \alpha$ so that $M = L_\beta$. Since $\kappa + 1 \subseteq Y$, we see that $\pi \restriction \kappa + 1 = \mathrm{id}_{\kappa+1}$ (verify this by induction!) and hence if $\bar{X} \in L_\beta$ with $\pi(\bar{X}) = X$, we have $X = \bar{X} \in L_\beta$.

Also we have that $|\beta| = |L_\beta| = |Y| = \kappa$ and hence $\beta < \kappa^+$. $\square$

It follows that $2^\kappa = |\mathcal{P}(\kappa)| \leq |L_{\kappa^+}| = \kappa^+$, so we are done. $\square$

24.04.24

7.3 The $\diamondsuit$-principle

The constructible universe can be analyzed in much more detail. The "natural" statementes we know to be independent of ZFC tend to fall into one of two categories:

- Statements about the "height" of the universe, i.e. demanding the existence of a certain type of "large cardinal" (which we will deal with later).

- Statements about the "width" of the universe, for example the Continuum hypothesis.

Empirically, every such statement[9] about the width of the universe is decidable in L, e.g. we have seen that GCH holds in L.

Definition 7.17 (Jensen)**.** The **diamond principle** $\diamondsuit_\kappa$ at a regular uncountable cardinal κ holds if there is a sequence $\langle a_\beta \mid \beta < \kappa$ so that

(i) $a_\beta \subseteq \beta$ for every $\beta < \kappa$ and

(ii) for any $X \subseteq \kappa$, the set

$$\{\beta < \kappa \mid a_\beta = X \cap \beta\}$$

is stationary in κ.

We also write $\diamondsuit$ instead of $\diamondsuit_{\omega_1}$.

The $\diamondsuit$-principle is an instance of "guessing principles", in this case there is a sequence of length κ which guesses all the 2^κ-many subsets of κ correctly on a big, i.e. stationary, set.

The $\diamondsuit$-principle on a successor cardinal is connected to the GCH on the previous cardinal.

Lemma 7.18. *If κ is an infinite regular cardinal such that $\diamondsuit_\kappa$ holds then $2^\lambda \leq \kappa$ for all infinite cardinals $\lambda < \kappa$.*

Proof. Exercise. □

Shelah has shown that, surprisingly, this is somewhat reversible.

Theorem 7.19 (Shelah)**.** *If κ is an uncountable cardinal with $2^\kappa = \kappa^+$ then $\diamondsuit_{\kappa^+}$ holds.*

However, this is **not** true for $\kappa = \omega$: it is consistent with ZFC that CH holds, yet $\diamondsuit$ fails.

Theorem 7.20. *Assume $V = L$. Then $\diamondsuit_\kappa$ holds for any uncountable regular cardinal κ.*

In the following proof, a simple observation will be key which we state explicitly now.

Proposition 7.21. *If $\alpha \in \mathrm{Lim}$ then $(<_L)^{L_\alpha} = <_L \restriction L_\alpha$. That is, the canonicial wellorder on L is absolute between L and L_α.*

The proof is by inspection of the definition of $<_L$ and is better left to the reader.

[9] Of course Con(ZFC) is not decided by ZFC $+ V = L$, but such a statement is arguably not natural.

Proof of Theorem 7.20. This argument is a good example of another "archetype" of Set Theoretical proofs: putting "localized" counterexamples together onto a sequence for long enough and show that in the end, there is no "full" counterexample.

We construct sequences $\vec{C} = \langle C_\beta \mid \beta \in \kappa \cap \mathrm{Lim} \rangle$ and $\vec{a} = \langle a_\beta \mid \beta \in \kappa \cap \mathrm{Lim} \rangle$ by recursion so that $C_\beta \subseteq \beta$ is club and $a_\beta \subseteq \beta$. Suppose $a \upharpoonright \beta$ and $\vec{C} \upharpoonright \beta$ are defined. We split into two cases.

Case 1: There is some $a \subseteq \beta$ and a club $C \subseteq \beta$ so that $a_\gamma \neq a \cap \gamma$ for all $\gamma \in C$. Then let (C_β, a_β) be the $<_L$-least such pair.

Case 2: Case 1 fails. Then let $C_\beta = a_\beta = \beta$.

We claim that the sequence $\langle a_\beta \mid \beta \in \kappa \cap \mathrm{Lim} \rangle$ witnesses $\diamondsuit_\kappa$ (or technically it does so after filling it up with the empty set at successor indices). Suppose not and let (C, a) be the $<_L$-least pair so that $C \subseteq \kappa$ is a club, $a \subseteq \kappa$ and $a_\beta \neq a \cap \beta$ for all $\beta \in C$.

Claim 7.22. *There is a elementary substructure $X \prec L_{\kappa^+}$ such that $X \cap \kappa \in \kappa$.*

Proof. Such a substructure may be constructed as the union along a $\subseteq$-increasing union $(X_n)_{n<\omega}$ so that $X_0 \prec L_{\kappa^+}$ is a countable elementary substructure and $X_{n+1} \prec L_{\kappa^+}$ satisfies $|X_{n+1}| < \kappa$ and $\sup X_n \cap \kappa \subseteq X_{n+1}$. Note that since κ is regular uncountable, $\sup X_n \cap \kappa < \kappa$.

Then X is an elementary substructure by Tarski's chain lemma and $X \cap \kappa \in \kappa$ follows from $\mathrm{cof}(\kappa) > \omega$. $\square$

Note that $\vec{C}, \vec{a}, (C, a) \in X$ as the wellorder $<_L \upharpoonright L_{\kappa^+}$ is definable over L_{κ^+}.

By condensation, let $\alpha < \kappa^+$ so that L_α is the Mostowski collapse of the set X and let $\pi \colon X \to L_\alpha$ be the collapse map. Let $\beta = X \cap \kappa$.

Claim 7.23. $\beta \in C$.

Proof. We have $\pi(\kappa) = \{\pi(\gamma) \mid \gamma \in X \cap \kappa\} = \beta$ and $\pi(C) = \{\pi(\gamma) \mid \gamma \in X \cap C\} = C \cap \beta$. As π is an isomorphism, $L_\alpha \models$ "$\pi(C) \subseteq \beta$ is club" and hence β is a limit point of C. As C is closed, $\beta \in C$. $\square$

One similarly sees that $\pi(a) = a \cap \beta$, $\pi(\vec{a}) = \vec{a} \upharpoonright \beta$ and $\pi(\vec{C}) = \vec{C} \upharpoonright \beta$. We have that

$$L_{\kappa^+} \models \text{“}(C,a) \text{ is the } <_L \text{-least pair } (D,b) \text{ with } D \subseteq \kappa \text{ club and } b \subseteq \kappa \text{ with } b \cap \gamma \neq a_\gamma \text{ for } \gamma \in D\text{”}.$$

and by applying the isomorphism π, we have

$$L_\alpha \models \text{“}(C \cap \beta, a \cap \beta) \text{ is the } <_L \text{-least pair } (D,b) \text{ with } D \cap \beta \subseteq \beta \text{ club and } b \subseteq \beta \text{ with } b \cap \gamma \neq a_\gamma \text{ for } \gamma \in D\text{”}.$$

But then by Proposition 7.21, $(C \cap \beta, a \cap \beta)$ is really the $<_L$-least such pair and hence we put $C_\beta = C \cap \beta$ and $a_\beta = a \cap \beta$. But this is a contradiction as $\beta \in C$. $\square$

30.04.24

7.4 Suslin's Hypothesis

We will assume the Axiom of Choice again for the remainder of this section.

We are now going to see a famous application of the $\diamondsuit$-principle. Recall that the real line is the unique (up to isomorphism) linear order without endpoints which is

(i) dense, i.e. for all $x < y$ there is z with $x < z < y$,

(ii) complete, i.e. any bounded subset has admits an infimum and supremum and

(iii) separable, i.e. has a countable dense subset.

The Russian mathematician Mikhail Suslin asked whether separability can be weakened in this characterization.

Definition 7.24. Let $\mathcal{X}$ be a topological space.

(i) An **antichain** in $\mathcal{X}$ is a collection $\mathcal{A}$ of open sets so that $O \cap O'$ does not contain a non-empty open set for $O \neq O'$ both in $\mathcal{A}$.

The space $\mathcal{X}$ satisfies the **countable (anti-)chain condition (c.c.c.)** every antichain in $\mathcal{X}$ is countable.

Observe that any separable topological space satisfies the c.c.c. (this is essentially the proof that any monotonous function $f : \mathbb{R} \to \mathbb{R}$ is not continuous at at most countably many points).

Definition 7.25. Suslin's Hypothesis (SH) holds if any complete dense linear order without endpoints which satisfies the c.c.c. is isomorphic to $(\mathbb{R}, <)$.

It turns out that Suslin's hypothesis is not decided by ZFC, but it is decided by ZFC + $\diamondsuit$.

Theorem 7.26 (Jensen)**.** *If $\diamondsuit$ holds then Suslin's Hypothesis fails.*

We will now prove (most of) this theorem. First of all, Set Theorists like working with trees, as we have seen in the introduction. We transform SH into a statement about trees.

Definition 7.27. Suppose $\mathcal{T} = (T, <_T)$ is a partial order. Then $\mathcal{T}$ is a **tree** if $\mathcal{T}$ has a minimum and $\mathrm{pred}_{<_T}(t)$ is wellordered by $<_t$ for all $t \in T$.

Now suppose $\mathcal{T}$ is a tree.

(i) $\mathrm{ht}(t) = \mathrm{otp}(\mathrm{pred}_{<_T}(t))$ is the **height** of $t \in T$.

(ii) The **height** of $\mathcal{T}$ is $\mathrm{ht}(\mathcal{T}) = \sup\{\mathrm{ht}(t) + 1 \mid t \in T\}$.

(iii) For $\alpha < \mathrm{ht}(T)$, the set $T_\alpha = \{t \in T \mid \mathrm{ht}(t) = \alpha\}$ is the **α-th level of $\mathcal{T}$**.

(iv) A **branch** through $\mathcal{T}$ is a downwards closed subset of T linearly ordered by $<_T$.

(v) A branch $b \subseteq T$ is **maximal** if there is no branch c through $\mathcal{T}$ with $b \subsetneq c$. $\partial\mathcal{T}$ is the set of maximal branches through $\mathcal{T}$.

(vi) A branch $b \subseteq T$ is **cofinal** if $\operatorname{otp}((b, <_T)) = \operatorname{ht}(\mathcal{T})$. $[\mathcal{T}]$ is the set of cofinal branches through $\mathcal{T}$.

We will usually confuse T and $\mathcal{T}$.

For example $\omega^{<\omega}$, the set of all functions $f\colon n \to \omega$ for some $n < \omega$, is a tree of height ω when ordered by inclusion. The n-th level of $\omega^{<\omega}$ is ${}^n\omega$. Any function $g\colon \omega \to \omega$ induces a cofinal branch $\{g \restriction n \mid n < \omega\}$ (and every cofinal branch is of this form).

Every cofinal branch through a tree $\mathcal{T}$ is a maximal branch and any tree $\mathcal{T}$ has a maximal branch (assuming the axiom of choice), but not every tree has a cofinal branch. For example, consider the subtree of $\omega^{<\omega}$ of all functions $f\colon n \to \omega$ with $n = 0$ or $n < f(0)$.

Definition 7.28. A **Suslin tree** is a tree $(T, <_T)$ of height ω_1 without uncountable chains or antichains. This means that

(i) T has no cofinal branch and

(ii) if $\mathcal{A} \subseteq T$ so that $s \not\leq_T t$ and $t \not\leq_T s$ for $s \neq t$ in T, then $\mathcal{A}$ is countable.

Note that for any tree T, any level T_α of T is a maximal antichain of T, so if T is a Suslin tree then T_α is countable for all $\alpha < \omega_1$.

It is easy to construct a tree of height ω_1 without any cofinal branch: let $T = \{0\} \cup \{(\alpha, \beta) \mid \alpha < \beta < \omega_1\}$ so that 0 is the minimum point and $(\alpha, \beta) \leq (\gamma, \delta)$ iff $\alpha \leq \gamma$ and $\beta = \delta$. This is just wellorders of all countable lengths stuck together with a minimal point. Not that for $0 < \alpha < \omega_1$, $|T_\alpha| = \aleph_1$, so this tree is quite thick.

In general, there is some tension between the non-existence of a cofinal branch and how thin the tree is. A lot of interesting things live at the sweet spot of this tension, a Suslin tree is one such example.

The relevance of Suslin trees to Suslin's hypothesis should be clear form the following.

Theorem 7.29. *The following are equivalent.*

(i) *Suslin's hypothesis holds.*

(ii) *There is no Suslin tree.*

We will take this as given, the relevant implication will appear on the next exercise sheet.

Lemma 7.30. *Assume $\diamondsuit$ holds. Then there is a Suslin tree.*

The $\diamondsuit$-principle will help us as follows: The construction lasts ω_1-many steps and we want to diagonalize against 2^{ω_1} possible antichains. In each step, we may only deal with at most one such antichain, so it seems like we run out of time. However, the $\diamondsuit$-sequence allows us to reduce this problem to diagonalizing against ω_1-many guesses instead, which will suffice.

