Set Theory

Jacques Duparc

Bâtiment Anthropôle
CH - 1015 Lausanne

Jacques.Duparc@unil.ch
Jacques.Duparc@epfl.ch

EPFL

Contents

Introduction

These lecture notes are very much inspired from Kunen's *Set Theory an Introduction to Independence Proofs* [22].

The basic requirements for this course are contained in the "Mathematical Logic" course. Among other things, you should have a clear understanding of each of the following: first order language, signature, terms, formulas, theory, proof theory, models, completeness theorem, compactness theorem, Löwenheim-Skolem theorem. Ici la belle citation [26]
It **makes no sense** to take this course **without a solid background on first order logic** — see [2, 3, 4, 5, 6, 33].
We use the following notations

Notation 1 (Formulas). *Given $\mathcal{L}$ any first order language, the* set of all *$\mathcal{L}$-formulas is the $\subseteq$-least $X \subseteq \mathcal{L}^{\omega}$ that satisfies:*

- *all atomic formulas belong to X.*
- *If φ and ψ belong to X, then*

$$\neg\varphi,\ (\varphi \wedge \psi),\ (\varphi \vee \psi),\ (\varphi \rightarrow \psi) \text{ and } (\varphi \leftrightarrow \psi)$$

  also belong to X.
- *If x is any variable and φ belongs to X, then $\forall x\ \varphi$ and $\exists x\, \varphi$ also belong to X.*

For better readability, we usually omit the outermost parentheses. For instance we write

$$\exists x\, x \in y \longrightarrow x \neq y$$

instead of

$$(\exists x\, x \in y \longrightarrow x \neq y).$$

Notation 2. *Given any first order language $\mathcal{L}$, and any set of $\mathcal{L}$-formulas Γ, and any $\mathcal{L}$-formula φ, we write*

- $\Gamma \models \varphi$ *for "φ holds in all $\mathcal{L}$-structures that satisfy Γ.*
- $\Gamma \vdash \varphi$ *for "Γ proves φ (in classical logic)".*

Part I

The Theory Itself

Chapter 1

The Axioms

Zermelo-Fraenkel Set Theory is a first order theory with equality whose language only contains a single binary relation symbol denoted by "$\in$" and called epsilon. The objects over which the variables range are called "sets". When the relation $x \in y$ holds, we say that "x is an element of y" or that "x belongs to y".

Whether we consider Zermelo Set Theory (denoted **Z**), Zermelo-Fraenkel Set Theory (denoted **ZF**), or Zermelo-Fraenkel Set Theory plus the Axiom of Choice (denoted **ZFC**), all three of them consist in infinitely many axioms.

- **Z** corresponds to axioms (0)– (7)
- **ZF** corresponds to axioms (0)– (8)
- **ZFC** corresponds to axioms (0)– (9)

ZFC

Axioms of Zermelo-Fraenkel + Choice

0. **Set Existence**

$$\exists x \;\; x = x$$

(1) **Extensionality**

$$\forall x \; \forall y \; \big(\forall z \, (z \in x \longleftrightarrow z \in y) \rightarrow x = y)\big)$$

(2) **Comprehension Schema**

$$\forall z \forall w_1 ... \forall w_n \exists y \forall x \big(x \in y \longleftrightarrow (x \in z \ \wedge \ \varphi)\big)$$

where $\varphi := \varphi_{(x,z,w_1 ... w_n)}$ is any formula whose free variables — if any — are among $x, z, w_1, \ldots, w_n$. Moreover, y is not free in φ.

(3) **Pairing**

$$\forall x \forall y \exists z (x \in z \ \wedge \ y \in z)$$

(4) **Union**

$$\forall a \exists b \forall x \forall y \big((x \in y \ \wedge \ y \in a) \rightarrow x \in b\big)$$

(5) **Infinity**

$$\exists x (\exists y \, (y \in x \ \wedge \ \forall z \, z \notin y) \ \wedge \ \forall z (z \in x \rightarrow z \cup \{z\} \in x))$$

(6) **Power Set**

$$\forall x \, \exists y \, \forall z \, \big(\forall u \ \ (u \in z \rightarrow u \in x) \rightarrow z \in y\big)$$

(7) **Foundation**

$$\forall x \, \Big(\exists y \, y \in x \rightarrow \exists y \, \big(y \in x \ \wedge \ \neg \exists z \, \big(z \in x \ \wedge \ z \in y\big)\big)\Big)$$

(8) **Replacement Schema**

$$\forall A \ \forall w_1 \ldots \forall w_n \ \Big[\forall x \, \big(x \in A \longrightarrow \exists ! y \ \varphi\big) \longrightarrow \exists B \ \forall x \, \big(x \in A \longrightarrow \exists y \, (y \in B \ \wedge \ \varphi)\big)\Big]$$

where $\varphi := \varphi_{(x,y,A,w_1,\ldots,w_n)}$ is any formula with free variables among $x, y, A, w_1, \ldots, w_n$, and where $\exists ! y \, \varphi$ abbreviates

$$\exists y \, \Big(\varphi_{(x,y,A,w_1,\ldots,w_n)} \ \wedge \ \forall z \, \big(\varphi(x,z,A,w_1,\ldots,w_n) \longrightarrow z = y)\big)\Big).$$

(9) **Choice**

$$\forall x \left(\left(\begin{array}{c} \forall y \, (y \in x \longrightarrow \exists z \, z \in y) \\ \bigwedge \\ \forall y \forall y' \left(\left(\begin{array}{c} y \in x \\ \wedge \\ y' \in x \\ \wedge \\ \exists z \, (z \in y \ \wedge \ z \in y') \end{array} \right) \longrightarrow y = y' \right) \end{array} \right) \longrightarrow \exists c \forall y \, \Big(y \in x \longrightarrow \exists ! z \, (z \in y \ \wedge \ z \in c) \Big) \right).$$

1.1 Extensionality and Comprehension

Set Existence is not really needed since first order logic proves it without any hypothesis. i.e., formally, one has

$$\vdash \exists x\, x = x.$$

So, it is redundant and should as well be omitted.

Extensionality says that what makes a set is its elements. This means that two sets that have the same elements are the same sets. In other words, if two sets are different, there must exist one element that belongs to one of them but not the other one.

Comprehension Schema is not a single axiom, but one axiom for each $\mathcal{L}$-formula φ. So these are countably many axioms. The idea is that given a "property P", to form the "set of all sets that satisfy the property P". Since the theory we work in is expressed in first order logic, such a property should be conveyed by a first order formula — e.g. φ — with a single free variable, say x, and the "set of all sets that satisfy the property P" becomes now

$$\{x \mid \varphi\}$$

which is a notation for $\exists y\, \forall x\ (x \in y \longleftrightarrow \varphi)$. But this does not work for it leads to Russell's paradox[1]:

take $\varphi := x \notin x$ — which is a notation for $\neg x \in x$ — and form $y = \{x \mid \varphi\}$; i.e.,

$$y = \{x \mid x \notin x\}\,.$$

One comes to the contradiction

$$y \in y \longleftrightarrow y \notin y.$$

For this reason, instead of claiming the existence of $\{x \mid \varphi\}$, an instance of the **Comprehension Schema** requires a preexisting set z in order to form

$$\{x \in z \mid \varphi\}\,.$$

In other words, it is the *universal closure* of

$$\exists y \forall x \big(x \in y \longleftrightarrow (x \in z\ \wedge\ \varphi)\big).$$

Notice that it is also required that y be not free in φ in order to avoid paradoxes that would arise for instance from

$$\exists y \forall x \big(x \in y \longleftrightarrow (x \in z\ \wedge\ x \notin y)\big)$$

as soon as z would be a non-empty set.

Since there exists a set, we are entitled to form a set y that satisfies $\forall x\ \ x \notin y$.

[1]The same paradox was discovered a year before by Ernst Zermelo who did not publish it but rather make it known to members of the University of Göttingen such as David Hilbert.

Definition 3. *The empty set is the* unique *set y defined from any set z by*

$$\varnothing := \{x \in z \mid x \neq x\}.$$

Notice that the empty set is unique by **Extensionality**.

On the other hand, there is no *Universal Set*. Or in other words, the collection of all sets is not a set.

Theorem 4.

$$\mathbf{Z} \vdash_c \neg\exists z\, \forall x \;\; x \in z.$$

Proof of Theorem 4: Towards a contradiction, we assume $\exists z\, \forall x \;\; x \in z$ and form, by an instance of the **Comprehension Schema**,

$$\{x \in z \mid x \notin x\}$$

which is no different than

$$\{x \mid x \notin x\}$$

which leads to Russell's paradox.

$\square$ 4

Notice that **Set Existence**, **Extensionality** and **Comprehension Schema** together prove that the empty set exist. But they cannot prove that another set exists.
Indeed, consider the following $\mathcal{L}$-structure $\mathcal{M} = \left\langle \{\varnothing\}, \in^{\mathcal{M}} \right\rangle$ where $\in^{\mathcal{M}} = \varnothing$, i.e.,

$$\mathcal{M} \models \varnothing \notin \varnothing.$$

One has

- $\mathcal{M} \models$ **Set Existence**
- $\mathcal{M} \models$ **Extensionality**
- $\mathcal{M} \models$ **Comprehension Schema**.

Hence,

$$\mathcal{M} \models \{\textbf{Set Existence}, \textbf{Extensionality}, \textbf{Comprehension Schema}\}$$

and

$$\mathcal{M} \models \forall x \;\; x = \varnothing.$$

Therefore

$$\{\textbf{Set Existence}, \textbf{Extensionality}, \textbf{Comprehension Schema}\} \not\models \exists x \;\; x \neq \varnothing.$$

1.2 Pairing, Union and Replacement

Pairing is an axiom that, given two sets, provides a pair of sets. It reads

$$\forall x \forall y \exists z (x \in z \ \wedge \ y \in z)$$

which only says that given sets x and y, there exists some set z that contains both of them. In fact, this axiom does not say $z = \{x, y\}$ but $\{x, y\} \subseteq z$. In order to obtain exactly the set $\{x, y\}$, it is enough to apply **Comprehension Schema**:

$$\{x, y\} = \{a \in z \mid a = x \ \vee \ a = y\}$$

This set is unique by **Extensionality**.

Example 5. *So, now we can build infinitely many sets, for instance*

- $\{\varnothing\} = \{\varnothing, \varnothing\}$
- $\{\{\varnothing\}\} = \{\{\varnothing\}, \{\varnothing\}\}$
- $\{\{\{\varnothing\}\}\} = \{\{\{\varnothing\}\}, \{\{\varnothing\}\}\}$, *etc.*

Notation 6. *We recall that given any set y and any formula φ,*

- $\exists x \in y \, \varphi$ *abbreviates* $\exists x (x \in y \ \wedge \ \varphi)$
- $\forall x \in y \, \varphi$ *abbreviates* $\forall x (x \in y \longrightarrow \varphi)$.

Notation 7. *Given any sets, x, y,*

- $x \subseteq y$ *abbreviates* $\forall z \in x \ \ z \in y$
- $x \subsetneq y$ *abbreviates* $\forall z \in x \ \ z \in y \ \wedge \ \exists z \in y \ \ z \notin x$.

Pairing allows to define the *ordered pair* which is a starting point to defining sequences of sets.

Definition 8. *Given two sets, x and y, the* ordered pair (x, y) *as the set*

$$\{\{x\}, \{x, y\}\}.$$

What, we need is to recover x and y, from any set of the form $\{\{x\}, \{x, y\}\}$. Indeed, in case $x \neq y$, then we have $\{\{x\}, \{x, y\}\}$ is a set that contains two different elements: a singleton and a pair. The element inside the singleton is x and the one inside the pair that is not x

is y. In case $x = y$, then $\{\{x\},\{x,y\}\} = \{\{x\},\{x,x\}\} = \{\{x\},\{x\}\} = \{\{x\}\}$ is a singleton whose element is also a singleton containing x.

Formally, we show

Lemma 9.

$$(x,y) = (x',y') \iff \begin{cases} x = x' \\ \textit{and} \\ y = y'. \end{cases}$$

Proof of Lemma 9: $(\Leftarrow)$ is obvious. For $(\Rightarrow)$ we distinguish between $x = y$ and $x \neq y$.

- If $x = y$, then $(x,y) = \{\{x\},\{x,y\}\} = \{\{x\},\{x,x\}\} = \{\{x\},\{x\}\} = \{\{x\}\}$. Therefore, by **Extensionality** one has,

$$\{\{x\}\} = \left\{\left\{x'\right\},\left\{x',y'\right\}\right\} \iff \{x\} = \left\{x'\right\} \ \textit{and} \ \left\{x'\right\} = \left\{x',y'\right\}.$$

  which leads to both $x = x'$ and $y' = x' = x$.
- If $x \neq y$, then $(x,y) = \{\{x\},\{x,y\}\}$ and $\{x\} \neq \{x,y\}$. Therefore, since $\{\{x\},\{x,y\}\} = \{\{x'\},\{x',y'\}\}$, one must have $x' \neq y'$ and also $\{x\} = \{x'\}$ and $\{x,y\} = \{x',y'\}$ which leads to to both $x = x'$ and $y' = x'$.

$\square$ 9

Union claims that given any set a, there exists some set b that contains — at least — all elements of elements of a:

$$\forall a \exists b \forall x \forall y \big((x \in y \ \wedge \ y \in a) \rightarrow x \in b\big).$$

By **Comprehension Schema** and **Extensionality**, we define the unique following set

$$\bigcup a = \{x \in b \mid \exists y \in a \ \ x \in y\}$$

Example 10.

- $\bigcup \varnothing = \varnothing$
- $\bigcup \{\varnothing\} = \varnothing$
- $\bigcup \{\{\varnothing\}\} = \{\varnothing\}$
- $\bigcup \{\{\{\varnothing\}\}\} = \{\{\varnothing\}\}$
- $\bigcup \{\{\varnothing,\{\varnothing\}\}\} = \{\varnothing,\{\varnothing\}\}$
- $\bigcup \{\{\varnothing\},\{\varnothing,\{\varnothing\}\}\} = \{\varnothing,\{\varnothing\}\}.$

Notation 11. *Given any non-empty set a,*

$$\bigcap a = \left\{ x \in \bigcup a \mid \forall y \in a \ \ x \in y \right\}.$$

Notation 12. *Given any sets, x, y,*

- $x \cup y$ *abbreviates* $\bigcup\{x, y\}$
- $x \cap y$ *abbreviates* $\bigcap\{x, y\}$
- $x \smallsetminus y$ *abbreviates* $\{z \in x \mid z \notin y\}$.

Replacement Schema talks about a given a formula φ that *"behaves as a functional"* on any given set A. This means that for every element x in A there exists a unique element y (somewhere in the universe!) that φ relates to x. So, φ behaves as if it were a function that maps elements of A to other elements of the universe. **Replacement Schema** then says that there exists a set B that contains all elements of the form y related to some x from A. Namely,

$$\forall A \ \forall w_1 \ldots \forall w_n \ \ \Big[\forall x \left(x \in A \longrightarrow \exists! y \ \varphi \right) \longrightarrow \exists B \ \forall x \left(x \in A \longrightarrow \exists y \left(y \in B \ \wedge \ \varphi \right) \right) \Big]$$

where $\varphi := \varphi_{(x,y,A,w_1,\ldots,w_n)}$ is any formula with free variables among $x, y, A, w_1, \ldots, w_n$, and where $\exists! y \, \varphi$ abbreviates

$$\exists y \left(\varphi_{(x,y,A,w_1,\ldots,w_n)} \ \wedge \ \forall z \left(\varphi(x, z, A, w_1, \ldots, w_n) \longrightarrow z = y \right) \right).$$

By **Comprehension Schema** and **Extensionality**, given any $A, w_1, \ldots, w_n$, one obtains a set that behaves as the range of A by this functional. Namely,

$$\left\{ y \in B \mid \exists x \in A \ \varphi_{(x,y,A,w_1,\ldots,w_n)} \right\}.$$

From there, given any sets A and B, one can obtain the cartesian product $A \times B$.