In fact, we will use a $\diamondsuit$-sequence to guess local maximal antichains and make sure they will not grow further. This is achieved via the following observation.

Proposition 7.31. *Suppose T is a tree, $\alpha < \mathrm{ht}(T)$ and $\mathcal{A} \subseteq T_{<\alpha} \coloneqq \bigcup_{\beta<\alpha} T_\beta$ is a maximal antichain in $T_{<\alpha}$. Further assume that every node $t \in T_\alpha$ is above a node of $\mathcal{A}$. Then $\mathcal{A}$ is a maximal antichain in T.*

Proof. We have to show that any $t \in T$ is comparable with some node of $\mathcal{A}$. This is clear if $\mathrm{ht}_T(t) < \alpha$. If $\mathrm{ht}_T(t) \geq \alpha$, let $s \leq_T t$ be the unique node below t in T_α. By assumption, s is above a node of $\mathcal{A}$ and hence so is t. $\square$

Proof of Lemma 7.30. Let us fix a $\diamondsuit$-sequence $\langle a_\beta \mid \beta < \omega_1\rangle$. We build a tree $(T, <_T)$ level by level and make sure that $T_\alpha \subseteq \omega \cdot (\alpha+1) \setminus \omega \cdot \alpha$, so that $T \subseteq \omega_1$.

The tree T will have additional nice properties: it will be **normal**, i.e. every node $t \in T$

- has ω-many immediate successors and
- can be extended to arbitrary high levels, that is for any $\mathrm{ht}(t) \leq \alpha < \omega_1$ there is some $t \leq_T t^+ \in T_\alpha$.

Moreover, T will be extensional.

We let 0 be the minimum on T so that $T_0 = \{0\}$. Now if T_β is defined, then T_β is countable and we give each $t \in T_\beta$ ω-many successors in the countably infinite set $\omega \cdot (\beta + 1) \setminus \omega \cdot \beta$.

Next, let us deal with the limit step $\beta \in \mathrm{Lim}$. Let $T_{<\beta} \coloneqq \bigcup_{\gamma<\beta} T_\gamma$.

Claim 7.32. *For $t \in T_{<\beta}$ there is a cofinal branch b through $T_{<\beta}$ with $t \in b$.*

Proof. Let $(\beta_n)_{n<\omega}$ be cofinal and increasing with supremum β and $\mathrm{ht}(t) = \beta_0$. By induction, $T_{<\beta}$ is normal and so we can recursively define a $\leq_T$-increasing sequence $(t_n)_{n<\omega}$ with $t_0 = t$ and $t_n \in T_{\beta_n}$. The downwards closure of $(t_n)_{n<\omega}$ is then a cofinal branch. $\square$

We now let $A_\beta = a_\beta$ in case a_β is a maximal antichain of $T_{<\beta}$ and $A_\beta = \{0\}$ otherwise. A_β is a maximal antichain in any case.

We will define T_β so that the antichain a_β is "sealed", i.e. cannot possibly grow any longer. To do this, we make sure that any $t \in T_\beta$ has some point in a_β as it's predecessor. On the other hand, we have to make sure that T stays normal. For any $t \in T_{<\beta}$ which is $\leq_T$-above some point in A_β, we chose a cofinal branch b_t through $T_{<\beta}$ with $t \in b_t$. We then add a point to T_β with $\mathrm{pred}_{<_T}(t) = b_t$.

As $T_{<\beta}$ is countable, the next level T_β is countable as well, so we may assume $T_\beta \subseteq \omega \cdot (\beta + 1) \setminus \omega \cdot \beta$.

This completes the construction of T.

Claim 7.33. *$(T, \leq_T)$ has no uncountable antichains.*

Proof. Let $A \subseteq T$ be an antichain and we may assume that A is a maximal antichain. For each $t \in T$, let $g(t)$ be the minimal β so that there is some $a \in A$ compatible with t and $\mathrm{ht}(a) = \beta$.

Define $f\colon \omega_1 \to \omega_1$ via
$$f(\alpha) = \sup g[T_\alpha].$$
Then
$$C_f = \{\alpha < \omega_1 \mid f[\alpha] \subseteq \alpha\} = \{\alpha < \omega_1 \mid A \cap T_{<\alpha} \text{ is a maximal antichain in } T_{<\alpha}\}$$
is a club. Hence there is some $\beta \in C_f$ which guesses A correctly, i.e. $a_\beta = A \cap \beta$. But then our definition of T_β made sure that any $t \in T$ of height $\geq\beta$ is above some element of $A \cap T_{<\beta}$. By Proposition 7.31, $A = A \cap T_{<\beta}$ is countable. □

Since T is normal, this also implies that T has no cofinal branch: If b were such a branch, then for each $t \in b$, choose an immediate successor $a_t \in T$ of t which is not in b. The set $\{a_t \mid t \in b\}$ is then an uncountable antichain, contradiction. □

Theorem 7.26 is an immediate consequence of Theorem 7.29 and Lemma 7.30.

7.5 Relative Constructibility

We briefly mention two variants of the L-construction. The first one makes sure by force that a certain set gets into the final model.

Definition 7.34. Suppose that A is a set. The $L_\alpha(A)$-hierarchy is defined by recursion.

- $L_0(A) = \mathrm{tc}(\{\mathrm{A}\})$, the transitive closure of $\{A\}$,
- $L_{\alpha+1}(A) = \mathrm{Def}(L_\alpha(A))$ and
- $L_\alpha(A) = \bigcup_{\beta<\alpha} L_\beta(A)$ if $\alpha \in \mathrm{Lim}$.

We let $L(A) = \bigcup_{\alpha\in\mathrm{Ord}} L_\alpha(A)$.

So the construction is exactly as the construction of L, except that we put A into the first step while still making sure that all levels are transitive.

Lemma 7.35. *Let $A \in V$. Then $A \in L(A)$ and $L(A)$ is a transitive model of* ZF*. Indeed $L(A)$ is the smallest such model with all ordinals, i.e. if W is any transitive model of* ZF *with* $\mathrm{Ord} \cup \{A\} \subseteq W$ *then* $L(A) \subseteq W$*.*

The proof is exactly the same as for L. We remark that the axiom of choice may fail in $L(A)$, for example the existence of sufficiently large cardinals implies that $L(\mathbb{R})$ contains no wellorder of $\mathbb{R}$. In general, $L(A) \models \mathrm{AC}$ iff $L(A)$ contains a wellorder of $\mathrm{tc}(\{A\})$.

The second variant has the advantage that it always produces models of ZFC and that it can be carried out relative to a proper class. For $A \subseteq M$, $(M, \in, A)$

is the structure in the language $\{\in, \dot{A}\}$, where $\dot{A}$ is an unary relational symbol interpreted as A in $(M, \in, A)$. This strucutre can "ask" which of its elements belong to A via the atomic formula $\dot{A}(x)$, which is usually denoted by $x \in \dot{A}$ instead.

Definition 7.36. Suppose that A is a class. The $L_\alpha[A]$-hierarchy is defined by recursion.

- $L_0[A] = \emptyset$,
- $L_{\alpha+1}[A] = \mathrm{Def}((L_\alpha[A]; \in A \cap L_\alpha[A]))$ and
- $L_\alpha(A) = \bigcup_{\beta<\alpha} L_\beta[A]$ if $\alpha \in \mathrm{Lim}$.

We let $L[A] = \bigcup_{\alpha \in \mathrm{Ord}} L_\alpha[A]$.

Lemma 7.37. *Let A be a class. Then $L[A] \models \mathrm{ZFC}$.*

Once again, the proof is exactly as for L.

However, we may not have that $A \in L[A]$, even if A is a set. For example, it is easily checked that $L[\{x\}] = L$ for any set x, so if $x \notin L$ then $\{x\} \notin L[\{x\}]$. However, A is a set of ordinals then $A \in L[A]$. In general, $A \cap L[A] \in L[A]$ and $L[A] = L(A \cap L[A])$ for any set A.

Since $L[A]$ is always a model of ZFC, it makes sense to ask whether GCH holds in $L[A]$. In general, this is not true. Nonetheless, a minor variation of the condensation lemma shows:

Theorem 7.38. *Suppose $A \subseteq \kappa$ for an ordinal κ. Then* GCH *above κ holds in $L[A]$, i.e. $L[A] \models \forall \lambda \geq \kappa\ 2^\lambda = \lambda^+$.*

8 Large Cardinals

07.05.24

We assume the Axiom of Choice in this section. Consider the collection $\mathcal{T}$ of all consistent computable theories[10] T extending ZFC, we want to order $\mathcal{T}$ with respect to the "strength of the theories". A theory $T \in \mathcal{T}$ has a **higher consistency strength than** $S \in \mathcal{T}$, denoted by $S \leq_{\mathrm{cons}} T$, if

$$\mathrm{ZFC} \vdash \mathrm{Con}(T) \to \mathrm{Con}(S),$$

where $\mathrm{Con}(T)$ is the $\in$-formula "T is consistent". For $S, T \in \mathcal{T}$, let $S \sim_{\mathrm{cons}} T$ iff $S \leq_{\mathrm{cons}} T$ and $T \leq_{\mathrm{cons}} S$. In this case, we say that S, T are **equiconsistent**. We also write $S <_{\mathrm{cons}} T$ if $S \leq_{\mathrm{cons}} T$ and S, T are not equiconsistent. In practice, we usually show $S <_{\mathrm{cons}} T$ by showing $T \vdash \mathrm{Con}(S)$.

The partial order $(\mathcal{T}, \leq_{\mathrm{cons}})$ is *very* complicated, for example every countable partial order can be embedded into this one.

Nonetheless, there is a nice chain $\mathrm{LC} \subseteq \mathcal{T}$ known as the **large cardinal hierarchy** we are about to explore.

[10]To be pedantic, we should work with computable representations of these theories instead of the theories themselves. We will glance over this distinction.

The following miraculous observation holds true empirically:

Every "natural" theory $T \in \mathcal{T}$ is belongs to the chain LC.

Of course, the word natural has no precise meaning here and roughly refers to all theories $T \in \mathcal{T}$ which are actually interesting to mathematicians. This is not helped by the fact that the term "large cardinal" itself does not have a precise meaning.

Roughly, a large cardinal property is a natural first order property ϕ, so that

- if $\varphi(\kappa)$ holds then κ is a cardinal,
- ZFC does not prove the existence of a cardinal κ with $\varphi(\kappa)$ and
- $\mathrm{ZFC} + \exists\kappa\ \varphi(\kappa)$ is not known to be inconsistent.

We then have $\mathrm{LC} = \{[\mathrm{ZFC} + \exists\kappa\ \varphi(\kappa)] \mid \varphi \text{ is a large cardinal property}\}$. In such a situation, it is better to go by example.

8.1 Worldly and inaccessible cardinals

Definition 8.1. A cardinal κ is **worldly** if $V_\kappa \models \mathrm{ZFC}$.

This is arguably the smallest instance of a large cardinal. By Gödels 2nd incompleteness theorem, ZFC does not prove the existence of a worldly cardinal.

Definition 8.2. A cardinal κ is a **strong limit cardinal** if $2^\lambda < \kappa$ for all $\lambda < \kappa$.

Definition 8.3. A **weakly inaccessible** cardinal is a regular uncountable limit cardinal. A **strongly inaccessible** (or just inaccessible) cardinal is a regular strong limit cardinal.

Clearly, any strongly inaccessible cardinal is weakly inaccessible, but a weakly inaccessible cardinal may not be strongly inaccessible.

Lemma 8.4. *The following relations hold:*

$$\begin{aligned}
&\mathrm{ZFC} + \exists\kappa\ \text{“}\kappa \textit{ is worldy”} \\
<_{\mathrm{cons}}&\mathrm{ZFC} + \exists\kappa\ \text{“}\kappa \textit{ is weakly inaccessible”} \\
\sim_{\mathrm{cons}}&\mathrm{ZFC} + \exists\kappa\ \text{“}\kappa \textit{ is strongly inaccessible”}.
\end{aligned}$$

Proof. You have seen in the exercises that if κ is strongly inaccessible then $V_\kappa \models \mathrm{ZFC}$. By applying the reflection theorem inside V_κ, we find that for every $\in$-formula φ there is a club $C_\varphi \subseteq \kappa$ so that φ is absolute between V_α and V_κ for all $\alpha \in C$. Hence if $\lambda \in \bigcap_{\varphi\in\mathrm{Fml}_\in} C_\varphi$ then $V_\lambda \prec V_\kappa$ λ is a cardinal so that λ is worldly. Clearly, $V_\kappa \models$ "λ is worldly" and hence V_κ is a model of $\mathrm{ZFC}+\exists\kappa$ "κ is worldly". Since V_κ is a set, $\mathrm{Con}(\mathrm{ZFC}+\exists\kappa$ "κ is worldly") holds.

It remains to show that the existence of a weakly inaccessible cardinal is equiconsistent to the existence of a strongly inaccessible cardinal. Suppose now that κ is weakly inaccessible. Then, since regularity is a Π_1-property, $L \models$ "κ is regular". Similarly, any cardinal in V is a cardinal in L and it follows that

$$L \models \text{“}\kappa \text{ is weakly inaccessible”}.$$

But $L \models$ GCH, so that κ must be strongly inaccessible in L. $\square$

Usually, large cardinal properties can be expressed equivalently in terms of the existence of certain nice elementary substructures. Or similarly, in terms of elementary embeddings.