Definition 13. *Given two sets, A and B, the cartesian product $A \times B$ is defined as*

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

This is well defined, since for each $b \in B$ we have

$$\forall a \in A \ \exists! x \ \ x = (a, b)$$

So, by one instance of the **Replacement Schema** and one instance of the **Comprehension Schema**, we obtain

$$A \times \{b\} = \{x \mid \exists a \in A \;\; x = (a,b)\}\,.$$

Again,

$$\forall b \in B \; \exists! x \;\; X = A \times \{b\}$$

So, by one more instance of the **Replacement Schema** and another one instance of the **Comprehension Schema**, we obtain

$$\{A \times \{b\} \mid b \in B\}$$

By **Union**, we obtain

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

1.3 Relations and Functions

Definition 14. *Given any set R, we define*

- $dom\,(R) = \{x \mid \exists y \;\; (x,y) \in R\} \quad \left(= \left\{ x \in \bigcup\bigcup R \mid \exists y \;\; (x,y) \in R \right\} \right)$
- $ran\,(R) = \{y \mid \exists x \;\; (x,y) \in R\} \quad \left(= \left\{ y \in \bigcup\bigcup R \mid \exists x \;\; (x,y) \in R \right\} \right).$
- *A set that satisfies $R \subseteq dom\,(R) \times ran\,(R)$ is called a (binary) relation.*

Notation 15. *Given any relation R, we write*

- R^{-1} *for*

$$R^{-1} = \{(y,x) \in ran\,(R) \times dom\,(R) \mid (x,y) \in R\}\,.$$

- xRy *for* $(x,y) \in R$.

Definition 16. *A set f is a function iff f is a relation that satisfies*

$$\forall x \in dom\,(f) \;\; \exists! y \in ran\,(f) \quad (x,y) \in f.$$

Notation 17. *Given any sets f, A, B we write $f : A \longrightarrow B$ to say that f is a function with $dom(f) = A$ and $ran(f) \subseteq B$*

- *for $x \in A$, $f(x)$ stands for the unique $y \in B$ such that $(x, y) \in f$;*
- *for $C \subseteq A$,*
  - *$f \restriction C$ stands for $f \cap (C \times B)$*
  - *$f[C]$ stands for $\{y \in B \mid \exists x \in C \quad f(x) = y\}$ (i.e., $f[C] = \{f(x) \mid x \in C\}$).*

Since we deal with sets, a subset of a given set may also be an element of the same set. Therefore, as shown in the next example, we need to distinguish between $f[C]$ and $f(C)$.

Example 18. *Let $f : \{\varnothing, \{\varnothing\}\} \longrightarrow \{\varnothing, \{\varnothing\}\}$ be such that $f(\varnothing) = \varnothing$ and $f(\{\varnothing\}) = \varnothing$.* $\left(\textit{i.e.,}\ f = \{(\varnothing, \varnothing), (\{\varnothing\}, \varnothing)\} = \left\{\{\{\varnothing\}, \{\varnothing, \varnothing\}\}, \left\{\{\{\varnothing\}\}, \{\{\varnothing\}, \varnothing\}\right\}\right\} = \left\{\{\{\varnothing\}\}, \left\{\{\{\varnothing\}\}, \{\{\varnothing\}, \varnothing\}\right\}\right\}\right)$

We have

- $f(\{\varnothing\}) = \varnothing$
- $f[\{\varnothing\}] = \{\varnothing\}$.

Chapter 2

Well-orderings and Ordinals

2.1 Well-orderings

Definition 19. *A function* $f : A \longrightarrow B$ *is*

- ***injective*** *if* $f^{-1} \subseteq ran(f) \times A$ *is a function*[1]
- ***surjective*** *if* $ran(f) = B$
- ***bijective*** *if* f *is both injective and surjective.*

Definition 20 (ordering)**.**

- *A* ***(partial) ordering*** *is a pair* (A, R) *such that* R *is a relation that satisfies*

  transitivity *:* $\forall x \in A \; \forall y \in A \; \forall z \in A \quad ((xRy \wedge yRz) \longrightarrow xRz)$
  irreflexivity *:* $\forall x \in A \quad \neg xRx.$

- (A, R) *is a* ***total ordering*** *of* A *if in addition to being an ordering, one has*

$$\forall x \in A \; \forall y \in A \quad \big(x \neq y \longrightarrow (xRy \vee yRx)\big).$$

[1]This is equivalent to say that $\forall x \in A \; \forall x' \in A \quad (f(x) = f(x') \longrightarrow x = x')$. Injective functions are also called one-to-one (1-1).

Definition 21 (order isomorphism)**.** *Given orderings* $(A, <_A)$, $(B, <_B)$, *an order isomorphism is any bijection* $f : A \longrightarrow B$ *that satisfies*

$$\forall x \in A \; \forall y \in A \; \Big(x <_A y \longleftrightarrow f(x) <_B f(y)\Big).$$

Notation 22. *We write* $(A, <_A) \approx (B, <_B)$ *when there exists an order isomorphism* $f : A \longrightarrow B$.

Definition 23 (well-ordering)**.** *We say* $<_A$ *well orders* A *— or equivalently* $(A, <_A)$ *is a well-ordering — if* $(A, <_A)$ *is a total ordering that satisfies*

$$\forall B \subseteq A \; \big(B \neq \varnothing \longrightarrow \exists x \in B \; \forall y \in B \;\; y \nless_A x\big).$$

In other words, a total ordering of A is a well-ordering if any non-empty subset of A admits a $<_A$-minimal element[2].

Example 24.

- $(\mathbb{N}, <)$ *is a well-ordering;*
- $(\mathbb{Q}, <)$ *is not a well-ordering;*
- $(\mathbb{Z}, <)$ *is not a well-ordering;*
- $(\mathbb{R}, <)$ *is not a well-ordering;*
- *every finite total ordering is a well-ordering;*
- *no dense ordering (with at least two elements) is a well-ordering.*

Notation 25. *If* $x \in A$, *and* $(A, <)$ *is an ordering, we write* $\lceil x \rceil_A^<$ *for the set of* $<$*-predecessors of* x *inside* A. *Namely,*

$$\lceil x \rceil_A^< = \{y \in A \mid y < x\}\,.$$

[2] This $<_A$-minimal element is unique since $<_A$ is total on A.

Lemma 26. *If* $(A,<)$ *is a well-ordering and* $x \in A$*, then* $(\lceil \overset{x}{-} \rceil^{<}_{A}, <)$ *is a well-ordering.*

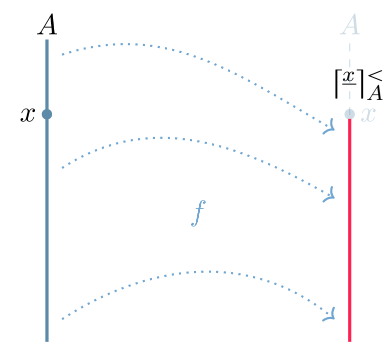
Proof of Lemma 26: Immediate.

$\square$ 26

Lemma 27. *If* $(A,<)$ *is a well-ordering and* $x \in A$*, then*

$$(A,<) \not\approx (\lceil \overset{x}{-} \rceil^{<}_{A}, <).$$

Proof of Lemma 27: Towards a contradiction, we assume $f : A \longrightarrow \lceil \overset{x}{-} \rceil^{<}_{A}$ is an isomorphism.

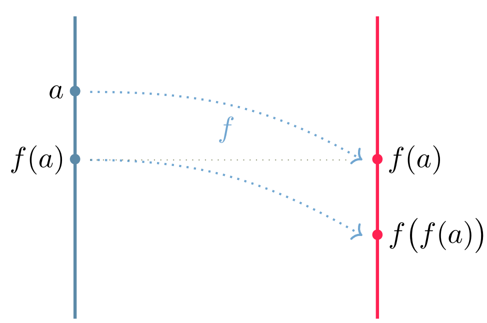


We consider the following set (which is non-empty since it contains x)

$$\{y \in A \mid f(y) \neq y\}$$

together with a its $<$-least element and discuss whether $f(a) < a$ or $a < f(a)$.

- if $f(a) < a$ holds,

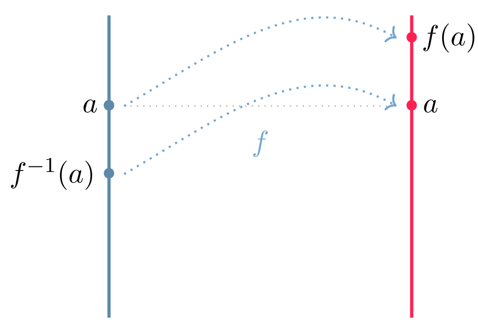


then we have both

- $f\big(f(a)\big) < f(a)$ since f is an isomorphism, and
- $f\big(f(a)\big) = f(a)$ since $f(a) \notin \{y \in A \mid f(y) \neq y\}$ holds by $<$-minimality of a.

A contradiction.

◦ if $a < f(a)$ holds,



then we have

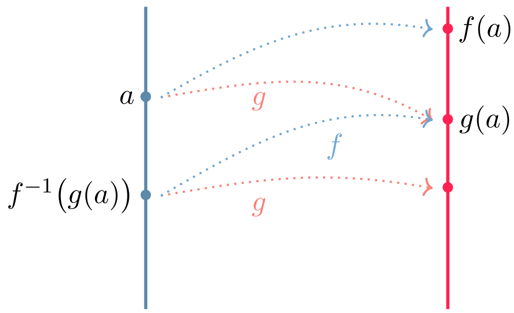
$$\begin{aligned} f^{-1}(a) < a \quad & \longleftrightarrow \quad f\big(f^{-1}(a)\big) < f(a) \\ & \longleftrightarrow \quad a < f(a) \end{aligned}$$

which leads to both $f^{-1}(a) < a$ and $f^{-1}(a) \in \{y \in A \mid f(y) \neq y\}$, contradicting the $<$-minimality of a.

□ 27

Lemma 28. *If $(A, <_A)$ and $(B, <_B)$ are well-orderings and $(A, <_A) \approx (B, <_B)$, then the isomorphism is unique.*

Proof of Lemma 28: Towards a contradiction, we assume that there exist two different isomorphisms $f : A \longrightarrow B$ and $g : A \longrightarrow B$.
Take a the $<_A$-least element inside $\{x \in A \mid f(x) \neq g(x)\}$.
We first assume $g(a) <_B f(a)$:



We consider $f^{-1}\big(g(a)\big)$: since f is an isomorphism we have

$$\begin{aligned} f^{-1}\big(g(a)\big) <_A a \quad &\longleftrightarrow \quad f\Big(f^{-1}\big(g(a)\big)\Big) <_B f(a) \\ &\longleftrightarrow \quad g(a) <_B f(a) \end{aligned}$$

which shows that $f^{-1}\big(g(a)\big) <_A a$ holds; which also, by $<_A$-minimality, means that

$$f^{-1}\big(g(a)\big) \notin \{x \in A \mid f(x) \neq g(x)\}\,.$$

Therefore, we have

$$\begin{aligned} g\Big(f^{-1}\big(g(a)\big)\Big) &= f\Big(f^{-1}\big(g(a)\big)\Big) \\ &= g(a). \end{aligned}$$

Since g is injective, this leads to $f^{-1}\big(g(a)\big) = a$, contradicting $f^{-1}\big(g(a)\big) <_A a$.
The other case: $f(a) <_B g(a)$ leads *mutatis mutandis* to the same contradiction.

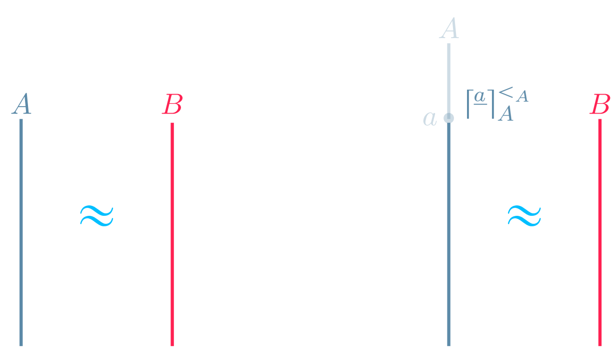
□ 28

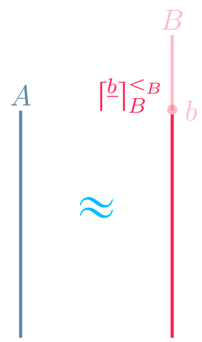
Theorem 29. *If $(A, <_A)$ and $(B, <_B)$ are well-orderings, then only one of the following occurs:*

(1) $(A, <_A) \approx (B, <_B)$

(2) $\exists a \in A \quad (\lceil \underline{a} \rceil_A^{<_A}, <_A) \approx (B, <_B)$

(3) $\exists b \in B \quad (A, <_A) \approx (\lceil \underline{b} \rceil_B^{<_B}, <_B)$





Proof of Theorem 29: The idea of the proof is that an isomorphism must necessarily map the $<_A$-least element of A to the $<_B$-least element of B, then the $<_A$-least element of $A \setminus \{a\}$ to the $<_B$-least element of $B \setminus \{b\}$, etc. If we need to stop the process because of lack of element either on the "A side" or on the "B side", then we fall into cases (2) or (3), otherwise we reach case (1).
Formally, we set

$$f = \left\{(x, y) \in A \times B \mid \left(\lceil \underline{x} \rceil_A^{<_A}, <_A\right) \approx \left(\lceil \underline{y} \rceil_B^{<_B}, <_B\right)\right\}.$$

Towards a contradiction, we suppose that f is not a function. Hence, there exist $(x, y) \in f$ and $(x, y') \in f$ with $y \neq y'$. By symmetry, we assume $y <_B y'$. We have

- $\lceil \underline{x} \rceil_A^{<_A} \approx \lceil \underline{y} \rceil_B^{<_B}$
- $\lceil \underline{x} \rceil_A^{<_A} \approx \lceil \underline{y'} \rceil_B^{<_B}$
- $\lceil \underline{y} \rceil_B^{<_B} = \lceil \underline{y} \rceil_{\lceil \underline{y'} \rceil_B^{<_B}}^{<_B}$

from which we obtain

- $\lceil \underline{y'} \rceil_B^{<_B} \approx \lceil \underline{y} \rceil_B^{<_B}$
- $\lceil \underline{y'} \rceil_B^{<_B} = \lceil \underline{y} \rceil_{\lceil \underline{y'} \rceil_B^{<_B}}^{<_B}$

The latter contradicts Lemma 27, therefore f is a **function**.

Towards a contradiction, we now assume that f is not injective. i.e., there exist $(x, y) \in f$ and $(x', y) \in f$ with $x \neq x'$. By symmetry, we assume $x <_A x'$. We have We have

- $\lceil \underline{x} \rceil_A^{<_A} \approx \lceil \underline{y} \rceil_B^{<_B}$
- $\lceil \underline{x'} \rceil_A^{<_A} \approx \lceil \underline{y} \rceil_B^{<_B}$
- $\lceil \underline{x} \rceil_A^{<_A} = \lceil \underline{x} \rceil_{\lceil \underline{x'} \rceil_A^{<_A}}^{<_A}$

from which we obtain

- $\lceil \underline{x'} \rceil_A^{<_A} \approx \lceil \underline{x} \rceil_A^{<_A}$
- $\lceil \underline{x'} \rceil_A^{<_A} = \lceil \underline{x} \rceil_{\lceil \underline{x'} \rceil_B^{<_A}}^{<_A}$

The latter contradicts Lemma 27, therefore f is **injective**.

We now show that $dom(f)$ is an $<_A$-initial segment of A. i.e.,

$$(x' \in dom(f) \wedge x <_A x') \longrightarrow x \in dom(f).$$

Indeed, consider $(x', y') \in f$ and $x <_A x'$. Then, since $(x', y') \in f$, one has

$$\left(\lceil \underline{x'} \rceil_A^{<_A}, <_A\right) \approx \left(\lceil \underline{y'} \rceil_B^{<_B}, <_B\right)$$

so that there exists some *unique* isomorphism $g : \lceil \underline{x'} \rceil_A^{<_A} \to \lceil \underline{y'} \rceil_B^{<_B}$.
For $y = g(x)$, it appears that $g \restriction \lceil \underline{x} \rceil_A^{<_A} : \lceil \underline{x} \rceil_A^{<_A} \to \lceil \underline{y} \rceil_B^{<_B}$ is an isomorphism, hence $(x, y) \in f$, which completes the proof that $dom(f)$ is an $<_A$-initial segment of A.
We show that $ran(f)$ is an $<_B$-initial segment of B. i.e.,

$$(y' \in ran(f) \ \wedge \ y <_B y') \longrightarrow y \in ran(f).$$

Indeed, we consider $(x', y') \in f$ and $y <_B y'$. Since $(x', y') \in f$, one has

$$\left(\lceil \underline{x'} \rceil_A^{<_A}, <_A\right) \approx \left(\lceil \underline{y'} \rceil_B^{<_B}, <_B\right)$$

so that there exists some *unique* isomorphism $g : \lceil \underline{x'} \rceil_A^{<_A} \to \lceil \underline{y'} \rceil_B^{<_B}$.
For $x = g^{-1}(y)$, we obtain that $g \restriction \lceil \underline{x} \rceil_A^{<_A} : \lceil \underline{x} \rceil_A^{<_A} \to \lceil \underline{y} \rceil_B^{<_B}$ is an isomorphism, hence $(x, y) \in f$, which completes the proof that $ran(f)$ is an $<_B$-initial segment of B.

To prove that f is an isomorphism, it remains to show that

$$\forall x \in dom(f) \ \forall y \in dom(f) \quad \big(x <_A y \longleftrightarrow f(x) <_B f(y)\big)$$

- if $x <_A y$, then there exists a necessarily unique isomorphism $g : \lceil \underline{y} \rceil_A^{<_A} \to \lceil \underline{f(y)} \rceil_B^{<_B}$. Now consider $g(x) \in ran(g) = \lceil \underline{f(y)} \rceil_B^{<_B}$. Therefore, we have $g(x) <_B f(y)$ and since $g \restriction \lceil \underline{x} \rceil_A^{<_A} : \lceil \underline{x} \rceil_A^{<_A} \to \lceil \underline{g(x)} \rceil_B^{<_B}$ is an isomorphism, it follows by unicity of the isomorphism that $g(x) = f(x)$, hence $f(x) <_B f(y)$.

- if $f(x) <_A f(y)$, then take the unique isomorphism $g : \lceil \underline{y} \rceil_A^{<_A} \to \lceil \underline{f(y)} \rceil_B^{<_B}$ and consider $g^{-1}\big(f(x)\big) \in dom(g) = \lceil \underline{y} \rceil_A^{<_A}$. $g \restriction \lceil \underline{g^{-1}\big(f(x)\big)} \rceil_A^{<_A} : \lceil \underline{g^{-1}\big(f(x)\big)} \rceil_A^{<_A} \to \lceil \underline{f(x)} \rceil_B^{<_B}$ is an isomorphism. Hence, by unicity of the isomorphism $f\Big(g^{-1}\big(f(x)\big)\Big) = f(x)$, so that $g^{-1}\big(f(x)\big) = x$ and $x \in dom(g) = \lceil \underline{y} \rceil_A^{<_A}$, which gives $x < y$.

So, we have proved that f is an isomorphism from some initial segment of $(A, <_A)$ to some initial segment of $(B, <_B)$. By construction — the very definition of f — both initial segments cannot be proper, which yields to the three cases of the theorem.

$\square$ 29

2.2 Ordinals

If someone is left with a box of apples and a box of oranges and needs to answer the question whether there are the same number of apples than oranges, then that person has two different ways of coping with that problem:

(1) She can count on one hand how many apples there are, and on the other hand how many oranges there are, then compare these two numbers.

(2) She can also take one apple and one orange and throw them away, and repeat that process until either there is no more apple but still oranges, or no more oranges but still apples, or both boxes become empty at the same time.

The first way of dealing with her issue requires to be able to well-order the apples as well as the oranges. She needs to decide which apple is the first, which other one is the second, etc. The second way of solving this problem relies on the existence of a bijection between two sets (counting up to one is enough!).

- The first notion corresponds to the notion of *ordinal number*,
- the second one to the notion of *cardinal number*.

A canonical presentation of ordinal numbers is through *transitive sets* that are *well-ordered* by the membership relation ($\in$). Where a *transitive set* is nothing bu a set on which the membership relation is transitive.

Definition 30 (Transitive set)**.** *Let A be any set. A is* transitive *if it satisfies*

$$\forall x\, \forall y \ \ \big((x \in y \ \wedge \ y \in A) \longrightarrow x \in A\big).$$

or equivalently

$$\forall x \ \ \big(x \in A \longrightarrow x \subseteq A\big).$$

Example 31.

- *The following sets are transitive:*
  - $\varnothing$
  - $\{\varnothing\}$
  - $\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$
- *The following sets are not transitive:*
  - $\{\{\varnothing\}\}$
  - $\{\varnothing, \{\{\varnothing\}\}\}$
  - $\{\{\varnothing\}, \{\{\varnothing\}\}\}$

Definition 32 (Ordinal)**.** *Let α be any set. α is an* ordinal *if and only if*

(1) α *is transitive, and*

(2) $(\alpha, \in_\alpha)$ *is a well-ordering*[3].

Example 33.

- *The following sets are ordinals:*
  - $\varnothing$
  - $\{\varnothing\}$
  - $\{\varnothing, \{\varnothing\}\}$
- *The following sets are not ordinals:*
  - $\{\{\varnothing\}\}$
  - $\{\{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$
  - $\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$

Theorem 34.

(1) If α *is an ordinal and* $x \in \alpha$*, then* x *is an ordinal and* $x = \lceil \underline{x} \rceil_\alpha^{\in_\alpha}$

(2) If α *and* β *are ordinals and* $(\alpha, \in_\alpha) \approx (\beta, \in_\beta)$*, then* $\alpha = \beta$*.*

(3) If α *and* β *are ordinals, then only one of the following assertions is satisfied:*

(a) $\alpha = \beta$ *(b)* $\alpha \in \beta$ *(c)* $\beta \in \alpha$*.*

(4) If α, β, γ *are ordinals, then*

$$\big(\alpha \in \beta \ \wedge \ \beta \in \gamma\big) \longrightarrow \alpha \in \gamma.$$

(5) If A *is a non-empty set of ordinals, then*

$$\exists \alpha \in A \ \forall \beta \in A \ \ \big(\alpha \in \beta \ \vee \ \alpha = \beta\big).$$

Proof of Theorem 34:

(1) We assume α is an ordinal and $x \in \alpha$.

[3] $\in_\alpha$ stands for $\{(x,y) \in \alpha \times \alpha \mid x \in y\}$.

- $(x, \in_x)$ is a total ordering:

  irreflexivity: given any $y \in x$, since $y \in x \in \alpha$ and α is transitive it follows $y \in \alpha$, and since $(\alpha, \in_\alpha)$ is a well-ordering, this gives $y \notin y$.

  transitivity: Given any $y_1 \in x, y_2 \in x, y_3 \in x$, since $x \in \alpha$ and α is transitive it follows y_1, y_2, y_3 belong to α, and since $(\alpha, \in_\alpha)$ is a well-ordering, this gives

  $$(y_1 \in y_2 \ \wedge \ y_2 \in y_3) \longrightarrow y_1 \in y_3.$$

  .

  totality: Given any $y \in x$ and $z \in x$, since $x \in \alpha$ and α is transitive, we have $y \in \alpha$ and $z \in \alpha$ and since $(\alpha, \in_\alpha)$ is a well-ordering we obtain

  $$y \in z \ \vee \ y = z \ \vee \ z \in y.$$

  well-foundedness: Given any $y \subseteq x$ with $y \neq \varnothing$, by transitivity of α, for all $z \in y$, one has $z \in \alpha$. Therefore, $y \subseteq \alpha$ and since $(\alpha, \in_\alpha)$ is a well-ordering, there exists a $\in_\alpha$-minimal inside y. Since $\in_x \subseteq \in_\alpha$, a is also $\in_x$-minimal inside y.

- x is transitive:

  Consider $z \in y$ and $y \in x$, then by transitivity of α, one has x, y, z all belong to α. Since $(\alpha, \in_\alpha)$ is an ordering, one has

  $$(z \in y \ \wedge \ y \in x) \longrightarrow z \in x.$$

- $x = [\underline{x}]_\alpha^{\in_\alpha}$ holds by the very definition of $[\underline{x}]_\alpha^{\in_\alpha}$.

(2) We assume α and β are ordinals, $(\alpha, \in_\alpha) \approx (\beta, \in_\beta)$. Let $f : (\alpha, \in_\alpha) \to (\beta, \in_\beta)$ be the (unique) isomorphism. If $f(\gamma) = \gamma$ holds for all $\gamma \in \alpha$, then

$$\beta = ran(f) = \{f(\gamma) \mid \gamma \in \alpha\} = \{\gamma \mid \gamma \in \alpha\} = \alpha.$$

So, towards a contradiction, we assume f is not the identity function and consider γ the $\in$-least element inside α such that $f(\gamma) \neq \gamma$. Notice that since f is an isomorphism, for every θ one has $\theta \in \gamma \longleftrightarrow f(\theta) \in f(\gamma)$. Therefore, $\theta \in \gamma \longrightarrow \theta \in f(\gamma)$, i.e., $\gamma \subseteq f(\gamma)$. Also, for every θ one has $f^{-1}(\theta) \in \gamma \longleftrightarrow \theta \in f(\gamma)$. Hence if $\theta \in f(\gamma)$, then $f^{-1}(\theta) \in \gamma$ and $f^{-1}(\theta) \in \gamma$ implies $f\big(f^{-1}(\theta)\big) = f^{-1}(\theta)$, so that $\theta = f^{-1}(\theta)$, hence $\theta \in \gamma$, which shows $f(\gamma) \subseteq \gamma$. So we have come to $\gamma = f(\gamma)$, which contradicts our hypothesis.

(3) If α and β are ordinals, then $(\alpha, \in_\alpha)$ and $(\beta, \in_\beta)$ are well-orderings. By Theorem 29 only one of the following occurs:

  (a) $(\alpha, \in_\alpha) \approx (\beta, \in_\beta)$ which leads to $\alpha = \beta$ by Theorem 34 (2).

  (b) $\exists \gamma \in \alpha \quad ([\underline{\gamma}]_\alpha^{\in_\alpha}, \in_\alpha) \approx (\beta, \in_\beta)$. By Theorem 34 (1) γ is an ordinal and $\gamma = [\underline{\gamma}]_\alpha^{\in_\alpha}$. So that $(\gamma, \in_\alpha) \approx (\beta, \in_\beta)$ holds which leads to $\gamma = \beta$ by Theorem 34 (2). Hence $\beta \in \alpha$ is satisfied.

(c) $\exists \gamma \in \beta \ \ (\alpha, \in_\alpha) \approx (\lceil \underline{\gamma} \rceil_\beta^{\in_\beta}, \in_\beta)$. By Theorem 34 (1) γ is an ordinal and $\gamma = \lceil \underline{\gamma} \rceil_\beta^{\in_\beta}$. So that $(\alpha, \in_\alpha)\gamma(\beta, \in_\beta)$ holds and by Theorem 34 (2) $\gamma = \alpha$ holds as well, which shows $\alpha \in \beta$.

(4) holds because γ being an ordinal is a transitive set.

(5) Since A is a non-empty set of ordinals, take any $\alpha \in A$. Then, either

$$\forall \beta \in A \ \ \big(\alpha \in \beta \ \vee \ \alpha = \beta\big)$$

in which case we are done, or

$$\exists \beta \in A \ \ \big(\alpha \notin \beta \ \wedge \ \alpha \neq \beta\big).$$

By Theorem 34 (3) this leads to $\exists \beta \in A \ \ \beta \in \alpha$. i.e., $A \cap \alpha \neq \varnothing$. Let δ be the $\in$-minimal element in $A \cap \alpha$, it satisfies

$$\forall \beta \in A \ \ \big(\delta \in \beta \ \vee \ \delta = \beta\big).$$

$\square$ 34

Example 35. *Notice that there is a formula $\varphi_{\mathbf{On}}$ with a single free variable x such that $\varphi_{\mathbf{On}}(x)$ if and only if x is an ordinal. For instance $\varphi_{\mathbf{On}}$ is the conjunction of the following 5 formulas:*

(1) $\forall z \forall y \ \ \big((z \in y \ \wedge \ y \in x) \longrightarrow z \in x\big)$ — *x is a transitive set*

(2) $\forall y \, (y \in x \longrightarrow \neg y \in y)$ — *$(x, \in_x)$ is irreflexive*

(3) $\forall y_1 \in x \ \forall y_2 \in x \ \forall y_3 \in x \ \ \big((y_1 \in y_2 \ \wedge \ y_2 \in y_3) \longrightarrow y_1 \in y_3\big)$ — *$(x, \in_x)$ is transitive*

(4) $\forall y \in x \ \forall z \in x \ (y \in z \ \vee \ z \in y \ \vee \ y = z)$ — *$(x, \in_x)$ is total*

(5) $\forall y \subseteq x \ \big(y \neq \varnothing \longrightarrow \exists z \in y \, \forall z' \in y \ \ (z \in z' \ \vee \ z = z')\big)$ — *$(x, \in_x)$ is well-founded*

From now on, we may use the formulation "*x is an ordinal*" for $\varphi_{\mathbf{On}}(x)$.

Theorem 36 (There is no set that contains all ordinals)**.**

$$\mathbf{ZF} \vdash_c \neg \exists A \forall \alpha \ \ \big(\alpha \text{ “is an ordinal”} \longrightarrow \alpha \in A\big).$$

Proof of Theorem 36: Towards a contradiction, we work within the theory

$$\mathbf{ZF} \cup \left\{\exists A \forall \alpha \quad \left(\alpha \text{ ``is an ordinal''} \longrightarrow \alpha \in A\right)\right\}.$$

By **Comprehension Schema** and **Extensionality** this theory yields

$$\exists A \forall \alpha \quad \left(\alpha \text{ ``is an ordinal''} \longleftrightarrow \alpha \in A\right).$$

Let us call **On** this supposedly set of all ordinals. we have

- **On** is transitive by Theorem 34 (1).
- $(\mathbf{On}, \in_{\mathbf{On}})$ is a well-ordering since it satisfies
  - irreflexivity: for all ordinal α, $\alpha \notin \alpha$ holds for $(\alpha, \in_\alpha)$ is irreflexive;
  - transitivity: holds by Theorem 34 (3);
  - totality: holds by Theorem 34 (4).
  - well-foundedness: holds by Theorem 34 (5).

Since it is both transitive and well-ordered by $\in$, **On** is an ordinal, which leads to $\mathbf{On} \in \mathbf{On}$, a contradiction to irreflexivity.

□ 36

Lemma 37 (A transitive set of ordinals is an ordinal)**.**
If A is a set of ordinals that satisfies

$$\forall \alpha \in A \ \forall \beta \in \alpha \quad \beta \in A,$$

then A is an ordinal.

Proof of Lemma 37: A is transitive and by Theorem 34 $(A, \in_A)$ is a well-ordering, therefore A is an ordinal.

□ 37

Theorem 38. *If $(A, <_A)$ is a well-ordering, then there exists some unique ordinal α such that*

$$(A, <_A) \approx (\alpha, \in_\alpha).$$

Proof of Theorem 38:

Unicity: assume that α and β are two ordinals that satisfy

$$(A, <_A) \approx (\alpha, \in_\alpha) \text{ and } (A, <_A) \approx (\beta, \in_\beta).$$

This leads to

$$(\alpha, \in_\alpha) \approx (\beta, \in_\beta)$$

and to $\alpha = \beta$ by Theorem 34 (2).

Existence: we consider

$$B = \left\{a \in A \mid \exists\alpha \text{ “ordinal” } (\lceil \underline{a} \rceil_A^{<_A}, <_A) \approx (\alpha, \in_\alpha)\right\}$$

and let f be the functional that assigns to each $a \in B$ the *unique* ordinal α such that $(\lceil \underline{a} \rceil_A^{<_A}, <_A) \approx (\alpha, \in_\alpha)$. By **Comprehension Schema**, $f[B]$ is a set.

- $f[B]$ is *transitive* because if $b \in a \in B$, then $(\lceil \underline{a} \rceil_A^{<_A}, <_A) \approx (\alpha, \in_\alpha)$ holds for some α. Let $g : \lceil \underline{a} \rceil_A^{<_A} \to \alpha$ be the unique isomorphism and $\beta = g(b)$. One has then $g \upharpoonright \lceil \underline{b} \rceil_A^{<_A} : \lceil \underline{b} \rceil_A^{<_A} \to \beta$ is an isomorphism that witnesses $(\lceil \underline{b} \rceil_A^{<_A}, <_A) \approx (\beta, \in_\beta)$.
- $f[B]$ is a set of ordinals.

Therefore, by Lemma 37, $f[B]$ is some ordinal α.

It remains to show that $B = A$. Towards a contradiction, we assume that $A \setminus B \neq \varnothing$ and let a be the $<_A$-least element in $A \setminus B$. It follows that $\lceil \underline{a} \rceil_A^{<_A}$ is an initial segment of $(A, <_A)$, therefore $f\left[\lceil \underline{a} \rceil_A^{<_A}\right]$ is an initial segment of $f[B] = \alpha$, hence a transitive set of ordinals, i.e., an ordinal δ. Since $(\lceil \underline{a} \rceil_A^{<_A}, <_A) \approx (\delta, \in_\delta)$, it follows that $a \in B$, a contradiction.