Definition 8.5. Suppose $\mathcal{M} = (M, \dots), \mathcal{N} = (N, \dots)$ are two structures in the same language $\mathcal{L}$. A map

$$j \colon M \to N$$

is a **elementary embedding** between $\mathcal{M}$ and $\mathcal{N}$ if for all formulas φ in the language $\mathcal{L}$ and all $a_0, \dots a_n \in M$,

$$\mathcal{M} \models \varphi(a_0, \dots, a_n) \text{ iff } \mathcal{N} \models \varphi(j(a_0), \dots, j(a_n)).$$

We usually confuse $\mathcal{M}$ with M, $\mathcal{N}$ with N and write $j \colon \mathcal{M} \to \mathcal{N}$.

Observe that if $j \colon \mathcal{M} \to \mathcal{N}$ is an elementary embedding that $\operatorname{ran}(j)$ is an elementary substructure of $\mathcal{N}$.

If $\mathcal{N} = (N, \in, \dots)$ is a $\in$-structure on a transitive set N (with possibly additional structure) then the elementary substructures of $\mathcal{N}$ are in $1-1$ correspondence to elementary embeddings $j \colon \mathcal{M} \to \mathcal{N}$ where $\mathcal{M}$ is also a structure on a transitive set M. This correspondence is given by the Mostowski (anti-)collapse. We say that such an embedding $j \colon M \to N$ is **non-trivial** if $j \restriction \mathrm{Ord}^M$ is not the identity on Ord^M. If M, N are models of (sufficiently much of) ZFC, this is equivalent to $j(x) \neq x$ for some $x \in M$. In case j is non-trivial, the **critical point of** j is $\operatorname{crit}(j) = \min\{\alpha \in \mathrm{Ord}^M \mid j(\alpha) \neq \alpha\}$. Note that $j(\alpha) \geq \alpha$ for all ordinals $\alpha \in M$ and hence $j(\operatorname{crit}(j)) > \operatorname{crit}(j)$.

We record some observations regarding critical points.

Proposition 8.6. *Suppose that M, N are transitive and $j \colon M \to N$ is a non-trivial elementary embedding. Then $M \models$ "$\operatorname{crit}(j)$ is regular".*

Proof. Let $\kappa = \operatorname{crit}(j)$ and suppose $\alpha < \operatorname{crit}(j)$ and $f \colon \alpha \to \operatorname{crit}(j)$ is cofinal with $f \in M$. By elementarity, $j(f) \colon j(\alpha) \to j(\kappa)$ is cofinal. As $\alpha < \operatorname{crit}(j)$, $j(\alpha) = \alpha$. Further, for any $\beta < \alpha$ we have

$$j(f)(\beta) = j(f(j(\beta)) = j(f(\beta)) = f(\beta).$$

The first and last equality hold as $\beta, f(\beta) < \operatorname{crit}(j)$ and the middle equality follows from elementarity of j. $\square$

We stress that the critical point of an elementary embedding need not truly be regular in V, a witness to singularity can merely not be in the domain of the embedding.

Proposition 8.7. *Suppose N is a (non-empty) transitive set, $\kappa \in N$ is an ordinal and $X \prec N$ is an elementary substructure with $\kappa \in X$ and $X \prec \kappa \in \kappa$. If $j\colon M \to N$ is the Mostowski anticollapse of X then* $\mathrm{crit}(j) = X \cap \kappa$.

Proof. By a moment of reflection. □

Definition 8.8. For an infinite cardinal θ, define

$$H_\theta = \{x \mid |\mathrm{tc}(x)| < \theta\}.$$

The H_θ's form a continuous hierarchy just as the V_α's.

Lemma 8.9. *For any infinite cardinal θ, H_θ is a set and if θ is regular uncountable then $H_\theta \models \mathrm{ZFC} - (Power)$.*

Proof. Any element of H_θ can be coded by a subset of $\theta \times \theta$ and hence H_θ is the range of a function with domain $\mathcal{P}(\theta \times \theta)$.

Now assume θ is regular uncountable. We only show $H_\theta \models$ (Replacement). If $F\colon x \to H_\theta$ is any function definable over H_θ, then

$$\mathrm{tc}(F[x]) = \bigcup_{y \in x} \mathrm{tc}(F(y)).$$

Since x has size $<\theta$, $\mathrm{tc}(F(y))$ has size $<\theta$ for all $y \in x$ and θ is regular, $F[x] \in H_\theta$. □

The advantage of the H-hierarchy over the V-hierarchy is that it is easy to find members of the H-hierarchy in which (Replacement) holds, at the pain of loosing (Power).

Lemma 8.10. *Let κ be an uncountable cardinal. The following are equivalent:*

(*i*) *κ is weakly inaccessible.*

(*ii*) *There is a elementary substructure $X \prec H_{\kappa^+}$ so that $X \cap \kappa$ is a cardinal $< \kappa$.*

(*iii*) *There is a transitive set M and an elementary embedding*

$$j\colon M \to H_{\kappa^+}$$

with $j(\mathrm{crit}(j)) = \kappa$ such that $\mathrm{crit}(j)$ is a cardinal.

Proof. $(i) \Rightarrow (ii)$: Since κ is regular uncountable, we can build a $\subseteq$-increasing sequence $\langle X_i \mid i < \kappa\rangle$ of elementary substructures of H_{κ^+} so that

- $|X_i| < \kappa$
- $\sup(X_i \cap \mathrm{Ord}) \subseteq X_{i+1}$
- $X_i = \bigcup_{j<i} X_j$ if $i \in \mathrm{Lim} \cap \kappa$

for all $i < \kappa$. If $\alpha \in \mathrm{Lim} \cap \kappa$ then $\delta_\alpha \coloneqq X_\alpha \cap \kappa \in \kappa$. The set $D = \{\delta_\alpha \mid \alpha \in \kappa \cap \mathrm{Lim}\}$ is club in κ and since κ is a limit cardinal, $C = \mathrm{Card} \cap \kappa$ is a club in κ as well. Since κ is regular, $C \cap D$ is nonempty and X_α witnesses (ii) for any α with $\delta_\alpha \in D$.

$(ii) \Rightarrow (iii)$: Let j, M be given by the Mostowski anticollapse of X and apply Proposition 8.7.

$(iii) \Rightarrow (i)$: Let $j\colon M \to H_{\kappa^+}$ witness (iii). Let $\lambda = \mathrm{crit}(j)$ and recall that $M \models$ "λ is regular" so that $H_{\kappa^+} \models$ "$\kappa = j(\lambda)$ is regular". But H_{κ^+} contains any function $f\colon \alpha \to \kappa$ for $\alpha < \kappa$, so that κ is truly regular. It remains to show that κ is a limit cardinal. Otherwise, $\kappa = \delta^+$ for some cardinal δ. As δ is definable from κ, $\delta \in \mathrm{ran}(j)$ so that $\delta = j(\bar{\delta})$ for some $\bar{\delta} \in M$. But then $\bar{\delta} < \lambda$ and hence $\bar{\delta} = \delta$. This implies $\delta < \mathrm{crit}(j) < \delta^+$, which is impossible as $\mathrm{crit}(j)$ is a cardinal. $\square$

Essentially the same argument can be used to characterize inaccessible cardinals.

Lemma 8.11. *Let κ be an uncountable cardinal. The following are equivalent:*

(i) *κ is inaccessible.*

(ii) *There is a elementary substructure $X \prec H_{\kappa^+}$ so that $\delta = X \cap \kappa$ is a strong limit cardinal $< \kappa$.*

(iii) *There is a transitive set M and an elementary embedding*

$$j\colon M \to H_{\kappa^+}$$

with $j(\mathrm{crit}(j)) = \kappa$ and $\mathrm{crit}(j)$ a strong limit cardinal.

We now give an example of a natural theory equiconsistent to the existence of an inaccessible cardinal.

Definition 8.12. The axiom of **Dependent Choice** (DC) holds if for any relation R on a set X so that for all $x \in X$ there is $y \in X$ with xRy, there is a sequence $\langle x_n \mid n < \omega\rangle$ with $x_n R x_{n+1}$ for all $n < \omega$.

The axiom DC is a weak version of AC which is sufficient to do Analysis, e.g. one can prove in ZF + DC that a function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous at a point x iff $\lim_{n\to\infty} f(x_n) = f(x)$ for all sequences $(x_n)_{n<\omega}$ converging to x (this is not provable in ZF alone!).

Theorem 8.13 (Shelah-Solovay)**.** *The following theories are equiconsistent.*

(i) $\mathrm{ZFC} + \exists\kappa$ "*κ is inaccessible*".

(ii) $\mathrm{ZF} + \mathrm{DC} +$ "*all sets of reals are Lebesgue measurable*".

8.2 Mahlo Cardinals

08.05.24

We now move on to slightly larger cardinals.

Definition 8.14. A cardinal κ is **weakly Mahlo** if κ is regular and $\kappa \cap \mathrm{Reg}$ is stationary in κ. κ is **strongly Mahlo**, or simply Mahlo, if additionally $2^\lambda < \kappa$ for all $\lambda < \kappa$.

Lemma 8.15. *Suppose κ is (weakly) Mahlo. Then κ is (weakly) inaccessible limit of (weakly) inaccessible cardinals. In fact,*

$$\{\lambda < \kappa \mid \lambda \textit{ is (weakly) inaccessible}\}$$

is stationary in κ.

Proof. It is immediate that κ is (weakly) inaccessible. Let C be the set of limit cardinal $<\kappa$ if κ if κ is weakly Mahlo and the set of strong limit cardinals if κ is Mahlo. In any case, C is club in κ and as κ is (weakly) Mahlo,

$$\kappa \cap \mathrm{Reg} \cap C = \{\lambda < \kappa \mid \lambda \text{ is (weakly) inaccessible}\}$$

is stationary in κ. $\square$

Being a weakly Mahlo cardinal is a Π_1-property so any weakly Mahlo cardinal is weakly Mahlo in L and hence Mahlo in L as GCH holds in L. We can thus prove the following results for Mahlo cardinals similarly as the respective results about inaccessible cardinals.

Lemma 8.16. *The following relations hold:*

$$\begin{aligned} &\mathrm{ZFC} + \exists\kappa \text{ “}\kappa \textit{ is inaccessible”} \\ <_{\mathrm{cons}} &\mathrm{ZFC} + \exists\kappa \text{ “}\kappa \textit{ is weakly Mahlo”} \\ \sim_{\mathrm{cons}} &\mathrm{ZFC} + \exists\kappa \text{ “}\kappa \textit{ is Mahlo”}. \end{aligned}$$

We will prove a version of Lemma 8.10 for Mahlo cardinals, but first let us mention an easy but nonetheless important observation.

Proposition 8.17. *Suppose $j\colon M \to N$ is a nontrivial elementary embedding and M, N are transitive. Suppose $A \subseteq \mathrm{crit}(j)$ and $A \in M$. Then*

$$j(A) \cap \mathrm{crit}(j) = A.$$

This means that subsets of the critical points get “stretched” by the elementary embedding.

Proof. For $\alpha < \mathrm{crit}(j)$, we have

$$\alpha \in A \Leftrightarrow j(\alpha) \in j(A) \Leftrightarrow \alpha \in j(A)$$

where the first equivalence follows from the elementarity of j and the second as $\alpha < \mathrm{crit}(j)$. $\square$

Lemma 8.18. *Let κ be an uncountable cardinal. The following are equivalent:*

(*i*) *κ is (weakly) Mahlo.*

(*ii*) *There is a elementary substructure $X \prec H_{\kappa^+}$ so that $X \cap \kappa$ is a (weakly) inaccessible cardinal $< \kappa$.*

(*iii*) *There is a transitive set M and an elementary embedding*

$$j\colon M \to H_{\kappa^+}$$

with $j(\mathrm{crit}(j)) = \kappa$ such that $\mathrm{crit}(j)$ is (weakly) inaccessible.

Proof. (*i*) $\Rightarrow$ (*ii*) is similar to the relevant argument of Lemma 8.10. Note that we showed that we got a club of possible ordinals $X \cap \kappa$ in the argument there, so as the (weakly) inaccessibles $<\kappa$ are stationary now, (*ii*) follows.

(*ii*) $\Rightarrow$ (*iii*) is as before, so assume that $j\colon M \to H_{\kappa^+}$. We already know that κ is (weakly) inaccessible and it remains to see that $\kappa \cap \mathrm{Reg}$ is stationary. If not, there is a club $C \subseteq \kappa$ which does not contain any regular cardinal. By elementarity of j, we may assume that $C \in \mathrm{ran}(j)$ so that $C = j(\bar{C})$ for some $\bar{C} \in M$. By Proposition 8.17, $\bar{C} = C \cap \mathrm{crit}(j)$ and since $\bar{C} \subseteq \mathrm{crit}(j)$ is unbounded and C is closed, $\mathrm{crit}(j) \in C$. But this contradicts our assumption that $\mathrm{crit}(j)$ is (weakly) inaccessible. $\square$

9 Measurable Cardinals

As we have seen in the case of inaccessible and Mahlo cardinals, large cardinal properties can usually be expressed in terms of elementary embeddings $j\colon M \to N$ between transitive sets. The property in question gets stronger the more M and N resemble V (which can be achieved in various ways). The ultimate large cardinal property, in this sense, is an elementary embedding $j\colon V \to V$. We will see later that this is inconsistent with ZFC, but we may have an elementary embedding $j\colon V \to M$ with M a transitive class and such embeddings correspond to measurable cardinals.

Definition 9.1. Let X be any set. A set $\mathcal{F} \subseteq \mathcal{P}(X)$ is a **filter on** X if

(*i*) $X \in \mathcal{F}, \emptyset \notin \mathcal{F}$,

(*ii*) if $x \in \mathcal{F}$ and $x \subseteq y \subseteq X$ then $y \in \mathcal{F}$ and

(*iii*) if $x, y \in \mathcal{F}$ then $x \cap y \in \mathcal{F}$.

Definition 9.2. Let $\mathcal{F}$ be a filter on a set X.

(*i*) $\mathcal{F}$ on X **measures** a subset $x \subseteq X$ if $x \in \mathcal{F}$ or $X \setminus x \in \mathcal{X}$.

(*ii*) $\mathcal{F}$ is an **ultrafilter** (on X) if it measures every subset of X.

(*iii*) $\mathcal{F}$ is **non-principal** if $\bigcap F = \emptyset$.

For any $X \neq \emptyset$ and $x \in X$, $\mathcal{F}_x = \{A \subseteq X \mid x \in A\}$ is an example of a principal ultrafilter on X and in fact, any principal ultrafilter on X is of this form. In some sense, these are trivial and not that interesting.

A filter $\mathcal{F}$ on X can also be viewed as a finitely additive $0-1$ (partial) measure $\pi\colon \mathcal{P}(X) \xrightarrow{\text{partial}} 2$ via

$$\pi(a) = \begin{cases} 1 & \text{if } a \in \mathcal{F} \\ 0 & \text{if } X \setminus a \in \mathcal{F}. \end{cases}$$

We have $\pi(a) \leq \pi(b)$ for $a, b \in \mathrm{dom}(\pi)$ with $a \subseteq b$, $\pi(X) = 1$, $\pi(\emptyset) = 0$ and if $(a_i)_{i \leq n}$ are pairwise disjoint sets in $\mathrm{dom}(\pi)$ then $\pi(\bigcup_{i=0}^n a_i) = \sum_{i=0}^n \pi(a_i)$. If $\mathcal{F}$ is $<\kappa$-closed then π is $<\kappa$ additive.

ZF alone does not prove the existence of any non-principal ultrafilter. However, with the help of the axiom of choice, the following holds.

Lemma 9.3. *Any filter $\mathcal{F}$ can be extended to an ultrafilter $\mathcal{F} \subseteq \mathcal{U}$.*

Proof. Exercise. □

For any infinite set X, the **Freéchet filter** on X is $\mathcal{F} = \{A \subseteq X \mid X \setminus A \text{ is finite}\}$. This is a non-principal filter on X and when extended to an ultrafilter, yields a non-principal ultrafilter.

Now suppose that κ is a regular uncountable cardinal. Then the **club filter** on κ is

$$\mathcal{C}_\kappa = \{A \subseteq \kappa \mid \exists C \subseteq A \; C \text{ is club in } \kappa\}.$$

By Lemma 5.6, if $\alpha < \kappa$ and $A_\beta \in \mathcal{C}_\kappa$ for $\beta < \alpha$ then $\bigcap_{\beta<\alpha} A_\beta \in \mathcal{C}_\kappa$. A filter with this property is called **$<\kappa$-closed**. Note that this $<\kappa$-closure of a filter $\mathcal{F}$ is equivalent to the following: If $A \in \mathcal{F}$ and $A = \bigcup_{\beta<\alpha} B_\beta$ for some $\alpha < \kappa$ then $B_\beta \in \mathcal{F}$ for some $\beta < \alpha$.

Definition 9.4. Let κ be an uncountable cardinal. Then κ is **measurable** if there is a $<\kappa$-closed non-principal ultrafilter on κ.

An ultrafilter witnessing the measurability of κ is called a **measure** on κ.

The following result describes measurable cardinals in terms of elementary embeddings.

Theorem 9.5. *Let κ be any cardinal. The following are equivalent.*

(i) κ is measurable.

(ii) There is[11] *a transitive class M and a class j so that $j\colon V \to M$ is a elementary embedding with* $\mathrm{crit}(j) = \kappa$.

We will first only show one direction.

[11] We note that this is not a first order $\in$-statement. A similar but equivalent statement is the existence of a elementary embedding $j\colon H_{\kappa^+} \to M$ with M a transitive set and $\mathrm{crit}(j) = \kappa$.

Proof of (ii) ⇒ (i). Let $j\colon V \to M$ witness (ii). First note that $\kappa > \omega$ as every natural number as well as ω are Δ_0-definable and hence fixed by j. By Proposition 8.6, κ is regular and hence κ is uncountable. We define $\mathcal{U}$ via

$$\mathcal{U} = \{A \subseteq \kappa \mid \kappa \in j(A)\}.$$

Clearly $\emptyset \notin \mathcal{U}$ and since $\kappa = \mathrm{crit}(j)$, $\kappa < j(\kappa)$ so that $\kappa \in \mathcal{U}$. If $A \subseteq B \subseteq \kappa$ and $A \in \mathcal{U}$ then by elementarity of j, $\kappa \in j(A) \subseteq j(B)$ and hence $B \in \mathcal{U}$.

If $A \subseteq \kappa$ then $\kappa = A \cup (\kappa \setminus A)$ and thus $j(\kappa) = j(A) \cup j(\kappa \setminus A)$ by elementarity of j. It follows that $\kappa \in j(A)$ or $\kappa \in j(\kappa \setminus A)$, so one of these sets is in $\mathcal{U}$.

It remains to show that $\mathcal{U}$ is $< \kappa$-closed. Suppose that $\vec{A} \coloneqq \langle A_i \mid i < \beta\rangle$ is a sequence of sets in $\mathcal{U}$ and $\beta < \kappa$.

Claim 9.6. $j(\vec{A}) = \langle j(A_i) \mid i < \beta\rangle$.

Proof. $\vec{A}$ is a sequence of length β and hence

$$M \models \text{“}j(\vec{A}) \text{ is a sequence of length } j(\beta) = \beta\text{”}$$

by elementarity. This is a Δ_0-statement and hence true in V as M is transitive. So $j(\vec{A}) = \langle \tilde{A}_i \mid i < \beta\rangle$ for some $\tilde{A}_i$'s. Further, for $i < \beta$, the i-th entry in $\vec{A}$ is A_i and hence by elementarity of j, the $j(i) = i$-th entry in $j(\vec{A})$ is $j(A_i)$ (holds in M and hence in V). So $\tilde{A}_i = j(A_i)$ and the claim follows. □

Hence, by elementarity of j, we have

$$j\left(\bigcap_{i<\beta} A_i\right) = \bigcap_{i<\beta} j(A_i) \ni \kappa$$

and $\bigcap_{i<\beta} A_i \in \mathcal{U}$. □

The ultrafilter $\mathcal{U}$ above is the **ultrafilter/measure derived from** j. It has an additional nice property.

Definition 9.7. A filter $\mathcal{F}$ on a cardinal κ is **normal** if it is closed under diagonal intersections. That is, if $\langle A_i \mid i < \kappa\rangle$ is a sequence of elements of $\mathcal{F}$ then $\triangle_{i<\kappa} A_i \in \mathcal{F}$.

We remark that a filter $\mathcal{F}$ on κ is normal iff the following property reminiscent of Fodor's lemma holds: For any $A \in \mathcal{F}^+ \coloneqq \{B \subseteq \kappa \mid \forall C \in \mathcal{F}\ B \cap C \neq \emptyset\}$ and regressive $f\colon A \to \kappa$, there is $B \subseteq A$, $B \in \mathcal{F}^+$ so that $f \restriction B$ is constant.

Lemma 9.8. *The measure derived from a non-trivial elementarity embedding $j\colon V \to M$ with M transitive is normal.*

Proof. Let $\kappa = \mathrm{crit}(j)$ and $\mathcal{U}$ the derived measure. The argument is almost the same as showing that $\mathcal{U}$ is $<\kappa$-closed. Suppose $\vec{A} \coloneqq \langle A_i \mid i < \kappa\rangle$ so that all $A_i \in \mathcal{U}$. As before, we see that $j(\vec{A}) = \langle \tilde{A}_i \mid i < j(\kappa)\rangle$ with $\tilde{A}_i = j(A_i)$ for $i < \kappa$. But then $\kappa \in \bigcap_{i<\kappa} \tilde{A}_i$ so that $\kappa \in \triangle_{i<j(\kappa)} \tilde{A}_i = j(\triangle_{i<\kappa} A_i)$. □

9.1 The Ultrapower Construction

The proof of $(i) \Rightarrow (ii)$ of Theorem 9.5 will occupy us for the rest of this section. The main tool we will develop is that of an ultrapower construction.

Definition 9.9. Suppose M is a transitive class model of (sufficiently much of) ZFC and $\mathcal{U}$ is a M-ultrafilter on $X \in M$, i.e. $\mathcal{U}$ is a filter which measures every subset of X in M. The **ultrapower** $\mathrm{Ult}(M,\mathcal{U})$ **of** M **by** $\mathcal{U}$ is defined as follows: For $f, g\colon \kappa \to M$, $f \sim_{\mathcal{U}} g$ iff $\{\alpha < \kappa \mid f(\alpha) = g(\alpha)\} \in \mathcal{U}$. This is an equivalence relation on ${}^{\kappa}M \cap M$ and we denote the equivalence class of f by $[f]_{\mathcal{U}}$. We let $f E_{\mathcal{U}} g$ iff

$$\{\alpha < \kappa \mid f(\alpha) \in g(\alpha)\} \in \mathcal{U}$$

and note that this is well-defined. We then set[12]

$$\mathrm{Ult}(M,\mathcal{U}) = (({}^{\kappa}M \cap M)/\sim_{\mathcal{U}}, E_{\mathcal{U}}).$$

Lemma 9.10. *Suppose $\mathcal{U}$ is a measure on a measurable cardinal κ. Then* $\mathrm{Ult}(V,\mathcal{U})$ *is well-founded.*

Proof. Suppose not. Then, making use of DC, we can find a sequence $(f_n)_{n<\omega}$ of functions so that $[f_{n+1}]_{\mathcal{U}} E_{\mathcal{U}} [f_n]_{\mathcal{U}}$ for all $n < \omega$. We have

$$A_n = \{\alpha < \kappa \mid f_{n+1}(\alpha) \in f_n(\alpha)\} \in \mathcal{U}$$

so that $\bigcap_{n<\omega} A_n \in \mathcal{U}$ as $\mathcal{U}$ is $<\kappa$-closed and κ is uncountable. But if α is any member of $\bigcap_{n<\omega} A_n$ then $(f_n(\alpha))_{n<\omega}$ is an infinite $\in$-descending chain, contradiction. $\square$

Theorem 9.11 (Łoś)**.** *Suppose that M is a transitive class model of* ZFC *and $\mathcal{U}$ is a M-ultrafilter on $\kappa \in M$. For φ a $\in$-formula and $f_0, \dots, f_n \in {}^{\kappa}M \cap M$, we have*

$$\mathrm{Ult}(M,\mathcal{U}) \models \varphi([f_0]_{\mathcal{U}}, \dots, [f_n]_{\mathcal{U}}) \Leftrightarrow \{\alpha < \kappa \mid M \models \varphi(f_0(\alpha), \dots, f_n(\alpha))\} \in \mathcal{U}.$$

Proof. We argue by induction along the complexity of φ. We will assume $n = 0$ for notational convenience (note that the general case follows from this nonetheless).

φ **is atomic:** This case follows immediately from the definition of $\sim_{\mathcal{U}}$ and $E_{\mathcal{U}}$.

$\varphi = \psi_0 \wedge \psi_1$**:** We have

$$\begin{aligned}
&\mathrm{Ult}(M,\mathcal{U}) \models \varphi([f]_{\mathcal{U}})\\
\Leftrightarrow&\mathrm{Ult}(M,\mathcal{U}) \models \psi_0([f]_{\mathcal{U}}) \wedge \mathrm{Ult}(M,\mathcal{U}) \models \psi_1([f]_{\mathcal{U}})\\
\Leftrightarrow&\{\alpha < \kappa \mid M \models \psi_0(f(\alpha))\} \in \mathcal{U} \wedge \{\alpha < \kappa \mid M \models \psi_1(f(\alpha))\} \in \mathcal{U}\\
\Leftrightarrow&\{\alpha < \kappa \mid M \models \psi_0(f(\alpha))\} \cap \{\alpha < \kappa \mid M \models \psi_1(f(\alpha))\} \in \mathcal{U}\\
\Leftrightarrow&\{\alpha < \kappa \mid M \models \varphi(f(\alpha))\} \in \mathcal{U}.
\end{aligned}$$

[12]Technically, $[f]_{\mathcal{U}}$ is a proper class if M is a proper class. In this case we make use of Scott's trick again and replace $[f]_{\mathcal{U}}$ by the set $[f]_{\mathcal{U}} \cap V_\alpha$ where α is least so that this set is non-empty.