$\square$ 38

2.3 Order Type, Successor Ordinals and Limit Ordinals

Definition 39 (Type)**.** *Given any well-ordering* $(A, <_A)$*,*

$$\textit{type}\,(A, <_A) \ \textit{ is the unique ordinal } \alpha \textit{ such that } (A, <_A) \approx (\alpha, \in_\alpha).$$

Definition 40 (Supremum)**.** *If A is a set of ordinals, then*

$$\sup A = \bigcup A.$$

Definition 41 (Infimum)**.** *If A is a non-empty set of ordinals, then*

$$\inf A = \bigcap A.$$

Notation 42. *From now on,*

- *we use greek letters such as $\alpha, \beta, \gamma, \delta$ to denote exclusively ordinals.*
- *We also write $\alpha < \beta$ for $\alpha \in \beta$, and $\alpha \leqslant \beta$ for $\alpha \in \beta \vee \alpha = \beta$.*

Lemma 43.

(1) $\forall\alpha\ \forall\beta\ \ (\alpha \leqslant \beta \longleftrightarrow \alpha \subseteq \beta)$.

(2) If A is a set of ordinals, then $\sup A$ *is the least ordinal α that satisfies $\forall\beta \in A\ \ \beta \leqslant \alpha$.*

(3) If A is a non-empty set of ordinals, then $\inf A$ *is the least ordinal of A.*

Proof of Lemma 43:

(1) holds since ordinals are transitive sets.

(2) $\sup A = \bigcup A$ is a transitive set of ordinals, hence an ordinal α. Clearly, if $\beta \in A$, then $\beta \subseteq \bigcup A$, hence by Lemma 43 (1), $\beta \leqslant \alpha$, which shows $\forall\beta \in A\ \ \beta \leqslant \alpha$. For all $\gamma < \alpha$, $\gamma \in \alpha = \bigcup A$, hence there exists some $\beta \in A$ such that $\gamma \in \beta$. Since $\gamma < \alpha$ implies $\gamma < \beta$ for some $\beta \in A$, α is the least ordinal that satisfies $\forall\beta \in A\ \ \beta \leqslant \alpha$.
least ordinal δ that satisfies $\forall\alpha \in A\ \ \alpha \leqslant \delta$.

(3) If A is a non-empty set of ordinals, let α be the least ordinal in A, then one has

$$\alpha \subseteq \bigcap A \subseteq \alpha.$$

.

□ 43

Definition 44 (Successor)**.** *Given any ordinal α,*

$$S\alpha = \alpha \cup \{\alpha\}\,.$$

Lemma 45. *For each ordinal α,*

(1) $S\alpha$ is an ordinal. *(2) $\alpha < S\alpha$.* *(3) $\forall \beta \ \ (\beta < S\alpha \rightarrow \beta \leqslant \alpha)$.*

Proof of Lemma 45: Immediate.

$\square$ 45

Definition 46 (Limit Ordinal)**.** *Given any ordinal α,*

- *α is a* successor ordinal *if $\alpha = S\beta$ for some ordinal β;*
- *α is a* limit ordinal *if $\alpha \neq \varnothing$ and α is not a successor ordinal.*

Example 47. *The following sets are successor ordinals:*

- $\{\varnothing\}$
- $\{\varnothing, \{\varnothing\}\}$
- $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\}$
- $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$

Notation 48. *We agree to use the following denominations for the following sets:*

- $0 := \varnothing$
- $1 := S\varnothing = \{\varnothing\} = \{0\}$
- $2 := S1 = S\{\varnothing\} = \{\varnothing, \{\varnothing\}\} = \{0, 1\}$
- $3 := S2 = S\{\varnothing, \{\varnothing\}\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = \{0, 1, 2\}$.
- *Etc.*

So, for the moment, we are able to produce every single integer. But we cannot exhibit the set of all integers, i.e., $\{0, 1, 2, 3, \ldots, n, Sn, \ldots\}$. For this we need the axiom of **Infinity**.

Infinity claims that there exists a set that contains the empty set and is closed under $x \mapsto Sx$. Namely,

$$\exists x \big(\exists y \, (y \in x \ \wedge \ \forall z \, z \notin y) \ \wedge \ \forall z (z \in x \rightarrow z \cup \{z\} \in x) \big)$$

Therefore, this set contains infinitely many elements since it contains — at least — all the ones we would describe as the integers.

Definition 49 (Integer). *An ordinal α is an* integer *if and only if*

$$\forall \beta \leqslant \alpha \quad \beta = \big(\varnothing \ \vee \ \exists \gamma \quad \beta = S\gamma \big).$$

Notation 50 ($\boldsymbol{\omega}$). *The set of all integers is denoted ω.*

By **Comprehension Schema** and **Extensionality**, we obtain ω. Notice that ω is a transitive set of ordinals, therefore it is an ordinal. Moreover, ω is a limit ordinal since for each $\alpha \in \omega$, α is an integer and so is $S\alpha$, hence ω is not a successor ordinal.

2.4 Classes

A collection of sets that can be described by a 1st order $\{\in, =\}$ −formula is called a *class*. Formally, a class $\mathbf{C}$ is nothing but a formula with one free variable — that may or may not have other free variables that behave as parameters — $\varphi_{\mathbf{C}}$. But we may now use expressions such as $x \in \mathbf{C}$ or $x \subseteq \mathbf{C}$ to denote respectively $\varphi_{\mathbf{C}}(x)$ or $\forall y \in x \ \ \varphi_{\mathbf{C}}(y)$.

Example 51.

- $\mathbf{V}$ *stands for the class of all sets. i.e.,*

$$\mathbf{V} = \{x \mid x = x\}.$$

- $\mathbf{On}$ *stands for the class of all ordinals. i.e.,*

$$\mathbf{On} = \{x \mid x \text{ “is an ordinal”}\}.$$

A class that is not a set is called a *proper class*. For instance, $\mathbf{V}$ and $\mathbf{On}$ are both proper classes. If $\mathbf{C}$ and $\mathbf{C}'$ are classes, we may write $\mathbf{C} \subseteq \mathbf{C}'$ which stands for $\forall x \, \big(\varphi_{\mathbf{C}}(x) \longrightarrow \varphi_{\mathbf{C}'}(x) \big)$ assuming $\varphi_{\mathbf{C}}$ and $\varphi_{\mathbf{C}'}$ are the two formulas that characterize respectively the classes $\varphi_{\mathbf{C}}$ and $\varphi_{\mathbf{C}'}$. We may also write $\mathbf{C} \cap \mathbf{C}'$ for the class characterized by the formula $\varphi_{\mathbf{C}} \ \wedge \ \varphi_{\mathbf{C}'}$. Or even $\mathbf{F} : \mathbf{C} \rightarrow \mathbf{C}'$ to denote a functional from $\mathbf{C}$ to $\mathbf{C}'$. i.e., a formulas $\varphi_{\mathbf{F}(x,y)}$ — with at least two free variables x, y — that satisfies $\forall x \, \Big(\varphi_{\mathbf{C}}(x) \longrightarrow \exists! y \ \ \big(\varphi_{\mathbf{C}'}(y) \ \wedge \ \varphi_{\mathbf{F}}(x, y) \big) \Big)$.

2.5 Transfinite Induction and Recursion

In the following we prove results about classes. Since these classes are defined by formulas and one cannot quantify over formulas but rather over sets, the following theorems are in fact not really theorems for $\mathbf{Z}$. Precisely, each such theorem is a *"theorem schema"*, i.e., a single theorem for each class that we consider.

Theorem 52 (Transfinite Induction)**.** *if* $\mathbf{C} \subseteq \mathbf{On}$ *is a non-empty class, then*

$$\mathbf{C} \text{ has a } \in\text{-least element.}$$

Proof of Theorem 52: Take any $\alpha \in \mathbf{C}$.

- if $\alpha \cap \mathbf{C} = \varnothing$, then α is the $\in$-least element in $\mathbf{C}$.
- if $\alpha \cap \mathbf{C} \neq \varnothing$, then by Theorem 34 (5) there exists some $\in$-least element inside α. This element is also the $\in$-least element inside $\mathbf{C}$.

□ 52

Theorem 53 (Transfinite Recursion)**.** *Given any* $\mathbf{F} : \mathbf{V} \to \mathbf{V}$*, there exists a unique* $\mathbf{G} : \mathbf{On} \to \mathbf{V}$ *such that for each ordinal* α

$$\forall \alpha \quad \mathbf{G}(\alpha) = \mathbf{F}(\mathbf{G} \restriction \alpha).$$

Proof of Theorem 53:

Uniqueness: Assume there exist two different functionals $\mathbf{G_1}$ and $\mathbf{G_2}$. By Theorem 52 the non-empty class $\{\alpha \in \mathbf{On} \mid \mathbf{G_1}(\alpha) \neq \mathbf{G_2}(\alpha)\}$ has a least element β. By construction, one comes to the following contradiction:

$$\mathbf{G_1}(\beta) = \mathbf{F}(\mathbf{G_1} \restriction \beta) = \mathbf{F}(\mathbf{G_2} \restriction \beta) = \mathbf{G_2}(\beta).$$

Existence: we construct functions that are approximations of $\mathbf{G}$ on some proper initial segment of the ordinals. i.e., for each ordinal β, we construct $g_\beta : \beta \to \mathbf{V}$ such that

$$\forall \alpha < \beta \quad g_\beta(\alpha) = \mathbf{F}(g_\beta \restriction \alpha).$$

So, g_β is a function with domain β and ranges over some set obtained by one instance of the **Replacement Schema**. Notice then, that for each β, such an approximation g_β is unique by the same argument as the one used in the proof of the uniqueness of $\mathbf{G}$. Also, notice that if $\beta < \gamma$, then, by uniqueness of g_β, one has $g_\gamma \restriction \beta = g_\beta$. Finally define $\mathbf{G}(\alpha) = g_\beta(\alpha)$ for some $\beta > \alpha$.

□ 53

Example 54. *The addition on ordinals* $\mathbf{Add} : \mathbf{On} \times \mathbf{On} \to \mathbf{On}$ *is defined by making use — for each ordinal* α *— of* $\mathbf{Add}_{\boldsymbol{\alpha}} : \mathbf{On} \to \mathbf{On}$*, defined by transfinite recursion*

- $\mathbf{Add}_{\boldsymbol{\alpha}}(0) = \alpha$
- $\mathbf{Add}_{\boldsymbol{\alpha}}(S\beta) = S\mathbf{Add}_{\boldsymbol{\alpha}}(\beta)$
- $\mathbf{Add}_{\boldsymbol{\alpha}}(\beta) = \sup\{\mathbf{Add}_{\boldsymbol{\alpha}}(\gamma) \mid \gamma < \beta\}$ *for* β *a limit ordinal.*

Finally, one defines $\mathbf{Add}(\alpha, \beta) = \mathbf{Add}_{\boldsymbol{\alpha}}(\beta)$

Of course we write $\alpha + \beta$ *instead of* $\mathbf{Add}(\alpha, \beta)$*.*

2.6 Ordinal Addition, Multiplication and Exponentiation

Definition 55. *The ordinal addition* $: \mathbf{On} \times \mathbf{On} \to \mathbf{On}$ *is defined by*

initial step: $\alpha + 0 = \alpha$

successor step: $\alpha + S\beta = S(\alpha + \beta)$

limit step: $\alpha + \beta = \sup\{\alpha + \gamma \mid \gamma < \beta\}$ *(for* β *a limit ordinal).*

Notice that

- the ordinal addition restricted to $\omega \times \omega$ is nothing but the usual addition on the integers.
- the addition is not commutative:
  - $1 + \omega = \sup\{1 + n \mid n < \omega\} = \omega$
  - $\omega + 1 = S\omega = \omega \cup \{\omega\}$.
- the addition is associative.

Exercise 56. *Given any ordinals* $\alpha \leqslant \beta$*, there exists some unique ordinal* δ *such that* $\alpha + \delta = \beta$ *i.e.,*

$$\forall \alpha, \beta \in \mathbf{On} \quad \big(\alpha \leqslant \beta \longrightarrow \exists! \delta \in \mathbf{On} \quad \alpha + \delta = \beta\big).$$

Exercise 57. *This definition of the ordinal addition via recursion is equivalent to the following: Given any ordinals α, β:*

$$\alpha + \delta = \mathit{type}\left(\alpha \times \{0\} \cup \beta \times \{1\}, \lessapprox\right)$$

where

$$(\xi, i) \lessapprox (\xi', i') \iff \begin{cases} i < i' \\ \text{or} \\ i = i' \text{ and } \xi < \xi'. \end{cases}$$

Definition 58. *The ordinal multiplication* $: \mathbf{On} \times \mathbf{On} \to \mathbf{On}$ *is defined by transfinite recursion by*

initial step: $\alpha \cdot 0 = 0$

successor step: $\alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha$

limit step: $\alpha \cdot \beta = \sup\{\alpha \cdot \gamma \mid \gamma < \beta\}$ *(for β a limit ordinal).*

Notice that

- the ordinal multiplication restricted to $\omega \times \omega$ is nothing but the usual multiplication on the integers.
- the multiplication is not commutative:
  - $2 \cdot \omega = \sup\{2 \cdot n \mid n < \omega\} = \omega$
  - $\omega \cdot 2 = (\omega \cdot 1) + \omega = \big((\omega \cdot 0) + \omega\big) + \omega = \big(0 + \omega\big) + \omega = \big(\sup\{0 + n \mid n < \omega\}\big) + \omega = \omega + \omega.$
- the multiplication is associative.

Exercise 59. *This definition of the ordinal multiplication by recursion is equivalent to the following: Given any ordinals α, β:*

$$\alpha \cdot \delta = \mathit{type}\left(\alpha \times \beta, \lessapprox\right)$$

where

$$(\xi, \zeta) \lessapprox (\xi', \zeta') \iff \begin{cases} \zeta < \zeta' \\ \text{or} \\ \zeta = \zeta' \text{ and } \xi < \xi'. \end{cases}$$

Exercise 60. *The Euclidean division: Given any ordinals* $0 < \alpha, \beta$, *there exist unique ordinals* γ *and* $\delta < \beta$ *such that*

$$\alpha = \beta \cdot \gamma + \delta.$$

Definition 61. *The ordinal exponentiation* : $\mathbf{On} \times \mathbf{On} \to \mathbf{On}$ *is defined by transfinite recursion by*

initial step: $\alpha^0 = 1$

successor step: $\alpha^{(\beta+1)} = \alpha^\beta \cdot \alpha$

limit step: $\alpha^\beta = \sup \{\alpha^\gamma \mid \gamma < \beta\}$ *(for* β *a limit ordinal).*

Proposition 62. *Given any ordinals* α, β, γ,

$$\alpha^\beta \cdot \alpha^\gamma = \alpha^{\beta+\gamma}.$$

Proof of Proposition 62: The proof is by induction on γ.

initial step: $\alpha^\beta \cdot \alpha^0 = \alpha^\beta \cdot 1 = \alpha^\beta = \alpha^{\beta+0}$

successor step: $\alpha^\beta \cdot \alpha^{(\gamma+1)} = \alpha^\beta \cdot (\alpha^\gamma \cdot \alpha) = (\alpha^\beta \cdot \alpha^\gamma) \cdot \alpha = \alpha^{\beta+\gamma} \cdot \alpha = \alpha^{(\beta+\gamma)+1} = \alpha^{\beta+(\gamma+1)}$

limit step: $\alpha^\beta \cdot \alpha^\gamma = \alpha^\beta \cdot \sup \left\{\alpha^\delta \mid \delta < \gamma\right\} = \sup \left\{\alpha^\beta \cdot \alpha^\delta \mid \delta < \gamma\right\} = \sup \left\{\alpha^{\beta+\delta} \mid \delta < \gamma\right\} = \alpha^{\beta+\gamma}$ (for γ a limit ordinal).

□ 62

Proposition 63. *Given any ordinals* α, β, γ,

$$\left(\alpha^\beta\right)^\gamma = \alpha^{\beta\cdot\gamma}.$$

Proof of Proposition 63: The proof is by induction on γ.

initial step: $\left(\alpha^{\beta}\right)^{0} = 1 = \alpha^{0} = \alpha^{\beta\cdot 0}$

successor step: $\left(\alpha^{\beta}\right)^{(\gamma+1)} = \left(\alpha^{\beta}\right)^{\gamma} \cdot \alpha^{\beta} = \alpha^{\beta\cdot\gamma} \cdot \alpha^{\beta} = \alpha^{\beta\cdot\gamma+\beta} = \alpha^{\beta\cdot(\gamma+1)}$

limit step: $\left(\alpha^{\beta}\right)^{\gamma} = \sup\left\{\left(\alpha^{\beta}\right)^{\delta} \mid \delta < \gamma\right\} = \sup\left\{\left(\alpha^{\beta\cdot\delta}\right) \mid \delta < \gamma\right\} = \alpha^{\beta\cdot\gamma}$ (for γ a limit ordinal).

$\square$ 63

Exercise 64. *The Cantor Normal Form: Given any ordinal* $0 < \alpha$, *there exist a unique integer* k, *a unique sequence of non-null integers* $(n_i)_{i\leqslant k}$ *and a unique sequence* $(\alpha_i)_{i\leqslant k}$ *that satisfies*

$$\alpha \geqslant \alpha_k > \ldots > \alpha_1 > \alpha_0 \geqslant 0$$

such that

$$\alpha = \omega^{\alpha_k} \cdot n_k + \ldots + \omega^{\alpha_0} \cdot n_0 \quad \left(\sum_{i=k}^{i=0} \omega^{\alpha_i} \cdot n_i\right).$$

Chapter 3

Extension by Definition and Conservative Extension

3.1 Extension by Definitions

This section deals with what we really do when we introduce new symbols that were not in the original language of set theory. For instance, so far we have introduced

- constant symbols such as

  (1) $\varnothing$ (2) ω (3) $1, 2, 3, \ldots$

- function symbols such as

  (1) $\{\ \} : \begin{array}{ccc} \mathbf{V} & \to & \mathbf{V} \\ x & \mapsto & \{x\} \end{array}$ (4) $\cap : \begin{array}{ccc} \mathbf{V}\times\mathbf{V} & \to & \mathbf{V} \\ (x,y) & \mapsto & x \cap y \end{array}$ (7) $+ : \begin{array}{ccc} \mathbf{On}\times\mathbf{On} & \to & \mathbf{On} \\ (\alpha,\beta) & \mapsto & \alpha+\beta \end{array}$

  (2) $\times : \begin{array}{ccc} \mathbf{V}\times\mathbf{V} & \to & \mathbf{V} \\ (x,y) & \mapsto & x \times y \end{array}$ (5) $S : \begin{array}{ccc} \mathbf{On} & \to & \mathbf{On} \\ \alpha & \mapsto & \mathrm{S}\alpha \end{array}$ (8) $\cdot : \begin{array}{ccc} \mathbf{On}\times\mathbf{On} & \to & \mathbf{On} \\ (\alpha,\beta) & \mapsto & \alpha\cdot\beta \end{array}$

  (3) $\cup : \begin{array}{ccc} \mathbf{V}\times\mathbf{V} & \to & \mathbf{V} \\ (x,y) & \mapsto & x \cup y \end{array}$ (6) $\bigcup : \begin{array}{ccc} \mathbf{V} & \to & \mathbf{V} \\ x & \mapsto & \bigcup x \end{array}$ (9) $\bigcap : \begin{array}{ccc} \mathbf{V} & \to & \mathbf{V} \\ x & \mapsto & \bigcap x \end{array}$

- relation symbols such as $\subseteq$

Are we entitled to do so? It is time to clarify this point. Indeed, we started with the original language of set theory whose signature — apart from equality — is reduced to $\{\in\}$, and gradually, we switched to the languages $\{\in, \varnothing\}$, $\{\in, \varnothing, \subseteq\}$, $\{\in, \varnothing, \subseteq, \{\}\}$, etc. But each of these introduction a new symbol was accompanied with a precise definition of it in a language that did not make use of it. For instance, when $\varnothing$ was introduced, it was as a shortening for the formula with one free variable: $\forall z \ \ z \notin x$. This made sense since we had proved the existence of the empty set:

namely $\mathbf{Z} \vdash_c \exists x\ \forall z\ \ z \notin x$. So, strictly speaking, we did not absolutely need the introduction of this constant symbol. We could have still worked in the original language by, instead of having a formula of the form $\varphi(x_1, \ldots, x_n, \varnothing)$, replacing it with the formula $\exists x \Big(\forall z\ \ z \notin x \ \wedge \varphi(x_1, \ldots, x_n, x)\Big)$.

This way of introducing new symbols to the language and modifying the original theory accordingly is called an *extension by definition*. It leads to the extension of the original theory to a richer language. But making sure that the extended theory — which of course proves strictly more formulas than the original one (it proves $\varnothing = \varnothing$ for instance) — does not prove any other formula in the original language than the ones that are provable in the original theory. In other words, the extended theory works as a *conservative extension* of the original one.

In the following, we identify constant symbols with with 0-arity function symbols.

Definition 65. *Let $\mathcal{L}$ be any first order language and $\mathscr{T}$ any $\mathcal{L}$-theory. An $\mathcal{L}'$-theory $\mathscr{T}'$ is an* ***extension by definition*** *of $\mathscr{T}$ if there exists some integer n, languages $(\mathcal{L}_i)_{i \leqslant n}$ and for each $i \leqslant n$ a $\mathcal{L}_i$-theory $\mathscr{T}_i$ such that*

- $\mathcal{L} = \mathcal{L}_0 \subsetneq \mathcal{L}_1 \subsetneq \ldots \subsetneq \mathcal{L}_n = \mathcal{L}'$
- $\mathscr{T} = \mathscr{T}_0 \subsetneq \mathscr{T}_1 \subsetneq \ldots \subsetneq \mathscr{T}_n = \mathscr{T}'$
- *for each $i < n$, $\mathcal{L}_{i+1} \setminus \mathcal{L}_i = \{R\}$ or $\mathcal{L}_{i+1} \setminus \mathcal{L}_i = \{f\}$ where R and f stand for respectively a k_i-relation symbol or a k_i-function symbol.*
    - *if $\mathcal{L}_{i+1} \setminus \mathcal{L}_i = \{R\}$, then there exists some $\mathcal{L}_i$-formula $\varphi_{(x_1,\ldots,x_{k_i})}$ whose free variables are among $x_1, \ldots, x_{k_i}$, such that*

    $$\mathscr{T}_{i+1} = \mathscr{T}_i \cup \Big\{\forall x_1 \ldots \forall x_{k_i} \quad \big(\varphi_{(x_1,\ldots,x_{k_i})} \longleftrightarrow R(x_1, \ldots, x_{k_i})\big)\Big\}$$

    - *if $\mathcal{L}_{i+1} \setminus \mathcal{L}_i = \{f\}$, then there exists some $\mathcal{L}_i$-formula $\varphi_{(x_1,\ldots,x_{k_i},y)}$ whose free variables are among $x_1, \ldots, x_{k_i}, y$, such that*

    $$\mathscr{T}_{i+1} = \mathscr{T}_i \cup \Big\{\forall x_1 \ldots \forall x_{k_i}\ \forall y\ \ \big(\varphi_{(x_1,\ldots,x_{k_i},y)} \longleftrightarrow f(x_1, \ldots, x_{k_i}) = y\big)\Big\}$$

    and

    $$\mathscr{T}_i \vdash_c \forall x_1 \ldots \forall x_{k_i}\ \exists! y\ \ \varphi_{(x_1,\ldots,x_{k_i},y)}.$$

Lemma 66. *Let $\mathcal{L}$ be any first order language, $\mathscr{T}$ any $\mathcal{L}$-theory and $\mathscr{T}'$ any extension by definition of $\mathscr{T}$.*
If $\psi'_{(x_1,\ldots,x_k)}$ is any $\mathcal{L}'$-formula whose free variables are among $x_1,\ldots,x_k$, then there exists some $\mathcal{L}$-formula $\psi_{(x_1,\ldots,x_k)})$ whose free variables are among $x_1,\ldots,x_k$ such that

$$\mathscr{T}' \vdash_c \forall x_1 \ldots \forall x_k \ \ \big(\psi'_{(x_1,\ldots,x_k)} \longleftrightarrow \psi_{(x_1,\ldots,x_k)}\big).$$

Proof of Lemma 66: We assume $\mathscr{T}'$ is as in Definition 65, in particular $\mathscr{T}' = \mathscr{T}_n$.
Without loss of generality, we may assume that every atomic formula occurring inside $\psi'_{(x_1,\ldots,x_k)}$ is of the following two forms:

- $R(z_1,\ldots,z_m)$ with $z_1,\ldots,z_m$ variables
- $t = z$ with t a term and z a variable.

Indeed,

- if $R(t_1,\ldots,t_m)$ is an atomic formula whose free variables are among $x_1,\ldots,x_k$ and $t_1,\ldots,t_m$ are terms, then for different variables $z_1 \ldots z_m$ such that $\{z_1,\ldots,z_m\} \cap \{x_1,\ldots,x_k\} = \varnothing$ we have

$$\vdash_c R(t_1,\ldots,t_m) \longleftrightarrow \forall z_1 \ldots \forall z_m \ \ \left(\left(\bigwedge_{1\leqslant j\leqslant m} t_j = z_j\right) \longrightarrow R(z_1,\ldots,z_m)\right)$$

  or equivalently

$$\vdash_c R(t_1,\ldots,t_m) \longleftrightarrow \exists z_1 \ldots \exists z_m \ \ \left(\left(\bigwedge_{1\leqslant j\leqslant m} t_j = z_j\right) \wedge R(z_1,\ldots,z_m)\right)$$

- if $t_1 = t_2$ is an atomic formula whose free variables are among $x_1,\ldots,x_k$ and t_1, t_2 are terms, then for different variables z_1, z_2 such that $\{z_1,z_2\} \cap \{x_1,\ldots,x_k\} = \varnothing$ we have

$$\vdash_c t_1 = t_2 \longleftrightarrow \forall z_1 \forall z_2 \ \ \big((t_1 = z_1 \wedge \ t_2 = z_2) \longrightarrow z_1 = z_2\big)$$

  or equivalently

$$\vdash_c t_1 = t_2 \longleftrightarrow \exists z_1 \exists z_2 \ \ \big(t_1 = z_1 \ \wedge \ t_2 = z_2 \ \wedge \ z_1 = z_2\big)$$

The proof is by induction on n.

if n = 0: $\psi = \psi'$ works.

if $\mathbf{n = i + 1}$: the proof is by induction on the height of $\psi'_{(x_1,\ldots,x_k)}$. If $\psi'_{(x_1,\ldots,x_k)}$ is already some $\mathcal{L}_i$-formula, then by induction hypothesis there exists some $\mathcal{L}$-formula that satisfies the requirements. So we assume that $\psi'_{(x_1,\ldots,x_k)}$ is not a $\mathcal{L}_i$-formula. The proof is now by induction on $ht(\psi')$.

if $ht(\psi') = 0$: we distinguish between $\psi' := R(x_1, \ldots, x_k)$ and $\psi' := t = z$.

if $\psi' := \mathbf{R}(\mathbf{x_1}, \ldots, \mathbf{x_k})$: there exists

- some $\mathcal{L}_i$-formula φ_R such that

$$\mathscr{T}_{i+1} = \mathscr{T}_i \cup \left\{\forall x_1 \ldots \forall x_k \quad \left(\varphi_R(x_1, \ldots, x_k) \longleftrightarrow R(x_1, \ldots, x_k)\right)\right\}$$

- and by induction hypothesis, some $\mathcal{L}$-formula ψ such that

$$\mathscr{T}_i \vdash_c \forall x_1 \ldots \forall x_k \quad \left(\psi_{(x_1,\ldots,x_k)} \longleftrightarrow \varphi_R(x_1, \ldots, x_k)\right).$$

which shows that

$$\mathscr{T}_{i+1} \vdash_c \forall x_1 \ldots \forall x_k \quad \left(\psi(x_1, \ldots, x_k) \longleftrightarrow R(x_1, \ldots, x_k)\right).$$

if $\psi' := \mathbf{t} = \mathbf{z}$: the proof goes by induction of $ht(t)$ (notice that $z \in \{x_1, \ldots, x_k\}$).

if $ht(\mathbf{t}) = 0$: one necessarily has $t := c$ where c is a constant symbol not in $\mathcal{L}_i$.

- some $\mathcal{L}_i$-formula φ_c such that

$$\mathscr{T}_{i+1} = \mathscr{T}_i \cup \left\{\forall y \quad \left(\varphi_c(y) \longleftrightarrow c = y)\right)\right\}$$

and

$$\mathscr{T}_i \vdash_c \exists! y \quad \varphi_c(y)$$

- and by induction hypothesis, some $\mathcal{L}$-formula ψ such that

$$\mathscr{T}_i \vdash_c \forall y \quad \left(\psi(y) \longleftrightarrow \varphi_c(y)\right).$$

which shows that

$$\mathscr{T}_{i+1} \vdash_c \forall z \quad \left(\psi(z) \longleftrightarrow c = z\right).$$

if $ht(\mathbf{t}) > 0$: there exist a function symbol f, some non-null integer r and terms $u_1, \ldots, u_r$ such that

$$t = f(u_1, \ldots, u_r).$$

if $\mathbf{f} \in \mathcal{L}_{\mathbf{i+1}} \smallsetminus \mathcal{L}_{\mathbf{i}}$: there exists some $\mathcal{L}_i$-formula $\varphi_f(x_1, \ldots, , x_r, y)$ such that

$$\begin{array}{rcl} \mathscr{T}_{i+1} & = & \mathscr{T}_i \cup \left\{\forall x_1 \ldots \forall x_r \ \forall y \quad \left(\varphi_f(x_1, \ldots, x_r, y) \longleftrightarrow f(x_1, \ldots, x_r) = y\right)\right\} \\ \mathscr{T}_i & \vdash_c & \forall x_1 \ldots \forall x_r \ \exists! y \quad \varphi_f(x_1, \ldots, , x_r, y). \end{array}$$

Since for all integers integer $1 \leqslant j \leqslant r$ the inequality $ht(u_j) < ht(t)$ holds, by induction hypothesis we have $\mathcal{L}_i$-formulas

$$\varphi_{u_1}(x_1, \ldots, x_k, y), \ldots, \varphi_{u_r}(x_1, \ldots, x_k, y),$$

such that for each integer $1 \leqslant j \leqslant r$

$$\begin{array}{lll} \mathscr{T}_{i+1} & \ni & \forall x_1 \ldots \forall x_k\ \forall y\ \big(\varphi_{u_j}(x_1, \ldots, x_k, y)) \longleftrightarrow u_1(x_1, \ldots, x_k) = y\big) \\ \mathscr{T}_i & \vdash_c & \forall x_1 \ldots \forall x_k\ \exists! y\ \ \varphi_{u_j}(x_1, \ldots, , x_k, y) \end{array}$$

and $\mathcal{L}$-formulas

$$\psi_{u_1}(x_1, \ldots, x_k, y), \ldots, \psi_{u_r}(x_1, \ldots, x_k, y)$$

such that for each integer $1 \leqslant j \leqslant r$

$$\begin{array}{lll} \mathscr{T}_{i+1} & \vdash_c & \forall x_1 \ldots \forall x_k\ \forall y\ \big(\varphi_{u_j}(x_1, \ldots, x_k, y) \longleftrightarrow \psi_{u_j}(x_1, \ldots, x_k, y)\big) \\ \mathscr{T}_{i+1} & \vdash_c & \forall x_1 \ldots \forall x_k\ \forall y\ \big(\varphi_{v_j}(x_1, \ldots, x_k, y) \longleftrightarrow \psi_{v_j}(x_1, \ldots, x_k, y).\big) \end{array}$$

Finally, since φ_f is a $\mathcal{L}_i$-formula, we obtain by induction hypothesis a $\mathcal{L}$-formula $\psi_f(x_1, \ldots, x_r, y)$ that satisfies

$$\mathscr{T}_i \quad \vdash_c \quad \forall x_1 \ldots \forall x_r\ \forall y\ \big(\varphi_f(x_1, \ldots, x_r, y) \longleftrightarrow \psi_f(x_1, \ldots, x_r, y)\big).$$