$\varphi = \neg\psi$**:** Since $A \in \mathcal{U}$ iff $\kappa \setminus A \notin \mathcal{U}$ for $A \subseteq \kappa$, we calculate

$$\begin{aligned}
&\mathrm{Ult}(M,\mathcal{U}) \models \neg\psi([f]_{\mathcal{U}}) \\
\Leftrightarrow&\neg(\mathrm{Ult}(M,\mathcal{U}) \models \psi([f]_{\mathcal{U}})) \\
\Leftrightarrow&\{\alpha < \kappa \mid M \models \psi(f(\alpha))\} \notin \mathcal{U} \\
\Leftrightarrow&\{\alpha < \kappa \mid M \models \neg\psi(f(\alpha))\} \in \mathcal{U}.
\end{aligned}$$

$\varphi = \exists x\psi$**:** First suppose $\mathrm{Ult}(M,\mathcal{U}) \models \exists x\ \psi(x, [f]_{\mathcal{U}})$ as witnessed by $[g]_{\mathcal{U}}$. We thus have

$$\{\alpha < \kappa \mid M \models \exists x\ \psi(f(\alpha))\} \supseteq \{\alpha < \kappa \mid M \models \psi(g(\alpha), f(\alpha))\} \in \mathcal{U}$$

and hence $\{\alpha < \kappa \mid M \models \psi(g(\alpha), f(\alpha))\} \in \mathcal{U}$.

Next, assume $\{\alpha < \kappa \mid \exists x\ \psi(x, f(\alpha))\} \in \mathcal{U}$. As $M \models \mathrm{ZFC}$, there is a function $g\colon \kappa \to M$, $g \in M$ so that if $M \models \exists x\ \psi(x, f(\alpha))$ then $M \models \psi(g(x), f(x))$. Hence

$$\{\alpha < \kappa \mid M \models \psi(g(\alpha), f(\alpha))\} = \{\alpha < \kappa \mid M \models \exists x\ \psi(x, f(\alpha))\} \in \mathcal{U}$$

and it follows that $\mathrm{Ult}(M,\mathcal{U}) \models \psi([g]_{\mathcal{U}}, [f]_{\mathcal{U}})$.

□

It follows immediately that $\mathrm{Ult}(M,\mathcal{U})$ and $(M, \in)$ are elementarily equivalent, that is $\mathrm{Ult}(M,\mathcal{U}) \models \varphi$ iff $M \models \varphi$ for any $\in$-sentence φ. In particular, $\mathrm{Ult}(M,\mathcal{U})$ is a model of ZFC and is extensional. It is easy to see that $\mathrm{Ult}(M,\mathcal{U})$ is set-like.

Convention *Suppose that* $\mathrm{Ult}(M,\mathcal{U})$ *is well-founded, e.g. if* $\mathcal{U}$ *is* $<\omega_1$*-closed. Then we identify* $\mathrm{Ult}(M,\mathcal{U})$ *with its transitive collapse.*

Definition 9.12. The **ultrapower embedding** given by M and $\mathcal{U}$ is

$$j^M_{\mathcal{U}}\colon M \to \mathrm{Ult}(M,\mathcal{U}),\ x \mapsto [c_x]_{\mathcal{U}}$$

where $c_x\colon \kappa \to \{x\}$ is the constant function with value x.

Usually, we write $j_{\mathcal{U}}$ (or just j) for $j^M_{\mathcal{U}}$ if M is clear from context, particularly if $M = V$.

Proposition 9.13. *The ultrapower embedding* $j_{\mathcal{U}}$ *is an elementary embedding.*

Proof. For any $\in$-formula φ and parameters $x_0, \ldots, x_n \in M$, we have

$$\begin{aligned}
&M \models \varphi(x_0, \ldots, x_n) \\
\Leftrightarrow&\{\alpha < \kappa \mid M \models \varphi(c_{x_0}(\alpha), \ldots, c_{x_n}(\alpha))\} \in \mathcal{U} \\
\Leftrightarrow&\mathrm{Ult}(M,\mathcal{U}) \models \varphi([c_0]_{\mathcal{U}}, \ldots, [c_n]_{\mathcal{U}})
\end{aligned}$$

by Łoś's theorem. □

Proof of Theorem 9.5 (i) ⇒ (ii). Suppose $\mathcal{U}$ is a measure on a measurable cardinal κ. Then $j_{\mathcal{U}} \to \mathrm{Ult}(V,\mathcal{U})$ is an elementary embedding and $\mathrm{Ult}(V,\mathcal{U})$ is a transitive class. It remains to show that $\mathrm{crit}(j_{\mathcal{U}}) = \kappa$.

Claim 9.14. $j_{\mathcal{U}} \upharpoonright \kappa = \mathrm{id}_\kappa$.

Proof. Clearly, $j_{\mathcal{U}}(\alpha) \geq \alpha$. Suppose on the other hand that $[f]_{\mathcal{U}} < [c_\alpha]_{\mathcal{U}} = j_{\mathcal{U}}(\alpha)$. We have

$$\bigcup_{\beta<\alpha} g^{-1}[\{\beta\}] = \{\beta < \kappa \mid g(\beta) < \alpha\} \in \mathcal{U}.$$

As $\mathcal{U}$ is $<\kappa$-closed, we find that $g^{-1}[\{\beta\}] \in \mathcal{U}$ for some $\beta < \alpha$ and hence $[g]_{\mathcal{U}} = j_{\mathcal{U}}(\beta) = \beta$ by induction. We have shown that $j_{\mathcal{U}}(\alpha) = \alpha$. □

On the other hand, we have $\alpha = j_{\mathcal{U}}(\alpha) = [c_\alpha]_{\mathcal{U}} < [\mathrm{id}_\kappa]_{\mathcal{U}} < [c_\kappa]_{\mathcal{U}} = j_{\mathcal{U}}(\kappa)$ for all $\alpha < \kappa$. Hence $j_{\mathcal{U}}(\kappa) > \kappa$.

□

Corollary 9.15. *Any measurable cardinal is a Mahlo cardinal and a limit of Mahlo cardinals.*

Proof. Let $j\colon V \to M$ be a non-trivial elementary embedding, M transitive and $\kappa = \mathrm{crit}(j)$. By Proposition 8.6, κ is regular. Suppose that $\lambda < \kappa$ and $\kappa \leq 2^\lambda$ as witnessed by a surjection $f\colon \mathcal{P}(\lambda) \to \kappa$. Note that since $\mathrm{crit}(j) > \lambda$, $j(X) = X$ for $X \subseteq \lambda$ and $\mathcal{P}(\lambda)^M = \mathcal{P}(\lambda)$. It follows that $j(f)$ is a surjection from $\mathcal{P}(\lambda)$ onto $j(\kappa)$. However, for $X \subseteq \lambda$, we have

$$j(f)(X) = j(f)(j(X)) = j(f(X)) = f(X)$$

since $f(X) < \mathrm{crit}(j)$ and j is elementary. But then e.g. $\kappa \notin \mathrm{ran}(j(f))$, contradiction.

The argument that κ is Mahlo is similar to the proof of $(iii) \Rightarrow (i)$ of Lemma 8.18: If $C \subseteq \kappa$ is a club then $j(C) \subseteq j(\kappa)$ is club, $C = j(C) \cap \kappa$ and hence $\kappa \in j(C)$. As κ is regular, κ is regular in M. By elementarity, C has a regular element.

Next, being Mahlo is a Π_1-property and hence, as $M \subseteq V$, $M \models$ "κ is Mahlo". For $X = \{\lambda < \kappa \mid \lambda \text{ is Mahlo}\}$, it follows that $X \in \mathcal{U}_j$, the ultrafilter derived from j. Since $\mathcal{U}_j$ is a measure on κ, X is unbounded in κ (and as $\mathcal{U}_j$ is normal, X is even stationary in κ). □

Corollary 9.16. *If $V = L$ then there are no measurable cardinals.*

Proof. Suppose towards a contradiction that there is a measurable cardinal and let κ be the least such. Suppose that $j\colon L \to M$ is a definable non-trivial elementary embedding with $\mathrm{crit}(j) = \kappa$. As $L = L^M \subseteq M \subseteq L$, we have $M = L$. But κ is the least measurable so that $\kappa < j(\kappa)$ is the least measurable of $M = L$. This is clearly a contradiction. □

We will now compute some basic properties of the ultrapower of V by a measure.

Lemma 9.17. *Suppose $\mathcal{U}$ is a measure on κ. Let $M = \mathrm{Ult}(V,\mathcal{U})$ and $j = j_{\mathcal{U}}$.*

(i) $V^M_{\kappa+1} = V_{\kappa+1}$.

(ii) $\mathcal{U} \notin M$ *and hence* $V^M_{\kappa+2} \subsetneq V_{\kappa+2}$.

(iii) $(\kappa^+)^M = \kappa^+$

(iv) $2^\kappa < j(\kappa) < (2^\kappa)^+$.

(v) *M is closed under sequences of length κ, i.e.* ${}^\kappa M \subseteq M$.

Note that $j(\kappa)$ is **not** a cardinal in V, even though κ is measurable in M.

Proof. (i): As κ is inaccessible, any element of V_κ has size $<\kappa$. It follows by induction that $j \restriction V_\kappa = \mathrm{id}_{V_\kappa}$, in particular $V_\kappa = V^M_\kappa$. If $A \in V_{\kappa+1}$ then $A \subseteq V_\kappa$ and hence $A = j(A) \cap V_\kappa \in M$. Hence $V_{\kappa+1} \subseteq M$.

(iii): It follows from (i) that $\mathcal{P}(\kappa)^M = \mathcal{P}(\kappa)$ and hence M contains all wellorders on κ (and agrees with V about which linear orders on κ are wellorders). As $\kappa^+ = \sup\{\mathrm{otp}(\prec) \mid \prec \text{ is a wellorder on } \kappa\}$, the claim follows.

(iv): Any $\alpha < j(\kappa)$ is represented by a function $f\colon \kappa \to \kappa$. There are 2^κ-such functions so that $|j(\kappa)| \leq 2^\kappa$ and $j(\kappa) < (2^\kappa)^+$. On the other hand, $\mathcal{P}(\kappa)^M = \mathcal{P}(\kappa)$ so that $2^\kappa \leq (2^\kappa)^M < j(\kappa)$, where the second inequality follows from the fact that $M \models$ "κ is measurable" and $\kappa < j(\kappa)$.

(ii): Suppose toward a contradiction that $\mathcal{U} \in M$. By (i), $H^M_{\kappa^+} = H_{\kappa^+}$. Hence

$$(\mathrm{Ult}(H_{\kappa^+},\mathcal{U}))^M = \mathrm{Ult}(H_{\kappa^+},\mathcal{U})$$

and if $k\colon H_{\kappa^+} \to \mathrm{Ult}(H_{\kappa^+},\mathcal{U})$ is the resulting elementary embedding, the argument of (iv) shows that $|k(\kappa)|^M \leq (2^\kappa)^M$. On the other hand, H_{κ^+} contains all functions $f\colon \kappa \to \kappa^+ = \mathrm{Ord} \cap H_{\kappa^+}$ so it follows that $k \restriction \kappa^+ = j \restriction \kappa^+$ (in fact $k = j \restriction H_{\kappa^+}$). But $j(\kappa)$ is measurable, in particular inaccessible in M, contradiction. As $\mathcal{U} \in V_{\kappa+2}$, $V^M_{\kappa+2} \subsetneq V_{\kappa+2}$.

(v): Suppose that $([f_\alpha]_{\mathcal{U}})_{\alpha<\kappa}$ is a sequence of elements of M. Define a function $g\colon \kappa \to V$ via

$$g(\alpha)\colon \alpha \to V,\ g(\alpha)(i) = f_i(\alpha).$$

It follows from Łoś's theorem that $[g]_{\mathcal{U}}$ is a function with domain $[\mathrm{id}_\kappa]_{\mathcal{U}} \geq \kappa$.

Claim 9.18. $[g]_{\mathcal{U}} \restriction \kappa = \langle [f_\alpha]_{\mathcal{U}} \mid \alpha < \kappa\rangle$.

Proof. For $\alpha < \kappa$, we have

$$\{\beta < \kappa \mid f_\alpha(\beta) = g(\beta)(c_\alpha(\beta))\} = \{\beta < \kappa \mid f_\alpha(\beta) = g(\beta)(\alpha)\} = (\alpha+1,\kappa) \in \mathcal{U}$$

and hence $[f_\alpha]_{\mathcal{U}} = [g]_{\mathcal{U}}([c_\alpha]_{\mathcal{U}}) = [g]_{\mathcal{U}}(\alpha)$ follows once again from Łoś's theorem. □

The restriction $[g]_{\mathcal{U}} \restriction \kappa$ is obviously in M, so (v) follows. □

9.2 The model $L[\mathcal{U}]$

The branch of Inner Model Theory constructs canonical inner models for large cardinals. Here, "canonical" has no precise meaning, it is a "you know it when you see it" kind of deal. We have seen that if κ is inaccessible/Mahlo then $L \models$ "κ is inaccessible/Mahlo" and indeed L is the canonical Model for these large cardinals.