Remember that $z \in \{x_1, \ldots, x_k\}$, so $z = x_l$ holds for some $1 \leqslant l \leqslant k$
This leads to

$$\mathscr{T}_{i+1} \quad \vdash_c \quad \forall x_1 \ldots \forall x_k \left[\begin{array}{c} f(u_1, \ldots, u_r) = x_l \\ \longleftrightarrow \\ \forall y_{u_1} \ldots \forall y_{u_r} \\ \left[\left[\begin{array}{c} \bigwedge\limits_{1 \leqslant j \leqslant r} \psi_{u_j}(x_1, \ldots, x_r, y_{u_j}) \\ \wedge \\ \bigwedge\limits_{1 \leqslant j \leqslant r} \psi_{v_j}(x_1, \ldots, x_r, y_{v_j}) \end{array} \right] \longrightarrow \psi_f(y_{u_1}, \ldots, y_{u_r}, x_l) \right] \end{array} \right].$$

if $\mathbf{f} \in \mathcal{L}_\mathbf{i}$: there exists some integer $i' < i$ and some $\mathcal{L}_{i'}$-formula $\varphi_f(x_1, \ldots, , x_r, y)$ such that

$$\begin{array}{lll} \mathscr{T}_{i'+1} & = & \mathscr{T}_i \cup \big\{\forall x_1 \ldots \forall x_r\ \forall y\ \big(\varphi_f(x_1, \ldots, x_r, y) \longleftrightarrow f(x_1, \ldots, x_r) = y\big)\big\} \\ \mathscr{T}_{i'} & \vdash_c & \forall x_1 \ldots \forall x_r\ \exists! y\ \ \varphi_f(x_1, \ldots, , x_r, y). \end{array}$$

hence,

$$\begin{array}{lll} \mathscr{T}_{i+1} & \ni & \forall x_1 \ldots \forall x_r\ \forall y\ \big(\varphi_f(x_1, \ldots, x_r, y) \longleftrightarrow f(x_1, \ldots, x_r) = y\big) \\ \mathscr{T}_i & \vdash_c & \forall x_1 \ldots \forall x_r\ \exists! y\ \ \varphi_f(x_1, \ldots, , x_r, y). \end{array}$$

Since for all integers integer $1 \leqslant j \leqslant r$ the inequality $ht\,(u_j) < ht\,(t)$ holds, by induction hypothesis we have $\mathcal{L}_i$-formulas

$$\varphi_{u_1}(x_1, \ldots, x_k, y), \ldots, \varphi_{u_r}(x_1, \ldots, x_k, y),$$

such that for each integer $1 \leqslant j \leqslant r$

$$\begin{array}{lll} \mathscr{T}_{i+1} & \ni & \forall x_1 \ldots \forall x_k \ \forall y \ \left(\varphi_{u_j}(x_1, \ldots, x_k, y)) \longleftrightarrow u_1(x_1, \ldots, x_k) = y\right) \\ \mathscr{T}_i & \vdash_c & \forall x_1 \ldots \forall x_k \ \exists! y \ \varphi_{u_j}(x_1, \ldots, , x_k, y) \end{array}$$

and $\mathcal{L}$-formulas

$$\psi_{u_1}(x_1, \ldots, x_k, y), \ldots, \psi_{u_r}(x_1, \ldots, x_k, y)$$

such that for each integer $1 \leqslant j \leqslant r$

$$\begin{array}{lll} \mathscr{T}_{i+1} & \vdash_c & \forall x_1 \ldots \forall x_k \ \forall y \ \left(\varphi_{u_j}(x_1, \ldots, x_k, y) \longleftrightarrow \psi_{u_j}(x_1, \ldots, x_k, y)\right) \\ \mathscr{T}_{i+1} & \vdash_c & \forall x_1 \ldots \forall x_k \ \forall y \ \left(\varphi_{v_j}(x_1, \ldots, x_k, y) \longleftrightarrow \psi_{v_j}(x_1, \ldots, x_k, y).\right) \end{array}$$

Finally, since φ_f is a $\mathcal{L}_{i'}$-formula, we obtain by induction hypothesis some $\mathcal{L}$-formula $\psi_f(x_1, \ldots, x_r, y)$ that satisfies

$$\mathscr{T}_i \ \vdash_c \ \forall x_1 \ldots \forall x_r \ \forall y \ \left(\varphi_f(x_1, \ldots, x_r, y) \longleftrightarrow \psi_f(x_1, \ldots, x_r, y)\right).$$

By setting $z = x_l$ (for some $1 \leqslant l \leqslant k$) we obtain

$$\mathscr{T}_{i+1} \ \vdash_c \ \forall x_1 \ldots \forall x_k \left[\begin{array}{c} f(u_1, \ldots, u_r) = x_l \\ \longleftrightarrow \\ \forall y_{u_1} \ldots \forall y_{u_r} \\ \left[\left[\begin{array}{c} \bigwedge_{1 \leqslant j \leqslant r} \psi_{u_j}(x_1, \ldots, x_r, y_{u_j}) \\ \wedge \\ \bigwedge_{1 \leqslant j \leqslant r} \psi_{v_j}(x_1, \ldots, x_r, y_{v_j}) \end{array} \right] \longrightarrow \psi_f(y_{u_1}, \ldots, y_{u_r}, x_l) \right] \end{array} \right].$$

if $\boldsymbol{ht}\left(\boldsymbol{\psi'}\right) > \mathbf{0}$ then

if $\boldsymbol{\psi' := \neg\theta}$ then by induction hypothesis there exists some $\mathcal{L}$-formula $\psi_{\theta(x_1,\ldots,x_k)}$ whose free variables are among $x_1, \ldots, x_k$ such that

$$\mathscr{T}_{i+1} \vdash_c \forall x_1 \ldots \forall x_k \ \left(\theta(x_1, \ldots, x_k) \longleftrightarrow \psi_{\theta(x_1,\ldots,x_k)}\right)$$

which immediately leads to

$$\mathscr{T}_{i+1} \vdash_c \forall x_1 \ldots \forall x_k \ \left(\psi'_{(x_1,\ldots,x_k)} \longleftrightarrow \neg\psi_{\theta(x_1,\ldots,x_k)}\right).$$

if $\boldsymbol{\psi' := \left(\theta_1 \ * \ \theta_2\right)}$ where $* \in \{\wedge, \vee, \longrightarrow, \longleftrightarrow\}$ then by induction hypothesis there exists $\mathcal{L}$-formulas $\psi_{\theta_1(x_1,\ldots,x_k)}$ and $\psi_{\theta_2(x_1,\ldots,x_k)}$ whose free variables are among $x_1, \ldots, x_k$ such that

$$\mathscr{T}_{i+1} \vdash_c \forall x_1 \ldots \forall x_k \ \left(\theta_1(x_1, \ldots, x_k) \longleftrightarrow \psi_{\theta_1(x_1,\ldots,x_k)}\right)$$

$$\mathscr{T}_{i+1} \vdash_c \forall x_1 \ldots \forall x_k \ \left(\theta_2(x_1, \ldots, x_k) \longleftrightarrow \psi_{\theta_2(x_1,\ldots,x_k)}\right)$$

which immediately leads to

$$\mathscr{T}_{i+1} \vdash_c \forall x_1 \ldots \forall x_k \ \left(\psi' \longleftrightarrow [\psi_{\theta_1} \ * \ \psi_{\theta_2}]_{(x_1,\ldots,x_k)}\right).$$

if $\boldsymbol{\psi' := \exists y\, \theta(y, x_1, \ldots, x_k)}$ then by induction hypothesis there exists some $\mathcal{L}$-formula $\psi_\theta(y, x_1, \ldots, x_k)$ whose free variables are among $y, x_1, \ldots, x_k$ such that

$$\mathscr{T}_{i+1} \vdash_c \forall x_1 \ldots \forall x_k \quad \big(\theta(y, x_1, \ldots, x_k) \longleftrightarrow \psi_\theta(y, x_1, \ldots, x_k)\big)$$

which immediately leads to

$$\mathscr{T}_{i+1} \vdash_c \forall x_1 \ldots \forall x_k \quad \big(\psi' \longleftrightarrow [\exists y\, \psi_\theta]\,(x_1, \ldots, x_k)\big).$$

□ 66

3.2 Conservative Extensions

Definition 67. *Let $\mathcal{L} \subseteq \mathcal{L}'$ be any first order languages $\mathscr{T}$ be any $\mathcal{L}$-theory and $\mathscr{T}'$ any $\mathcal{L}'$-theory such that $\mathscr{T} \subseteq \mathscr{T}'$.*
$\mathscr{T}'$ is a conservative extension *of $\mathscr{T}$ if for every closed $\mathcal{L}$-formula φ*

$$\mathscr{T}' \vdash_c \varphi \iff \mathscr{T} \vdash_c \varphi.$$

Theorem 68. *Let $\mathcal{L} \subseteq \mathcal{L}'$ be any first order languages $\mathscr{T}$ be any $\mathcal{L}$-theory and $\mathscr{T}'$ any $\mathcal{L}'$-theory such that $\mathscr{T} \subseteq \mathscr{T}'$.*

If $\mathscr{T}'$ is an extension by definition *of $\mathscr{T}$, then $\mathscr{T}'$ is a* conservative extension *of $\mathscr{T}$.*

Proof of Theorem 68:

($\Leftarrow$) is trivial.

($\Rightarrow$) A direct syntactic proof consists in for every $\mathcal{L}$-formula φ, translating every proof of $\mathscr{T}' \vdash_c \varphi$ into some proof of $\mathscr{T} \vdash_c \varphi$. This is extremely tedious.

We give a proof that relies on the completeness theorem for first order logic. In order to prove $\mathscr{T}' \vdash_c \varphi \Rightarrow \mathscr{T} \vdash_c \varphi$, we proceed by contradiction and show that $\mathscr{T} \nvdash_c \varphi \Rightarrow \mathscr{T}' \nvdash_c \varphi$ which comes down to showing $\mathscr{T} \not\models \varphi \Rightarrow \mathscr{T}' \not\models \varphi$. For this purpose, we consider some $\mathcal{L}$-structure $\mathcal{M}$ such that $\mathcal{M} \models \mathscr{T}$ but $\mathcal{M} \not\models \varphi$ and we simply extend this model to some $\mathcal{L}'$-structure $\mathcal{M}'$. In other words, $\mathcal{M}'$ restricted to the language $\mathcal{L}$ (by dropping the interpretation of the extraneous material) is nothing but $\mathcal{M}$. i.e.,

- $|\mathcal{M}'| = |\mathcal{M}|$ and
- for each relation symbol $R \in \mathcal{L}$, $R^{\mathcal{M}} = R^{\mathcal{M}'}$
- for each function symbol $f \in \mathcal{L}$, $f^{\mathcal{M}} = f^{\mathcal{M}'}$.

So we assume

- $\mathcal{L} = \mathcal{L}_0 \subsetneq \mathcal{L}_1 \subsetneq \ldots \subsetneq \mathcal{L}_{i+1} = \mathcal{L}'$
- $\mathscr{T} = \mathscr{T}_0 \subsetneq \mathscr{T}_1 \subsetneq \ldots \subsetneq \mathscr{T}_{i+1} = \mathscr{T}'$

The proof is by induction on the number of steps that lead from $\mathcal{L}$ to $\mathcal{L}'$. So we assume $\mathcal{M}_i$ is an extension of $\mathcal{M}$ from $\mathcal{L} = \mathcal{L}_0$ to $\mathcal{L}_i$.

if $\mathcal{L}_{i+1} \smallsetminus \mathcal{L}_i = \{R\}$ then there exists some $\mathcal{L}_i$-formula $\varphi_R(x_1, \ldots, x_{k_i})$ whose free variables are among $x_1, \ldots, x_{k_i}$, such that

$$\mathscr{T}_{i+1} = \mathscr{T}_i \cup \left\{\forall x_1 \ldots \forall x_{k_i} \;\; (\varphi_R(x_1, \ldots, x_{k_i}) \longleftrightarrow R(x_1, \ldots, x_{k_i}))\right\}.$$

Then $\mathcal{M}_i$ extends to $\mathcal{M}_{i+1}$ by setting:

- for every symbol from $\mathcal{L}_i$, the interpretation is unchanged and
- $R^{\mathcal{M}_{i+1}} = \{(a_1, \ldots, a_{k_i}) \mid \mathcal{M}_i \models \varphi_R\,[a_1/x_1, \ldots, a_{k_i}/x_{k_i}]\}$

if $\mathcal{L}_{i+1} \smallsetminus \mathcal{L}_i = \{f\}$ then there exists some $\mathcal{L}_i$-formula $\varphi_{(x_1,\ldots,x_{k_i}}, y)$ whose free variables are among $x_1, \ldots, x_{k_i}, y$, such that

$$\mathscr{T}_{i+1} = \mathscr{T}_i \cup \left\{\forall x_1 \ldots \forall x_{k_i}\ \forall y \;\; (\varphi_f(x_1, \ldots, x_{k_i}, y) \longleftrightarrow f(x_1, \ldots, x_{k_i}) = y)\right\}$$

and

$$\mathscr{T}_i \vdash_c \forall x_1 \ldots \forall x_{k_i}\ \exists! y \;\; \varphi_f(x_1, \ldots, x_{k_i}, y).$$

Then $\mathcal{M}_i$ extends to $\mathcal{M}_{i+1}$ by setting:

- for every symbol from $\mathcal{L}_i$, its interpretation is unchanged and
- since $\mathcal{M}_i \models \mathscr{T}_i$ and $\mathscr{T}_i \vdash_c \forall x_1 \ldots \forall x_{k_i}\ \exists! y \;\; \varphi_f(x_1, \ldots, x_{k_i}, y)$ both hold, it follows that
$$\mathcal{M}_i \models \forall x_1 \ldots \forall x_{k_i}\ \exists! y \;\; \varphi_f(x_1, \ldots, x_{k_i}, y).$$
we set for each $(a_1, \ldots, a_{k_i}) \in |\mathcal{M}_{i+1}|^{k_i}$
$$f^{\mathcal{M}_{i+1}}(a_1, \ldots, a_{k_i}) = \text{ the unique } b \text{ s.t. } \varphi_f\,[a_1/x_1, \ldots, a_{k_i}/x_{k_i}, b/y]\,.$$

$\square$ 68

Chapter 4

Axiom of Choice and Cardinals

4.1 Axiom of Choice, Zorn's Lemma and Zermelo's Theorem

The **Axiom of Choice** (**AC**) asserts that given any family $\mathcal{F}$ [1] of non-empty disjoint sets there exists a set that contains exactly one element from each element of the family.

$$\forall x \left(\left(\begin{array}{c} \forall y \, (y \in x \longrightarrow \exists z \, z \in y) \\ \bigwedge \\ \forall y \forall y' \left(\left(\begin{array}{c} y \in x \\ \wedge \\ y' \in x \\ \wedge \\ \exists z \, (z \in y \ \wedge \ z \in y') \end{array} \right) \longrightarrow y = y' \right) \end{array} \right) \longrightarrow \exists c \forall y \Big(y \in x \longrightarrow \exists ! z \, (z \in y \ \wedge \ z \in c) \Big) \right).$$

By **Comprehension Schema** and **Extensionality**, one easily obtains a set c formed of exactly one element from each element of $\mathcal{F}$ and no other.

It is also very easy to see, that given any family of the form $\mathcal{F} = \{A_i \mid i \in I\}$ — also denoted $\mathcal{F} = (A_i)_{i \in I}$ — where each set A_i is non-empty, there exists a choice function

$$f : I \to \bigcup_{i \in I} A_i$$

i.e., a mapping that satisfies $f(i) \in A_i$ for every $i \in I$.

Lemma 69 (**ZF**).

$$\mathbf{AC}$$

$$\Longleftrightarrow$$

[1] a family is nothing but a set. The notation $\mathcal{F} = (A_i)_{i \in I}$ stands for $\mathcal{F}$ is the range of some functional — or even function — whose domain is the set I.

$$\forall X \left(\forall x \in X \; x \neq \varnothing \longrightarrow \exists f : X \to \bigcup X \quad \forall x \in X \; f(x) \in x\right).$$

Proof of Lemma 69:

(AC $\Rightarrow \exists \mathbf{f} : \mathbf{X} \to \bigcup \mathbf{X}$ a choice function) if $X = \varnothing$, then the empty function[2] works. If X is a non-empty set whose elements are also non-empty, then we form

$$Y = \{\{x\} \times x \mid x \in X\}$$

which is obtained by defining a functional $\mathbf{F} : \mathbf{V} \to \mathbf{V}$ by $\mathbf{F}(x) = \{x\} \times x$ and setting $Y = \mathbf{F}[X]$.

If X is a non-empty set whose elements are non-empty, so is also Y, but in addition, any two different elements of Y are disjoint. So, by **AC**, there exists some set C that contains exactly one element from each set $\{x\} \times x$ that belongs to Y. Define $f : X \to \bigcup X$ by

$$f(x) = \pi_1\big(\mathbf{F}(x)\big)$$

where $\pi_1 : \bigcup Y \to X$ is the second projection[3].

($\exists \mathbf{f} : \mathbf{X} \to \bigcup \mathbf{X}$ a choice function $\Rightarrow$ AC) is immediate.

□ 69

Definition 70 (Partial Order)**.** *$(X, \leqslant)$ is a partial order if*

- *X is a non-empty set and*
- *$\leqslant \; \subseteq X \times X$ is a partial ordering on X, i.e.,*

reflexivity

$$\forall x \in X \quad x \leqslant x$$

anti-symmetry

$$\forall x \in X \; \forall y \in X \quad \big((x \leqslant y \;\wedge\; y \leqslant x) \longrightarrow \; x = y\big)$$

transitivity

$$\forall x \in X \; \forall y \in X \; \forall z \in X \quad \big((x \leqslant y \;\wedge\; y \leqslant z) \longrightarrow x \leqslant z\big).$$

[2] The empty function is $\varnothing$.
[3] $\pi_1 : \bigcup Y \to X$ is defined by $\pi_1(a_0, a_1) = a_1$.

Definition 71 (Chain)**.** *If $(X, \leqslant)$ is a partial order, then $C \subseteq X$ is a chain if $C \neq \varnothing$ and C is totally ordered by $\leqslant$, i.e.,*

$$\forall x \in C \ \forall y \in C \quad \big(x \leqslant y \ \vee \ y \leqslant x\big).$$

Definition 72 (Antichain)**.** *If $(X, \leqslant)$ is a partial order, then $A \subseteq X$ is an antichain if any two different elements of A are $\leqslant$-incomparable, i.e.,*

$$\forall x \in C \ \forall y \in C \quad \big((x \leqslant y \ \vee \ y \leqslant x) \longrightarrow x = y\big).$$

Definition 73 (Inductive partial order)**.** *Let $(X, \leqslant)$ be any partial order. $(X, \leqslant)$ is* inductive *if every chain $C \subseteq X$ admits an upper bound in X, i.e.,*

$$\exists x \in X \ \forall y \in C \quad y \leqslant x.$$

Kuratowski-Zorn's Lemma. *Every non-empty inductive partially ordered set $(X, \leqslant)$ admits some maximal element, i.e.,*

$$\exists x \in X \, \forall y \in X \ \ (x \leqslant y \longrightarrow x = y).$$

Zermelo's Theorem. *Every set can be well-ordered. i.e.,*

$$\forall A \ \exists <_A \subseteq A \times A \ \ (A, <_A) \text{ is a well-ordering.}$$

Theorem 76 (**ZF**)**.** *The following are equivalent:*

(1) Axiom of Choice *(2) Zorn's Lemma* *(3) Zermelo's Theorem*

Proof of Theorem 76**:**

(AC ⇒ ZL) Let $(X, \leqslant)$ be some non-empty inductive partial order. Towards a contradiction, we assume that $(X, \leqslant)$ does not admit any maximal element. We consider

$$\mathscr{C} = \{C \in \mathscr{P}(X) \mid C \textit{ is a chain}\}.$$

For each $C \in \mathscr{C}$, we set

$$M_C = \{x \in X \mid \forall y \in C \;\; y \lneqq x\}.$$

Notice that $M_C \neq \varnothing$ holds for otherwise any upper bound c of C would be a maximal element in X since for any element $x \in X$, if $c \lneqq x$ holds, then $x \in M_C$ holds as well. Hence, for every element $x \in X$, one has either x is incomparable with c or $x \leqslant c$. By **AC** there exists a choice function $u : \mathscr{C} \to X$ that associates to each $C \in \mathscr{C}$ an element $u_C \in M_C$, i.e., a strict upper bound on C. By transfinite recursion we define a functional $\mathbf{F} : \mathbf{On} \to X$ by

- $\mathbf{F}(0)$ is any element that belongs to X;
- $\mathbf{F}(\alpha + 1) = u_{\{\mathbf{F}(\xi) \mid \xi \leqslant \alpha\}}$;
- $\mathbf{F}(\alpha) = u_{\{\mathbf{F}(\xi) \mid \xi < \alpha\}}$ when α is limit.

By construction, $\mathbf{F}$ is injective, so we can also define a functional $\mathbf{G} : X \to \mathbf{On}$ by

- $\mathbf{G}(x) = \alpha$ if $\mathbf{F}(\alpha) = x$ holds for some α, and
- $\mathbf{G}(x) = \mathbf{F}(0)$ otherwise.

Since X is a set, it follows by the **Comprehension Schema** that **On** is a set as well which contradicts the fact **On** is a proper class (cf. Theorem 36).

(ZL ⇒ ZT) Let X be any non-empty set. We consider the set of well-orderings of subsets of X. Namely,

$$W = \{(Y, <_Y) \mid Y \subseteq X \text{ and } <_Y \subseteq Y \times Y \text{ is a well-ordering on } Y\}.$$

Notice that W is non-empty since for any $x \in X$, one has $(\{x\}, \varnothing)$ is a well-ordering. Then, we define the following partial ordering on W:

$$(Y, <_Y) \sqsubseteq (Y', <_{Y'}) \iff \begin{cases} Y \subseteq Y' \\ \textit{and} \\ \forall a \in y \, \forall b \in y \, (a <_Y b \longleftrightarrow a <_{Y'} b) \\ \textit{and} \\ \forall a \in y \, \forall b \in \left(Y' \smallsetminus Y\right) \;\; a <_{Y'} b. \end{cases}$$

$(W, \sqsubseteq)$ is an inductive partial order since for every chain $(Y_i, <_{Y_i})_{i \in I}$, the set

$$Y = \bigcup_{i \in I} Y_i$$

is well-ordered by

$$<_Y = \bigcup_{i \in I} <_{Y_i} .$$

So $(Y, <_Y)$ belongs to W and $(Y, <_Y)$ is an upper bound of the chain $(Y_i, <_{Y_i})_{i \in I}$.

Hence, by Zorn's Lemma there exists some maximal element $(Y, <_Y)$. If $Y = X$ holds, then $(X, <_Y)$ is a well-ordering. So, towards a contradiction, we assume there exists some $y \in X \smallsetminus Y$. We set

$$Y' = Y \cup \{y\} \ \textit{and} \ <_{Y'} = <_Y \cup \{(x, y) \mid x \in Y\} .$$

It follows that

$$(Y', <_{Y'}) \in W \ \textit{and} \ (Y', <_{Y'}) \sqsubseteq (Y, <_Y)$$

contradicting the maximality of $(Y, <_Y)$.

(ZT $\Rightarrow$ AC) We let X be a set of non-empty sets. By Lemma 69 it is enough to show that there exists a choice function $f : X \to \bigcup X$ such that $\forall x \in x\, f(x) \in x$ holds.

By Zermelo's Theorem, there exists some well-ordering $(\bigcup X, <_{\bigcup X})$. We define f as

$$f(x) = \textit{the} \ <_{\bigcup X} \textit{-least element in } x.$$

$\square$ 76

4.2 Cardinals

Cardinal numbers are defined as specific ordinal numbers. The definition relies on the existence or not of a bijection between some ordinal and one of its predecessors.

Notation 77. *We write*

- $A \simeq B \iff$ *there exists some* <u>*bijection*</u> $f : A \to B$*;*
- $A \lesssim B \iff$ *there exists some* <u>*injection*</u> $f : A \to B$*;*
- $A \lnsim B \iff$ *both* $A \lesssim B$ *and* $B \not\lesssim A$.

Cantor-Schröder-Bernstein Theorem (Z). *Let A and B be any sets.*

$$\mathbf{ZF} \vdash_c \big((A \lesssim B \ \wedge \ B \lesssim A) \longrightarrow A \simeq B\big).$$

Proof of the Cantor-Schröder-Bernstein Theorem: We assume that are given two injections $i : A \xrightarrow{1-1} B$ and $j : B \xrightarrow{1-1} A$.
We distinguish between $B \subseteq A$ and $B \nsubseteq A$

If B $\subseteq$ A: By recursion on the integers we build a sequence $(C_n)_{n<\omega}$:

- $C_0 = A \smallsetminus B$,
- $C_{n+1} = i[C_n]$,
- $C = \bigcup_{n\in\mathbb{N}} C_n$.

Then, we define $h : A \longrightarrow B$ as $\begin{cases} h \restriction (A \smallsetminus C) & = & Id. \\ h \restriction C & = & i \restriction C. \end{cases}$

Notice that

$\boldsymbol{ran}(\mathbf{h}) \subseteq \mathbf{B}$ holds since given any $x \in A$,

- if $x \in (A \smallsetminus C)$, then $x \notin C_0 = A \smallsetminus B$, hence $h(x) = Id(x) = x \notin B$.
- if $x \in C$, then $h(x) = i(x) \in B$.

h is injective since $h \restriction (A \smallsetminus C) = Id.$ which is injective
$h \restriction C = i \restriction C$ which is injective as well.

h is surjective since for each $y \in B$

- if $y \notin C$, then $h(y) = y$, and
- if $y \in C$, then there exists an integer n such that $y \in C_n$. Notice that $0 < n$ holds since $y \in B$ and $C_0 = A \smallsetminus B$. Therefore, there exists $x \in C_{n-1}$ such that $i(x) = y$.

If B $\nsubseteq$ A: We consider $B' = j[B]$. Since $B' \subseteq A$ holds, there exists a bijection $h : A \longleftrightarrow B'$. Since $j : B \longleftrightarrow B'$ is bijective, it follows that $j^{-1} \circ i : A \longleftrightarrow B$ is also bijective.

□ Cantor-Schröder-Bernstein Theorem

Definition 79. *Let A be any set that* *<u>can be well-ordered</u>*.

$$|A| = card(A) \text{ is the least ordinal } \alpha \text{ such that } \alpha \simeq A.$$

By Theorem 76, since **AC** is equivalent to **ZT**, it follows that under **AC**, $|A|$ is defined for every set A.
Notice that for every ordinal α the cardinal of α is defined without any mention of **AC**.
Under **AC**, given any two sets A, B, one has $A \simeq B \iff |A| = |B|$. In other words, the functional $|\ | : \mathbf{V} \to \mathbf{On}$ that maps a set A to its cardinal $|A|$ assigns one particular ordinal to each equivalent class determined by $\simeq$.
Moreover, $|\alpha| \leqslant \alpha$ always holds which justifies the following definition.

Definition 80. *Let $\alpha \in \mathbf{On}$.*

$$\alpha \textit{ is a cardinal} \iff |\alpha| = \alpha.$$

Notice that for every ordinal α,

$$|\alpha| = \alpha \iff \forall \beta < \alpha \quad \beta \not\simeq \alpha.$$

Lemma 81 (**ZF**)**.** *Let $\alpha, \beta \in \mathbf{On}$ and $n \in \omega$,*

(1) $|\alpha| \leqslant \beta \leqslant \alpha \longrightarrow |\alpha| = |\beta|$

(2) $n \not\simeq n + 1$

(3) $|n| = n$

(4) $|\omega| = \omega$

Proof of Lemma 81:

(1) Since $|\alpha| \leqslant \beta \leqslant \alpha$ holds, one has $|\alpha| \lesssim \beta \lesssim \alpha \lesssim |\alpha|$. Hence, by Cantor-Schröder-Bernstein Theorem[4] one has $|\alpha| \simeq \beta \simeq \alpha$.

(2) The proof goes by induction on n.

$\mathbf{n = 0}$ there is clearly no bijection $f : \{\varnothing\} \to \varnothing$.

$\mathbf{n = k + 1}$ towards a contradiction, assume there exists $f : k + 2 \overset{bij.}{\longleftrightarrow} k + 1$.

- if $f(k + 1) = k$, then $f \restriction k + 1 : k + 1 \to k$ is a bijection, which contradicts the induction hypothesis.
- if $f(k + 1) = p \neq k$, then take $q = f^{-1}(k)$ and define $g : k + 2 \longrightarrow k + 1$ by
  - $g(k + 1) = k$
  - $g(q) = p$
  - $g(r) = f(r)$ for $r \notin \{p, q\}$.

  Clearly, one has $g : k + 2 \overset{bij.}{\longleftrightarrow} k + 1$,hence $g \restriction k + 1 : k + 1 \overset{bij.}{\longleftrightarrow} k$ holds which contradicts the induction hypothesis.

(3) By Lemma 81 (2), there are no ordinals[5] $k < n < \omega$ such that $k \simeq n$. Thus $|n| = n$ holds.

(4) $|\omega| = \omega$ holds since for every integer n, $n \simeq \omega$ would yield $n \simeq n + 1$, contradicting Lemma 81 (2).

[4]This is Cantor-Schröder-Bernstein Theorem (page 55).
[5]Otherwise one would have $k \lesssim k + 1 \lesssim n \lesssim k$, hence $k \simeq k + 1 \simeq n$.

$\square$ 81

Definition 82. *Let A be any set that can be well-ordered.*

- *A is finite if $|A| < \omega$*
- *A is countable if $|A| \leqslant \omega$*
- *A is infinite if $|A| \nless \omega$*
- *A is uncountable if $|A| \nleqslant \omega$.*

Lemma 83 (**ZF**).

If A is a set of cardinal numbers, then $\sup A = \bigcup A$ is also a cardinal.

Proof of Lemma 83: By Lemma 43 (2), $\sup_{\alpha \in A} \alpha = \beta$ is an ordinal.

If $\boldsymbol{\beta} \in \mathbf{A}$, then β is a cardinal.

If $\boldsymbol{\beta} \notin \mathbf{A}$, then A is unbounded. i.e., for each ordinal $\gamma < \beta$, there exists some $|\alpha| \in A$ such that $\gamma < |\alpha|$. If we assume $|\beta| < \beta$, we obtain $|\beta| < |\alpha| < \beta$. Lemma 81(1) yields the contradiction $|\beta| = |\alpha|$.

$\square$ 83

Lemma 84 (**ZF**). *Let $\alpha \in$ **On**.*

If $\omega < \alpha$ and $|\alpha| = \alpha$, then α is a limit ordinal.

Proof of Lemma 84: If $\omega < \alpha$ holds, the following mapping $f : \alpha + 1 \overset{bij.}{\longleftrightarrow} \alpha$:

$$\begin{cases} f(\alpha) & = & 0 \\ f(\gamma) & = & 1 + \gamma \quad (\textit{any } \gamma < \alpha) \end{cases}$$

shows that infinite cardinals are limit ordinals.

$\square$ 84

For the moment, the only infinite cardinal we encountered is ω. The **Power Set** Axiom will provide us with plenty of infinite cardinals.

4.3 The Power Set Axiom

The **Power Set** Axiom claims that given any set x there exists some set y that contains all subsets of x:

$$\forall x\, \exists y\, \forall z \left(\forall u \ (u \in z \rightarrow u \in x) \rightarrow z \in y\right).$$

Again, by mean of an instance of the **Comprehension Schema** and **Extensionality**, one obtains the existence of a certain functional $\mathscr{P} : \mathbf{V} \rightarrow \mathbf{V}$.

Definition 85. *Let A be any set.*

$$\mathscr{P}(A) = \{B \mid B \subseteq A\}.$$

Cantor's Theorem (ZF). *Let A be any set.*

$$A \lnsim \mathscr{P}(A).$$

Proof of Cantor's Theorem: The mapping $g : A \rightarrow \mathscr{P}(A)$ defined by $g(x) = \{x\}$ is injective, hence $A \lesssim \mathscr{P}(A)$ holds. Towards a contradiction, we assume $\mathscr{P}(A) \lesssim A$ and also by Cantor-Schröder-Bernstein Theorem (page 55) $\mathscr{P}(A) \simeq A$. We let $f : A \overset{bij.}{\longleftrightarrow} \mathscr{P}(A)$ and set

$$B = \{a \in A \mid a \notin f(a)\} \quad \textit{and} \quad b = f^{-1}(B).$$

This leads to the contradiction $b \in B \iff b \notin B$.

□ Cantor's Theorem

Lemma 87 (ZF).

$$\forall \alpha \in \mathbf{On}\ \exists \beta \in \mathbf{On} \ \left(\alpha < \beta \ \wedge \ |\beta| = \beta\right).$$

(This result is a variant of Hartog's Lemma — see page 273.)