However, we have also proven that L has no measurable cardinals. In some sense, L is too small to be able to support measurable cardinals so we need to go to a larger model instead. Luckily, the canonical model for a measurable cardinal is relatively simple to construct, it is simply $L[\mathcal{U}]$ where $\mathcal{U}$ is a (normal) measure.

Lemma 9.19. *Let $\mathcal{U}$ be a normal measure on κ and $\mathcal{V} = \mathcal{U} \cap L[\mathcal{U}]$. Then $L[\mathcal{U}] \models$ "$\mathcal{V}$ is a normal measure on κ.*

Proof. Clearly, $\kappa \in L[\mathcal{U}] \cap \mathcal{U} = \mathcal{V}$ and $\emptyset \notin \mathcal{V}$. If $A \in \mathcal{V}$, $A \subseteq B \subseteq \kappa$ and $B \in L[\mathcal{U}]$ then $B \in \mathcal{U}$ and hence $B \in \mathcal{V}$. Now suppose $\langle A_i \mid i < \alpha \rangle \in L[\mathcal{U}]$ is a sequence of elements of $\mathcal{V}$ and $\alpha < \kappa$. Then $\bigcap_{i<\alpha} A_i \in \mathcal{U}$ as $\mathcal{U}$ is $<\kappa$-closed and $\bigcap_{i<\alpha} A_i \in L[\mathcal{U}]$ as $L[\mathcal{U}]$ is a model of ZFC (and the intersection is Δ_0-definable). Now suppose $A \subseteq \kappa$, $A \in L[\mathcal{U}]$. Then one of $A, \kappa \setminus A$ is in $\mathcal{U}$, hence in $\mathcal{V}$. We have shown that $\mathcal{V}$ is a $<\kappa$-closed normal ultrafilter on κ in $L[\mathcal{U}]$. Clearly, $\mathcal{V}$ is non-principal as well and proving that $\mathcal{V}$ is normal in $L[\mathcal{U}]$ is similar to proving the $<\kappa$-closure. □

So κ is a measurable cardinal in $L[\mathcal{U}]$. It looks like the model $L[\mathcal{U}]$ depends a lot on the normal measure $\mathcal{U}$, but it turns out that this is not the case. If $\mathcal{U}'$ is another normal measure on κ, then $L[\mathcal{U}] = L[\mathcal{U}']$. Moreover, if $\kappa < \lambda$ and $\mathcal{U}'$ is a normal measure on λ then $L[\mathcal{U}] \neq L[\mathcal{U}']$, but there is still an elementary embedding $j \colon L[\mathcal{U}] \to L[\mathcal{U}']$. In particular, the theory of $L[\mathcal{U}]$ does not depend on $\mathcal{U}$ at all. If time permits, we will prove these properties of $L[\mathcal{U}]$ later. For now, we show GCH in $L[\mathcal{U}]$.

Theorem 9.20. *Suppose $\mathcal{U}$ is a normal measure on κ. Then $L[\mathcal{U}] \models \mathrm{GCH}$.*

In general, the existence of a measurable cardinal κ has very little consequence on the continuum function. For example, the function $\lambda \mapsto 2^\lambda$ can be any function following the restrictions we already outlined as long as $2^\lambda < \kappa$ for all $\lambda < \kappa$.

Definition 9.21. Suppose that X is a set. For $a \in A \in X$, we define

$$X(a) = \{f(a) \mid f \in X \text{ is a function with } \mathrm{dom}(f) = A\}.$$

We will only be interested in $X(a)$ if X is a (not necessarily transitive) model of ZFC$^-$. In this case, the choice of A such that $a \in A \in X$ does not matter, so we did not make it explicit in the notation. Note that in this case also $X \subseteq X(a)$ and $a \in X(a)$.

Proposition 9.22. *Suppose M is a transitive model of* ZFC$^-$*, $X \prec M$ and $a \in A \in X$. Then $X(a) \prec M$.*

Proof. By Tarski's criterion, it suffices to show that whenever $M \models \exists x\ \varphi(x, b)$ for some $b \in X(a)$ then $\exists x \in X(a)\ M \models \varphi(x, b)$. So assume the premise holds true. We can find $f\colon A \to M$, $f \in X$ so that $b = f(a)$. As $M \models \mathrm{ZFC}^-$, there is a function[13] $g\colon A \to M$, $g \in M$ so that for $c \in A$ such that $M \models \exists x\ \varphi(x, f(c))$ then $M \models \varphi(g(c), f(c))$. By elementarity of X in M, we may assume w.l.o.g. that $g \in X$. But then $g(a) \in X(a)$ and $M \models \varphi(g(a), f(a))$. □

Lemma 9.23. *Suppose $\mathcal{U}$ is a normal measure on κ, M is a transitive model of* ZFC$^-$ *with $\mathcal{U} \in M$ and $X \prec M$ is an elementary substructure such that $\mathcal{U} \in X$ and $|X| < \kappa$.*
Then there is an $\alpha \in \kappa$ so that[14] $X(\alpha) \cap \alpha = X \cap \kappa$. In fact,

$$\{\alpha < \kappa \mid X(\alpha) \cap \alpha = X \cap \kappa\} \in \mathcal{U}.$$

Proof. Let $A = \bigcap(\mathcal{U} \cap X)$ and note that $A \in \mathcal{U}$ as $\mathcal{U}$ is $<\kappa$-closed and $|X| < \kappa$. Let $\alpha \in A$, we will show that $X(\alpha) \cap \alpha = X \cap \kappa$. As $(\beta, \kappa) \in X \cap \mathcal{U}$ for all $\beta \in X \cap \kappa$, we see that $\beta < \alpha$ and hence $\sup(X \cap \kappa) \leq \alpha$. Now suppose $\beta < \alpha$ and $\beta \in X(\alpha)$. Then there is a function $f\colon \kappa \to M$ with $f \in X$ and $\beta = f(\alpha)$. We may now assume that $f\colon \kappa \to \kappa$ is a regressive function, as otherwise we may replace f by the regressive function f' given by

$$f'(\gamma) = \begin{cases} f(\gamma) & \text{if } f(\gamma) < \gamma \\ 0 & \text{otherwise.} \end{cases}$$

As $\mathcal{U}$ is normal, there is some $A \in \mathcal{U}$ so that $f \restriction A$ is constant. Note that $A \in M$ as M is transitive and hence by elementarity of X in M, there is some $B \in \mathcal{U} \cap X$ so that $f \restriction B$ is constant. As $\alpha \in B$, β is the unique element of $f[B] \in X$ and hence $\beta \in X$. □

We will use the following elementary fact about normal measures.

Proposition 9.24. *Suppose $\mathcal{U}$ is a normal measure on κ. Then $\mathcal{C}_\kappa \subseteq \mathcal{U}$.*

Proof. Let $C \subseteq \kappa$ be a club. The function $f\colon (\kappa \setminus C) \setminus \min(C) \to \kappa$, $f(\alpha) = \max(C \cap \alpha)$ is well-defined and regressive since C is closed. If $B \subseteq \mathrm{dom}(f)$ so that $f \restriction B$ is constant then B is bounded in κ as C is unbounded. But then $B \notin \mathcal{U}$, so $\mathrm{dom}(f) \notin \mathcal{U}$ as $\mathcal{U}$ is normal. Hence $C \in \mathcal{U}$. □

Corollary 9.25. *Suppose $\mathcal{U}, \kappa, M, X$ are as in the statement of Lemma 9.23. Then for any $\delta < \kappa$ there is an elementary substructure $Y \prec M$ so that*

(*i*) $X \subseteq Y$,

(*ii*) $Y \cap \kappa \in \mathcal{U}$ *and*

[13] Recall that we made use of a similar function in the proof of Łoś's theorem.
[14] In other words, $\alpha \geq \sup(X \cap \kappa)$ and adding α to X does not add any new ordinals $<\alpha$.

(*iii*) $Y \cap \delta = X \cap \delta$.

Proof. Build $\langle X_i \mid i < \kappa \rangle$ by recursion via

- $X_0 = X$,
- $X_{i+1} = X_i(\alpha_i)$ where α_i is the least $\alpha \geq \delta$ so that $X_i(\alpha) \cap \alpha = X_i \cap \kappa$.
- $X_i = \bigcup_{j<i} X_j$ if $i \in \mathrm{Lim} \cap \kappa$.

By induction, we see that $|X_i| < \kappa$ and that α_i exists for all $i < \kappa$, so the construction does not break down. Also, $X_i \prec M$ for all $i < \kappa$ by Proposition 9.22 and Tarski's chain lemma. We now set $Y = \bigcup_{i<\kappa} X_i$. It is clear that $X \subseteq Y \prec M$ and $Y \cap \delta = X \cap \kappa$, so it remains to show that $Y \cap \kappa \in \mathcal{U}$. Suppose this is not the case.

Let $f\colon \kappa \to \kappa$ be given by $f(i) = \sup(X_i \cap \kappa)$ and let C_f be the club of closure points of f. By Proposition 9.24,

$$B \coloneqq (\kappa \setminus Y) \cap C_f \cap \mathrm{Lim} \setminus \delta \in \mathcal{U}.$$

Note that if $i \in B$ then $i = \sup X_i \cap \kappa$, hence $\alpha_i \geq i$. But $i \notin Y$, so $\alpha_i > i$. As we chose α_i minimally, this can only happen because $X_i(i) \cap i \neq X_i \cap i$, so there is an ordinal $\gamma_i < i$ which is new in $X_i(i)$. The map $i \mapsto \gamma_i$ is regressive on B, so constant on some $D \subseteq B$ with $D \in \mathcal{U}$ with value some $\gamma_* < \kappa$. For each $i \in D$, i is a limit so $X_i = \bigcup_{j<i} X_j$ so there is a minimal $j_i < i$ such that even $\gamma_* \in X_{j_i}(\alpha)$. The map $i \mapsto j_i$ is once again a regressive function, so there is some $i_* < \kappa$ and $E \subseteq D$, $E \in \mathcal{U}$ so that for all $\alpha \in E$, we have $\gamma_* \in X_{i_*}(\alpha)$. But $\gamma_* \notin X_{i_*}$, so this contradicts Lemma 9.23.

□

Lemma 9.26. *Suppose $U \subseteq \mathcal{P}(\alpha)$ for some ordinal α. Then "GCH above α" holds true in $L[U]$, i.e.*

$$L[U] \models \forall \lambda \in \mathrm{Card} \setminus \alpha \; 2^\lambda = \lambda^+.$$

Proof. For convenience, let us assume that $U \in L[U]$, otherwise replace U by $U \cap L[U]$. Suppose θ is a sufficiently large limit ordinal so that $U \in L_\theta[U]$.

Claim 9.27. *If $X \prec L_\theta[U]$ with $\alpha + 1 \subseteq X$ then $X \cong L_\gamma[U]$ for some $\gamma \leq \theta$.*

Proof. Let $\pi\colon X \to M$ be the Mostowski collapse map. Following the proof of the condensation lemma, there is some $\gamma \leq \theta$ so that $M = L_\gamma[\pi(U)]$. Note that $\pi \restriction \lambda = \mathrm{id}_\lambda$ since $\lambda \subseteq X$ and hence $\pi(A) = \pi[A] = A$ for $A \in U \cap X$. It follows that $\pi(U) = U \cap X = U \cap M$ and hence $M = L_\gamma[U \cap M]$. By induction on $\beta \leq \gamma$, we see that $L_\beta[U \cap M] = L_\beta[U]$ and hence $M = L_\gamma[U]$. □

We can now repeat the proof of Theorem 7.15. □

Proof of Theorem 9.20. Let us work in $L[\mathcal{U}]$. For notational convenience, assume $\mathcal{U} \in L[\mathcal{U}]$. Let λ be an infinite cardinal; it is our duty to prove $2^\lambda = \lambda^+$. If $\lambda \geq \kappa$ then $2^\lambda = \lambda^+$ by Lemma 9.26. So let us deal with the more difficult case $\lambda < \kappa$.

Suppose toward a contradiction that $2^\lambda > \lambda$. Let $<_{L[\mathcal{U}]}$ be the canonical global wellorder on $L[\mathcal{U}]$ (defined analogously to $<_L$). By our assumption, there is some $A \subseteq \lambda$ which is the λ^+-th subset of λ with respect to $<_{L[\mathcal{U}]}$. Let $\theta > \kappa$ be a regular cardinal such that $\mathcal{U} \in L_\theta[\mathcal{U}]$ and note that[15] $L_\theta[\mathcal{U}] \models \mathrm{ZFC}^-$. Next, let $X \prec L_\theta[\mathcal{U}]$ so that

(*i*) $\lambda + 1 \subseteq X$, $A, \mathcal{U} \in X$ and

(*ii*) $|X| = \lambda$.

By Corollary 9.25, there is some $Y \prec L_\theta[\mathcal{U}]$ with $X \subseteq Y$, $Y \cap \kappa \in \mathcal{U}$ and $Y \cap \lambda^+ = X \cap \lambda^+$. It follows that $Y \cap \lambda^+$ is bounded in λ^+. Let $\pi\colon Y \to M$ be the Mostowski collapse map of Y.

Claim 9.28. *$B \coloneqq \{\alpha < \kappa \mid \pi(\alpha) = \alpha\} \in \mathcal{U}$.*

Proof. Note that $\pi(\alpha) \leq \alpha$ for all $\alpha \in Y \cap \kappa$, so if $B \notin \mathcal{U}$ then necessarily $\{\alpha < \kappa \mid \pi(\alpha) < \alpha\} \in \mathcal{U}$ and π is regressive on this set. But π is injective so cannot possibly be constant on a set in $\mathcal{U}$. This contradicts the normality of $\mathcal{U}$. □

Claim 9.29. *$\pi(\mathcal{U}) = \mathcal{U} \cap M$.*

Proof. As $\mathcal{U}$ is an ultrafilter, it suffices to proof $\pi(C) \in \mathcal{U}$ for $C \in \mathcal{U} \cap Y$. By definition of the Mostowski-collapse, $\pi(C) = \pi[C \cap Y]$ and since

$$\pi(C) \supseteq \pi[C \cap Y] \cap B = C \cap Y \cap B \in \mathcal{U},$$

we find $\pi(C) \in \mathcal{U}$. □

Claim 9.30. *$M = L_\gamma[\mathcal{U}]$ for some ordinal γ.*

Proof. As in the proof of the Condensation Lemma 7.11, we see that $M = L_\gamma[\pi(\mathcal{U})]$ for some ordinal γ. Hence $M = L_\gamma[M \cap \mathcal{U}] = L_\gamma[\mathcal{U}]$. □

As $\lambda + 1 \subseteq Y$, $\pi \restriction \lambda + 1$ is the identity and it follows that $\pi(A) = A$. Since $Y \cap \lambda^+$ is bounded, $\pi(\lambda^+) = \pi[\lambda^+ \cap Y]$ is an ordinal of size $<\lambda^+$, so $\pi(\lambda^+) < \lambda^+$. On the other hand, similar as to $<_L$, the wellorder $<_{L[\mathcal{U}]}$ is local in the sense that $(<_{L[\mathcal{U}]})^{L_\gamma[\mathcal{U}]} = <_{L[\mathcal{U}]} \restriction L_\gamma[\mathcal{U}]$. By the elementarity of π,

$$L_\gamma[\mathcal{U}] \models \text{“}\pi(A) = A \text{ is the } \pi(\lambda^+)\text{-th subset of } \pi(\lambda) = \lambda\text{”}.$$

Further, it is clear form the definition of $<_{L[\mathcal{U}]}$ that if $A' <_{L[\mathcal{U}]} A$ then $A' \in L_\gamma[\mathcal{U}]$. So A really is the $\pi(\lambda^+)$-th subset of λ and consequently $\pi(\lambda^+) = \lambda^+$, contradiction. □

[15] The most difficult part is showing $L_\theta[\mathcal{U}] \models$(Collection). One can use the $L[\mathcal{U}]$-hierarchy and the regularity of θ to find the required witnesses

9.3 The Kunen inconsistency

We have mentioned earlier that there is no elementary embedding $j\colon V \to V$. However, we should make precise what exactly this means. After all, it is consistent (relative to some large cardinals) that there is an "external" nontrivial elementary $j\colon V \to V$.

Proposition 9.31. *Suppose there is a measurable cardinal. Then there is a non-trivial elementary embedding $j\colon L \to L$.*

Proof. Let $i\colon V \to M$ be a non-trivial elementary embedding and M transitive. Then $L = L^M$ and since L is definable without parameters, $j = i \restriction L\colon L \to L$ is a non-trivial elementary embedding. $\square$

In the example above, L does not "see" the embedding j, e.g. j is not definable over L. There are several ways to express that an elementary embedding $j\colon V \to V$ is "seen" or "close" to V and we will lead some of them to a contradiction.

Definition 9.32. (*i*) Suppose $\mathcal{L}$ is any first order language extending $\{\in\}$. The theory $\mathcal{L}$-ZF(C) is the $\mathcal{L}$-theory defined exactly as ZF(C), except that the instances of all schemes range over all $\mathcal{L}$-formulas instead of only over all $\in$-formulas.

(*ii*) Suppose $\mathcal{M} = (M, \in)$ is a model $\in$-model and $A \subseteq M$. Then $\mathcal{M}$ is a model of A-ZF(C) if the following holds: Let $\mathcal{M}_A = (M, \in, A)$ be the expansion of $\mathcal{M}$ to a model in the language $\mathcal{L} = \{\in, \dot{A}\}$ where $\dot{A}$ is a unary relation interpreted as A in $\mathcal{M}_A$. Then $\mathcal{M}_A$ is a model of $\mathcal{L}$-ZF(C).

For example, if $\mathcal{M}$ is any model of ZFC and A is definable over $\mathcal{M}$ then $\mathcal{M}$ is a model of A-ZFC.

Theorem 9.33 (Kunen)**.** *There is no non-trivial elementary embedding $j\colon V \to V$ so that V is a model of j-*ZFC*. In particular, there are no definable non-trivial elementary embeddings $j\colon V \to V$.*

In the statement above, j is not necessarily a class of V. Rather it only exists on the level of the meta-theory. The Kunen-inconsistency can be formalized more naturally in second order Set Theory. Before we prove the theorem, some preliminary observations.

Proposition 9.34. *Suppose $j\colon V \to M$ is an elementary embedding with M transitive. Then $j[\mathrm{Ord}]$ is ω-closed, i.e. if $\langle \alpha_n \mid n < \omega\rangle$ is an increasing sequence of ordinals in $\mathrm{ran}(j)$ then $\sup_{n<\omega} \alpha_n \in j[\mathrm{Ord}]$.*

Proof. Find β_n so that $j(\beta_n) = \alpha_n$ and let $\alpha_\omega = \sup_{n<\omega} \alpha_n$. Then $j(\langle \beta_n \mid n < \omega\rangle) = \langle \alpha_n \mid n < \omega\rangle$ and hence $j(\sup_{n<\omega} \beta_n) = \sup_{n<\omega} \alpha_n = \alpha_\omega$. $\square$

Proof of Theorem 9.33. (This proof is due to Woodin) Suppose otherwise. Let $\kappa_0 = \mathrm{crit}(j)$ and if κ_n is defined, let $\kappa_{n+1} = j(\kappa_n)$. The sequence $\vec{\kappa} := \langle \kappa_n \rangle_{n<\omega}$ is the **critical sequence of** j and since V is a model of j-ZFC, the recursion can be carried out within V so that $\vec{\kappa} \in V$. Let $\lambda = \sup_{n<\omega} \kappa_n$.

Claim 9.35. *$j(\lambda) = \lambda$ and $j(\lambda^+) = \lambda^+$.*

Proof. $j(\langle \kappa_n \mid n < \omega\rangle) = \langle j(\kappa_n) \mid n < \omega\rangle = \langle \kappa_{n+1} \mid n < \omega\rangle$ and hence

$$j(\lambda) = j(\sup_{n<\omega} \kappa_n) = \sup_{n<\omega} \kappa_{n+1} = \lambda.$$

Further, $j(\lambda^+) = j(\lambda)^+ = \lambda^+$. □

Let $\vec{S} := \langle S_i \mid i < \kappa_0\rangle$ be a splitting of the stationary set $E^{\lambda^+}_\omega = \{\alpha < \lambda^+ \mid \mathrm{cof}(\alpha) = \omega\rangle$ into κ_0-many sets which is possible by Theorem 5.14. By elementarity of j, $j(\vec{S})$ is a sequence of length $j(\kappa_0) = \kappa_1$ of pairwise disjoint stationary subsets of $j(E^{\lambda^+}_\omega) = E^{j(\lambda^+)}_{j(\omega)} = E^{\lambda^+}_\omega$, so we may write

$$j(\vec{S}) = \langle \tilde{S}_i \mid i < \kappa_1\rangle.$$

We will now take a closer look at $\tilde{S}_{\kappa_0}$. Let $C = \mathrm{Lim}(j[\lambda^+])$ which is a club in λ^+.

Claim 9.36. *$C \cap \tilde{S}_{\kappa_0} \subseteq j[\lambda^+]$.*

Proof. If $\alpha \in \tilde{S}_{\kappa_0} \cap C$ then α is the supremum of a strictly increasing sequence of elements of $j[\lambda^+]$ of length ω, hence is in $j[\lambda^+]$ by Proposition 9.34. □

As $\tilde{S}_{\kappa_0}$ is stationary, we can find some $\beta \in C \cap \tilde{S}_{\kappa_0}$. By the above claim, $\beta = j(\alpha)$ for some $\alpha < \lambda^+$. Note that $j(\mathrm{cof}(\alpha)) = \mathrm{cof}(j(\alpha)) = \mathrm{cof}(\beta) = \omega$ so that $\alpha \in E^{\lambda^+}_\omega$ which implies that there is a unique $i < \kappa_0$ so that $\alpha \in S_i$. But then $j(S_i) = \tilde{S}_{j(i)} = \tilde{S}_i$ and consequently $\alpha \in \tilde{S}_i \cap \tilde{S}_{\kappa_0}$. But the sequence $j(\vec{S})$ consists of pairwise disjoint sets, contradiction. □

Remark 9.37. The argument above really only uses that V is a model of j-ZFC to argue that λ exists (as opposed to $(\kappa_n)_{n<\omega}$ being cofinal in Ord) and to ensure $j \restriction \lambda^+ \in V$ (in order to construct C). The latter of which follows from the weaker assumption that j is amenable[16] to V.

The above proof makes use of the axiom of choice to ensure that λ^+ is regular and in the form of Solovay's Splitting Theorem 5.14. It is not known whether the existence of a non-trivial elementary $j\colon V \to V$ so that $V \models j$-ZF is inconsistent. The critical point of such an embedding is called a **Reinhardt cardinal**.

With a bit of coding, the proof of Kunen's inconsistency actually shows (in ZFC):

Theorem 9.38. *There is no λ and non-trivial elementary embedding $j\colon V_{\lambda+2} \to V_{\lambda+2}$.*

Proof. Exercise. □

[16]Some $A \subseteq M$ is amenable to M iff $x \cap A \in M$ for all $x \in M$.

The reason for the +2 is that in the proof of Theorem 9.33, we talk about λ^+ and stationary subsets of λ^+, which $V_{\lambda+2}$ can talk about (albeit not directly), but $V_{\lambda+1}$ cannot.

Definition 9.39. An I_1-cardinal[17] is the critical point of a non-trivial elementary embedding $j\colon V_{\lambda+1} \to V_{\lambda+1}$.

It is not known whether the existence of I_1-cardinals is consistent. If it is, the theory ZFC+ "there is an I_1-cardinal" is of very high consistency strength. Many prominent Set Theorists believe in the consistency of I_1-cardinals, but may nonetheless be uncomfortable if you ask them directly about it. The "choiceless large cardinals" are those large cardinals which are inconsistent with the axiom of choice, e.g. Reinhardt cardinals. Farmer Schlutzenberg has shown that there is a smooth transition between the large cardinals in ZFC and the choiceless large cardinals.

Theorem 9.40 (Schlutzenberg)**.** *If the theory* ZFC + "*there is a non-trivial elementary embedding* $j\colon L(V_{\lambda+1}) \to L(V_{\lambda+1})$ *with* $\mathrm{crit}(j) < \lambda$" *is consistent then so is* ZF + "*there is a non-trivial elementary* $j\colon V_{\lambda+2} \to V_{\lambda+2}$"*.*

This suggests that the use of the Axiom of Choice is in the proof of Theorem 9.33 is indeed not an accident!

9.4 Iterated Ultrapowers

Suppose $\mathcal{U}$ is a measure on a measurable cardinal κ. Then we can take the ultrapower $j_{\mathcal{U}}\colon V \to \mathrm{Ult}(V,\mathcal{U})$. But now $\mathrm{Ult}(V,\mathcal{U}) \models$ "$j_{\mathcal{U}}$ is a measure on $j_{\mathcal{U}}(\kappa)$", so we may continue and take another ultrapower

$$j_{j_{\mathcal{U}}^{\mathrm{Ult}(V,\mathcal{U})}(\mathcal{U})}\colon \mathrm{Ult}(V,\mathcal{U}) \to \mathrm{Ult}(\mathrm{Ult}(V,\mathcal{U}), j_{\mathcal{U}}(\mathcal{U})).$$

Continuing in this fashion, we can construct models M_n, elementary embedding $j_{n,n+1}\colon M_n \to M_{n+1}$, for $n < \omega$ so that $M_{n+1} = \mathrm{Ult}(M_n,\mathcal{U}_n)$, $M_0 = V$ and $\mathcal{U}_n = j_{n-1,n}(\mathcal{U}_{n-1})$ and $\mathcal{U}_0 = \mathcal{U}$. We can also continue at limit steps!

The category of first order structures with elementary embeddings as morphisms has co-limits. If $\preceq$ is a directed partial order on I and $\langle \mathcal{M}_i, \pi_{i,j} \mid i \preceq j\rangle$ is a commutative system of elementary embeddings (i.e. $\pi_{i,k} = \pi_{j,k} \circ \pi_{i,j}$ whenever $i \preceq j \preceq k$) then there is a co-limit $\langle \mathcal{M}_\infty, \pi_{i,\infty} \mid i \in J\rangle$ for this system (which we will call **direct limit**). The direct limit is unique (up to isomorphism) with the property that whenever $\mathcal{N}$ is a first order structure and $\mu\colon \mathcal{M}_i \to \mathcal{N}$ are elementary embeddings so that $\mu_i = \mu_j \circ \pi_{i,j}$ for all $i \preceq j$, then there is a unique elementary embedding $\mu_\infty\colon \mathcal{M}_\infty \to \mathcal{N}$ so that $\mu_i = \mu_\infty \circ \pi_{i,\infty}$ for all $i \in I$. This "universal property" is somewhat less relevant to us. In this category, the direct limit is alternatively uniquely characterized by the properties

(*i*) $\pi_{i,\infty}\colon \mathcal{M}_i \to \mathcal{M}_\infty$ is an elementary embedding,

[17] As the story goes, the I is supposedly short for "inconsistent

(*ii*) $\pi_{i,\infty} = \pi_{j,\infty} \circ \pi_{i,j}$ whenever $i \preceq j$ and

(*iii*) $\bigcup_{i \in I} \operatorname{ran}(\pi_{i,\infty})$ is (the universe of) $\mathcal{M}_\infty$.

Let us give an explicit construction for the interested reader.

Lemma 9.41. *The direct limit $\langle \mathcal{M}_\infty, \pi_{i,\infty} \mid i \in I\rangle$ exists.*

Proof. Define the universe of $\mathcal{M}_\infty$ as $\bigcup_{i\in I} M_i \times \{i\}/\sim$ where $(x,i) \sim (y,j)$ iff $\exists k\, i,j \preceq k \wedge \pi_{i,k}(x) = \pi_{j,k}(y)$. The map $\pi_{i,\infty}$ is then given by $\pi_{i,\infty}(x) = [(x,i)]_\sim$. The functions and relations in the underlying language are interpreted in $\mathcal{M}_\infty$ in the unique way which turns the $k_{i,\infty}$ into a homomorphism: For example for a n-ary relation R, we let $M_\infty \models R([x_0,i_0]_\sim, \dots, [x_{n-1}, i_{n-1}]_\sim)$ iff for any j so that $i_0, \dots, i_{n-1} \preceq j$ we have

$$\mathcal{M}_j \models R(\pi_{i_0,j}(x_0), \dots, \pi_{i_{n-1},j}(x_{n-1})).$$

It is not difficult to check that this does not depend on the choice of representatives.

It is clear that $\pi_{i,\infty} = \pi_{j,\infty} \circ \pi_{i,j}$ for $i \preceq j$ by definition of $\sim$, so it remains to check that the $\pi_{i,\infty}$ are indeed elementary embeddings and we do so by induction along the complexity of formulas. We only check the difficult direction in the quantifier case. So assume that $\mathcal{M}_\infty \models \exists x \varphi(x, \pi_i, \infty(y))$, we want to prove $\mathcal{M}_i \models \exists x\varphi(x,y)$. Say $[z,j]_\sim$ is a witness, i.e. $\mathcal{M}_\infty \models \varphi([z,j]_\sim, \pi_{i,\infty}(y))$. Then there is k with $i,j \preceq k$ and since $\pi_{i,\infty} = \pi_{k,\infty} \circ \pi_{i,k}$, we find that

$$\mathcal{M}_k \models \exists x \varphi(x, \pi_{i,k}(y))$$

as witnessed by $x = \pi_{j,k}(z)$ by our inductive assumption. As $\pi_{i,k}$ is elementary, it follows that $\mathcal{M}_i \models \exists x \varphi(x,y)$

$\square$

Definition 9.42. Suppose M is a transitive model of ZFC$^-$+$\mathcal{U}$ is an ultrafilter". The **iteration of M by $\mathcal{U}$ of length ξ** is denoted by

$$\langle \mathrm{Ult}^{(\alpha)}(V,\mathcal{U}), j_{\alpha,\beta} \mid \alpha \leq \beta < \xi\rangle$$

and defined recursively as follows:

(*i*) $\mathrm{Ult}^{(0)}(M,\mathcal{U}) = M$,

(*ii*) $j_{\alpha,\alpha}$ is the identity on $\mathrm{Ult}^{(\alpha)}(M,\mathcal{U})$,

(*iii*) $\mathrm{Ult}^{(\alpha+1)}(M,\mathcal{U}) = \mathrm{Ult}(\mathrm{Ult}^{(\alpha)}(M,\mathcal{U}), j_{0,\alpha}(\mathcal{U}))$ and $j_{\beta,\alpha+1} = j_{j_{0,\alpha}(\mathcal{U})}^{\mathrm{Ult}^{(\alpha)}(M,\mathcal{U})} \circ j_{\beta,\alpha}$ for $\beta \leq \alpha$ and

(*iv*) if $\alpha \in \mathrm{Lim} \cap \xi$ then $\langle \mathrm{Ult}^{(\alpha)}(M,\mathcal{U}), j_{\beta,\alpha} \mid \beta < \alpha\rangle$ is the direct limit along

$$\langle \mathrm{Ult}^{(\beta)}(M,\mathcal{U}), j_{\beta,\gamma} \mid \beta < \gamma < \alpha\rangle.$$

We extend our convention and replace $\mathrm{Ult}^{(\alpha)}(V,\mathcal{U})$ by its transitive collapse if it is wellfounded. We will prove next that this is always the case.

Technically, we have only defined $\mathrm{Ult}(M,\mathcal{U})$ under the assumption that M is transitive, so officially we stop the iteration above if we ever reach an illfounded model. It should nonetheless be clear how to define ultrapowers of non-transitive models as well. If we mention $\mathrm{Ult}^{(\alpha)}(M,\mathcal{U})$, it implicitly means that it exists as well.

Theorem 9.43. *Suppose $\mathcal{U}$ is a measure. Then $\mathrm{Ult}^{(\alpha)}(V,\mathcal{U})$ is wellfounded for all ordinals α.*

We will cover some basics first. We will see that there is no reason to worry about an ambiguous definition here. We could get by without it, but we will first prove **absoluteness of wellfoundedness**.

Lemma 9.44. *Suppose M is a transitive model of $\mathrm{ZFC}-$(Power) and $R \in M$ is a binary relation on X. Then R is wellfounded iff $M \models$ "R is wellfounded.*

Proof. If R is illfoundedness is a Σ_1-property in the parameter R, so if R is illfounded in M then R is truly illfounded. On the other hand, suppose that R is wellfounded in M.

Claim 9.45. *There is a function $f\colon X \to \mathrm{Ord}$, $f \in M$ so that xRy implies $f(x) < f(y)$.*

Proof. Applying the recursion theorem in M (which does not use (Power)), there is a function $f\colon X \to \mathrm{Ord}$ so that $f(x) = \sup\{f(y)+1 \mid yRx\}$. □

If R was illfounded in V, we can find a sequence $(x_n)_{n<\omega}$ so that $x_{n+1}Rx_n$ for all $n<\omega$. But then $(f(x_n))_{n<\omega}$ is an infinite descending sequence of ordinals, contradiction. □

Proposition 9.46. *Suppose $M \subseteq W$ are transitive model of ZFC^- and ZFC respectively, and M is definable in W. Suppose that $M \models$ "$\mathcal{U}$ is an ultrafilter". Then the iteration of M by $\mathcal{U}$ is absolute between V and W, that is*

$$\langle \mathrm{Ult}^{(\alpha)}(M,\mathcal{V}), j_{\alpha,\beta} \mid \alpha \leq \beta < \gamma\rangle = (\langle \mathrm{Ult}^{(\alpha)}(M,\mathcal{V}), j_{\alpha,\beta} \mid \alpha \leq \beta < \gamma\rangle)^W$$

for every $\gamma \in \mathrm{Ord} \cap W$.

Proof. All the relevant notions are absolute between V and W: $\sim_{\mathcal{U}}$, $E_{\mathcal{U}}$ as well as the direct limit construction are Δ_0 and wellfoundedness is absolute by Lemma 9.44. Further, any Mostowski collapse is unique so agrees between V and W. □

Lemma 9.47. *Suppose $\alpha + \beta = \gamma$ (in ordinal arithmetic) and $\mathcal{U}$ is a measure. Then*

$$\mathrm{Ult}^{(\gamma)}(V,\mathcal{U}) = \mathrm{Ult}^{(\beta)}(\mathrm{Ult}^{(\alpha)}(V,\mathcal{U}), j_{0,\alpha}(\mathcal{U}))$$

and $j_{\alpha,\gamma} = k_{0,\beta}$ where $\langle \mathrm{Ult}^{(\xi)}(\mathrm{Ult}^{(\alpha)}(V,\mathcal{U}), j_{0,\alpha}(\mathcal{U})), k_{\xi,\zeta} \mid \xi \leq \zeta \leq \beta\rangle$ is the iteration of $\mathrm{Ult}^{(\alpha)}(V,\mathcal{U})$ by $j_{0,\alpha}(\mathcal{U})$ of length $\beta+1$.

Proof. This is straightforward to show by induction on β. For the limit step, use the elementary fact about direct limits that if $\vec{\mathcal{M}} := \langle \mathcal{M}_i, \pi_{i,j} \mid i \preceq j \rangle$ is a directed system along a directed partial order I and $J \subseteq I$ is cofinal (i.e. for $i \in I$ there is $j \in J$ with $i \preceq j$) then $\vec{\mathcal{M}}$ and the subsystem $\langle \mathcal{M}_i, \pi_{i,j} \mid i, j \in J \wedge i \preceq j \rangle$ admit "the same" direct limit. That is, the models $\mathcal{M}_\infty$ are the same and the embeddings $\pi_{i,\infty}$ are the same for $i \in J$ (up to isomorphism). This easily follows from the defining property of the direct limit. □

Proof of Theorem 9.43. Let us write M_β instead of $\mathrm{Ult}^{(\beta)}(V, \mathcal{U})$.

Suppose γ is least so that $\mathrm{Ult}^{(\gamma)}(V, \mathcal{U})$ is illfounded. Putting together Lemma 9.10, Lemma 9.47 and 9.44, we see that γ is a limit ordinal.

Claim 9.48. *The embedding $j_{0,\gamma} \restriction \mathrm{Ord}\colon \mathrm{Ord} \to \mathrm{Ord}^{M_\gamma}$ is cofinal, i.e. for any $a \in \mathrm{Ord}^{M_\gamma}$ there is some β with $M_\gamma \models a \leq j_{0,\gamma}(\beta)$.*

Proof. By induction on $\beta \leq \gamma$, one sees that $j_{0,\beta} \restriction \mathrm{Ord}\colon \mathrm{Ord} \to \mathrm{Ord}^{M_\beta}$ is cofinal. For $\beta < \gamma$, $j_{\beta,\beta+1} \restriction \mathrm{Ord}\colon \mathrm{Ord} \to \mathrm{Ord}$ is order-preserving hence cofinal, so that $j_{0,\beta+1} \restriction \mathrm{Ord}\colon \mathrm{Ord} \to \mathrm{Ord}$ is cofinal. If $\beta \in \mathrm{Lim}$ then any $\alpha \in \mathrm{Ord}^{M_\beta}$ is of the form $j_{\delta,\beta}(\alpha')$ for some $\delta < \beta$. Hence there is some ξ with $j_{0,\delta}(\xi) \geq \alpha'$ and consequently

$$M_\beta \models j_{0,\beta}(\xi) = j_{\delta,\beta}(j_{0,\delta}(\xi)) \geq j_{\delta,\beta}(\alpha') = \alpha.$$

□

Note that Ord^{M_γ} is illfounded: if $(x_n)_{n<\omega}$ is E-descending then the "ranks" $(\mathrm{rk}(x_n))^{\mathcal{M}_\gamma}$ of the x_n as computed in M_γ are descending M_γ-ordinals. Here, E is the $\in$-relation as interpreted by M_γ.

By the claim above, there is some least ordinal ξ so that $j_{0,\gamma}(\xi)$ is in the illfounded part, i.e. so that there is a descending sequence $(a_n)_{n<\omega}$ of M_γ-ordinals with $a_0 = j_{0,\gamma}(\xi)$. We can now find some $\alpha < \gamma$ and ξ' so that $a_1 = j_{\alpha,\gamma}(\xi')$. Let β so that $\alpha + \beta = \gamma$. By Lemma 9.47 and Proposition 9.46,

$$M_\alpha \models \text{“}\beta \text{ is least so that } \mathrm{Ult}^{(\beta)}(M_\alpha, j_{0,\alpha}(\mathcal{U})) \text{ is illfounded”}$$

and further, we have

$$\begin{gathered} M_\alpha \models \text{“}j_{0,\alpha}(\xi) \text{ is the least ordinal } \zeta \text{ so that } j_{\alpha,\gamma}(\zeta) \text{ is} \\ \text{in the illfounded part of } \mathrm{Ult}^{(\beta)}(M_\alpha, j_{0,\alpha}(\mathcal{U}))\text{”} \end{gathered}$$

by elementarity of $j_{0,\alpha}$. But $j_{\alpha,\gamma}(\xi')$ is in the illfounded part of M_γ and $j_{\beta,\gamma}(\xi') = b_1 <^{M_\gamma} b_0 = j_{0,\gamma}(\xi)$ so that $\xi' < j_{0,\gamma}(\xi)$. Contradiction. □