Proof of Lemma 87: If α is an integer, $\beta = \omega$ gives the result. So we assume $\omega \leqslant \alpha$ and we consider

$$\mathcal{V} = \{<_\alpha \subseteq \mathscr{P}(\alpha \times \alpha) \mid (\alpha, <_\alpha) \textit{ is a well-ordering}\}$$

and the functional $\mathbf{F} : \mathcal{V} \rightarrow \mathbf{V}$ defined by

$$\mathbf{F}(<_\alpha) = (\alpha, <_\alpha).$$

We set
$$\mathcal{W} = \mathbf{F}[\mathcal{V}]$$
and
$$\mathcal{O} = \textit{type}\,[\mathcal{W}]$$
i.e.,
$$\mathcal{O} = \{\textit{type}\,(\alpha, <_\alpha) \mid <_\alpha \in \mathcal{V}\}\,.$$
Finally we consider $\sup \mathcal{O}$ and show it is a cardinal.
For this, notice first that $\sup \mathcal{O} \notin \mathcal{O}$ since $\mathcal{O}$ is closed under the successor operation because
$$\forall \alpha \in \mathbf{On} \;\; \big(\omega \leqslant \alpha \longrightarrow (\alpha + 1 \simeq 1 + \alpha \,\wedge\, 1 + \alpha = \alpha)\big).$$
Finally, towards a contradiction, we assume there exists some ordinal $\gamma < \sup \mathcal{O}$ that satisfies $\sup \mathcal{O} \lesssim \gamma$. We take any $\delta \in \mathcal{O}$ that satisfies $\gamma \leqslant \delta$. This leads to
$$\sup \mathcal{O} \lesssim \gamma \lesssim \delta \lesssim \sup \mathcal{O}$$
which comes down to
$$\gamma \simeq \delta \simeq \sup \mathcal{O}.$$
Since $\delta \in \mathcal{O}$ holds, it follows
$$\alpha \simeq \delta \simeq \sup \mathcal{O}.$$
Notice then that given any $f : \alpha \overset{bij.}{\longleftrightarrow} \sup \mathcal{O}$ the relation
$$<_f = \{(\gamma, \delta) \in \alpha \times \alpha \mid f(\gamma) < f(\delta)\}$$
satisfies
$$(\alpha, <_f) \;\; \textit{is a well-ordering}$$
and
$$\textit{type}\,(\alpha, <_f) = \sup \mathcal{O}.$$
which yields $\sup \mathcal{O} \in \mathcal{O}$, a contradiction.

□ 87

Lemma 87 motivates the following definitions.

Definition 88. *Given any $\alpha \in \mathbf{On}$, α^+ stands for the least cardinal $> \alpha$.*

Notice that for any $\alpha \in \mathbf{On}$, one has $\alpha^+ = |\alpha|^+$.

From now on, we use the letters κ and λ to denote cardinal numbers.

Definition 89. *Let κ be any cardinal.*

- κ *is a* successor *cardinal* $\iff \exists \alpha \in \mathbf{On} \ \ \kappa = \alpha^+$.
- κ *is a* limit *cardinal* $\iff \big(\omega < \kappa \wedge \forall \alpha \in \mathbf{On} \ \ \kappa \neq \alpha^+\big)$.

Definition 90. *By transfinite recursion, one defines the functional* $\aleph : \mathbf{On} \to \mathbf{On}$ *by:*

- $\aleph_0 = \omega$
- $\aleph_{\alpha+1} = \aleph_\alpha^+$
- $\aleph_\alpha = \sup \{\aleph_\beta \mid \beta < \alpha\}$ *when α is a limit ordinal.*

Notation 91. *One traditionally uses both notations $\aleph_\alpha$ or ω_α to name the same ordinal. The choice between ω_α and $\aleph_\alpha$ only depends on whether one wants to emphasise the cardinal side ($\aleph_\alpha$) or the ordinal side (ω_α).*

The need of such distinction will become clearer once we introduce the cardinal arithmetic operations which behave very differently from the ordinal arithmetic operations.

Lemma 92 (ZF).

(1) If $\alpha < \beta$, then $\aleph_\alpha < \aleph_\beta$.

(2) Every $\aleph_\alpha$ is an infinite cardinal.

(3) Every infinite cardinal is some $\aleph_\alpha$.

(4) $\aleph_\alpha$ is a successor cardinal $\iff$ α is a successor ordinal.

(5) $\aleph_\alpha$ is a limit cardinal $\iff$ α is a limit ordinal.

Proof of Lemma 92:

(1) This is immediate by induction on β.

(2) By induction on α:

- $\aleph_0 = \omega$ is an infinite cardinal,

- $\aleph_{\alpha+1} = \aleph_\alpha^+$ is also an infinite cardinal,
- when α is limit, $\aleph_\alpha = \sup\{\aleph_\beta \mid \beta < \alpha\}$ is also a cardinal for otherwise one would have $|\aleph_\alpha| < \aleph_\alpha$, hence there exists some $\beta < \alpha$ such that $|\aleph_\alpha| \leqslant \aleph_\beta < \aleph_\alpha$ holds. But then one would have
$$\aleph_\alpha \precsim |\aleph_\alpha| \precsim \aleph_\beta \precsim \aleph_{\beta+1} \precsim \aleph_\alpha$$
which yields the following contradiction:
$$\aleph_\beta \simeq \aleph_{\beta+1} = \aleph_\beta^+.$$

(3) Towards a contradiction, we assume there exists some infinite cardinal which is different from all $\aleph_\alpha$. Let κ be the least such cardinal.

- $\kappa = \omega$ is impossible since $\omega = \aleph_0$.
- $\kappa = \lambda^+$ is impossible since by induction hypothesis, $\lambda = \aleph_\alpha$ holds for some ordinal α, which gives
$$\kappa = \lambda^+ = \aleph_\alpha^+ = \aleph_{\alpha+1}.$$
- If κ is a limit cardinal, then for each ordinal $\alpha < \kappa$, $\alpha^+ < \kappa$ holds. So, *a fortiori* $|\alpha|^+ < \kappa$ holds as well and one has
$$\kappa = \sup_{\alpha<\kappa} \alpha = \sup_{\alpha<\kappa} \alpha^+ = \sup_{\alpha<\kappa} |\alpha|^+$$
By induction hypothesis, for each $\alpha < \kappa$, there exists some $\theta_\alpha \in \mathbf{On}$ such that $|\alpha| = \aleph_{\theta_\alpha}$. Since $|\alpha|^+ < \kappa$ holds for every $\alpha < \kappa$, the set $\{\theta_\alpha \mid \alpha < \kappa\}$ is closed under successor. Hence $\sup\{\theta_\alpha \mid \alpha < \kappa\} = \beta$ is a limit ordinal and
$$\kappa = \sup\{|\alpha| \mid \alpha < \kappa\} = \sup\{\aleph_{\theta_\alpha} \mid \alpha < \kappa\} = \aleph_\beta;$$
a contradiction.

(4) Immediate by induction on α.

(5) Immediate by induction on α.

$\square$ 92

4.4 Cardinal Addition, Multiplication and Exponentiation

We define addition, multiplication and exponentiation on cardinals. Since cardinals are specific ordinals, these operations should not be mistaken with the ordinal addition, multiplication and exponentiation. Formally, one should use different symbols such as for example $+, \cdot$ for ordinal addition and multiplication and $\oplus, \otimes$ for cardinal addition and multiplication. Nevertheless, we will use the same symbols and make sure the context takes care of which operation is in use. For instance, α^β deals with ordinal exponentiation while κ^λ refers to the cardinal one.

Definition 93 (Cardinal Addition)**.** *Let κ, λ be cardinals.*

$$\kappa + \lambda = |\kappa \times \{0\} \cup \lambda \times \{1\}\,|$$

Notice that this is well defined — even without **AC** — for there exists at least one ordinal α such that

$$\kappa \times \{0\} \cup \lambda \times \{1\} \simeq \alpha.$$

Namely the ordinal $\kappa + \lambda$ where $+$ stands for the ordinal addition.

When κ and λ are both integers, the cardinal addition is no different than the ordinal one which is exactly the addition on integers. Clearly, as the ordinal addition, the cardinal one is associative. But, contrary to its ordinal counterpart, cardinal addition is commutative.

Example 94.

- $1 + \aleph_0 = \aleph_0$
- $\aleph_0 + 1 = \aleph_0$
- $\aleph_0 + \aleph_1 = \aleph_1$
- $\aleph_1 + \aleph_1 = \aleph_1$
- $\aleph_\omega + \aleph_1 = \aleph_\omega$
- $\aleph_{\omega_1} + \aleph_{\omega_\omega} = \aleph_{\omega_\omega}$

Definition 95 (Cardinal Multiplication)**.** *Let κ, λ be cardinals.*

$$\kappa \cdot \lambda = |\kappa \times \lambda|$$

Notice that this is well defined — even without **AC** — for there exists at least one ordinal α such that

$$\kappa \times \lambda \simeq \alpha.$$

Namely the ordinal $\kappa \cdot \lambda$ where $\cdot$ stands for the ordinal multiplication.

When κ and λ are both integers, the cardinal multiplication is no different than the ordinal one which is exactly the multiplication on integers. Clearly, as the ordinal multiplication, the cardinal one is associative. But, contrary to its ordinal counterpart, cardinal multiplication is commutative.

Example 96.

- $1 \cdot \aleph_0 = \aleph_0$
- $\aleph_0 \cdot 1 = \aleph_0$
- $\aleph_0 \cdot \aleph_1 = \aleph_1$
- $\aleph_1 \cdot \aleph_1 = \aleph_1$
- $\aleph_\omega \cdot \aleph_1 = \aleph_\omega$
- $\aleph_{\omega_1} \cdot \aleph_{\omega_\omega} = \aleph_{\omega_\omega}$

Definition 97. *Given any sets A, B,*

$$^{A}B = \{f \in \mathscr{P}(A \times B) \mid f : A \to B\}.$$

i.e.,

$$^{A}B = \{f \subseteq A \times B \mid dom(f) = A,\ ran(f) \subseteq B \text{ and } f \text{ is a function}\}.$$

We also occasionally write B^A instead of ^{A}B. Notice that $^{A}B \in \mathscr{P}(\mathscr{P}(A \times B))$.

Definition 98 (AC). *Cardinal Exponentiation. Let κ, λ be cardinals.*

$$\kappa^\lambda = |^{\lambda}\kappa|$$

Notice that this definition requires **AC** because otherwise there may not be any ordinal α such that

$$^{\lambda}\kappa \simeq \alpha.$$

For instance, take $\kappa = \lambda = \aleph_0$, then $^{\lambda}\kappa$ is the set of functions from the integers to the integers. One may think of it as the set of irrational numbers since equipped with the product topology of the discrete topology on ω, this set becomes a topological space called the *Baire space*, which is homeomorphic to the subspace of the reals (equipped with the usual topology) formed of all the irrationals. With this in mind, finding an ordinal α which is equipotent to $^{\aleph_0}\aleph_0$ is the same as finding one equipotent with the irrationals.

Notice that when κ and λ are both integers, the cardinal exponentiation is no different than the ordinal one which is exactly the exponentiation on integers.

Lemma 99 (ZF). *Let κ be any infinite cardinal,*

$$\kappa \cdot \kappa = \kappa.$$

Proof of Lemma 99: Notice first that for all $\alpha \in \mathbf{On}$,

$$|\alpha \times \alpha| = \big||\alpha| \times |\alpha|\big|.$$

Notice that given any $f : \alpha \overset{bij.}{\longleftrightarrow} |\alpha|$, the mapping $g : \alpha \times \alpha \longrightarrow |\alpha| \times |\alpha|$ defined by $g(\xi, \xi') = \big(f(\xi), f(\xi')\big)$ is 1-1 and onto. So, one has

$$|\alpha \times \alpha| \simeq \alpha \times \alpha \simeq |\alpha| \times |\alpha| \simeq \big||\alpha| \times |\alpha|\big|.$$

We now show $\kappa \cdot \kappa = \kappa$. Since obviously, $\kappa \cdot \kappa \geqslant \kappa$ holds, it only remains to show $\kappa \cdot \kappa \leqslant \kappa$. For this purpose, we proceed by induction on κ. If $\kappa = \aleph_0$, then $f : \aleph_0 \times \aleph_0 \to \aleph_0$ defined by

$$f(n, m) = \frac{(n+m)(n+m+1)}{2} + m$$

is a bijection. This shows $\aleph_0 \cdot \aleph_0 = \aleph_0$.
For $\kappa > \aleph_0$, we describe the following well-ordering on $\kappa \times \kappa$:

$$(\alpha, \beta) \ \lhd \ (\alpha', \beta') \iff \begin{cases} \max\{\alpha, \beta\} < \max\{\alpha', \beta'\} \\ \bigvee \\ \begin{pmatrix} \max\{\alpha, \beta\} = \max\{\alpha', \beta'\} \\ \bigwedge \\ (\alpha, \beta) \ <_{lexic.} \ (\alpha', \beta'). \end{pmatrix} \end{cases}$$

Clearly, given any $\alpha, \beta < \kappa$,

$$\{(\zeta, \theta) \in \kappa \times \kappa \mid (\zeta, \theta) \lhd (\alpha, \beta)\} \subseteq (\max\{\alpha, \beta\} + 1) \times (\max\{\alpha, \beta\} + 1).$$

So, one has

$$\big|\{(\zeta, \theta) \in \kappa \times \kappa \mid (\zeta, \theta) \lhd (\alpha, \beta)\}\big| \leqslant \big|(\max\{\alpha, \beta\} + 1) \times (\max\{\alpha, \beta\} + 1)\big|$$

and since both

$$\big|(\max\{\alpha, \beta\} + 1) \times (\max\{\alpha, \beta\} + 1)\big| = \Big|\big|(\max\{\alpha, \beta\} + 1)\big| \times \big|(\max\{\alpha, \beta\} + 1)\big|\Big|$$

and

$$\big|(\max\{\alpha, \beta\} + 1)\big| \leqslant \lambda$$

hold for some infinite cardinal $\lambda < \kappa$. It follows by induction hypothesis that

$$\left|\{(\zeta,\theta) \in \kappa \times \kappa \mid (\zeta,\theta) \lhd (\alpha,\beta)\}\right| \leqslant |\lambda \times \lambda| = \lambda \cdot \lambda = \lambda.$$

Therefore,

$$\left|type\,(\kappa \times \kappa, \lhd)\right| \leqslant \sup_{(\alpha,\beta)\in\kappa\times\kappa} \left|\{(\zeta,\theta) \in \kappa \times \kappa \mid (\zeta,\theta) \lhd (\alpha,\beta)\}\right|$$

which leads to

$$\left|type\,(\kappa \times \kappa, \lhd)\right| \leqslant \sup_{\lambda<\kappa} |\lambda| = \kappa.$$

□ 99

This result immediately yields a very simple way of looking at the cardinal addition and multiplication.

Theorem 100 (**ZF**). *Let λ, κ be cardinal numbers.*

If $\boldsymbol{\lambda = 0}$ or $\boldsymbol{\kappa = 0}$ *then*

- $\lambda + \kappa = \max\{\lambda, \kappa\}$
- $\lambda \cdot \kappa = 0$

If $\boldsymbol{0 < \lambda}$ and $\boldsymbol{0 < \kappa}$

If $\boldsymbol{0 \leqslant \lambda < \aleph_0}$ and $\boldsymbol{0 \leqslant \kappa < \aleph_0}$ *then addition and multiplication are no different from addition and multiplication on the integers.*

If $\boldsymbol{\aleph_0 \leqslant \lambda}$ or $\boldsymbol{\aleph_0 \leqslant \kappa}$ *then*

- $\lambda + \kappa = \max\{\lambda, \kappa\}$
- $\lambda \cdot \kappa = \max\{\lambda, \kappa\}$

Proof of Theorem 100: We only consider the case $\lambda, \kappa > 0$ with either $\aleph_0 \leqslant \lambda$ or $\aleph_0 \leqslant \kappa$. We let $\kappa = \max\{\lambda, \kappa\}$. The following yields the result:

$$\kappa \leqslant \lambda + \kappa \leqslant \lambda \cdot \kappa \leqslant \kappa \cdot \kappa = \kappa.$$

□ 100

Lemma 101 (**ZFC**). *If λ, κ are cardinals such that $\aleph_0 \leqslant \kappa$ and $2 \leqslant \lambda \leqslant \kappa$, then*

$$\lambda^\kappa = 2^\kappa = \left|\mathscr{P}(\kappa)\right|.$$

Proof of Lemma 101: One has

$$^{\kappa}2 \lesssim {}^{\kappa}\lambda \lesssim \mathscr{P}(\kappa \times \lambda) \lesssim \mathscr{P}(\kappa \times \kappa) \lesssim \mathscr{P}(\kappa) \lesssim {}^{\kappa}2.$$

□ 101

Lemma 102 (ZFC). *Let κ, λ, μ be any cardinals.*

(1) $\kappa^{\lambda+\mu} = \kappa^{\lambda} \cdot \kappa^{\mu}$

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

Proof of Lemma 102:

(1) First, notice that given any set A one has

$$\left|{}^{|A|}\kappa\right| = \left|{}^{A}\kappa\right|$$

since this obviously holds if $A = \varnothing$ and otherwise, given any $f : |A| \overset{bij.}{\longleftrightarrow} A$, one easily defines a bijection $h : \left|{}^{|A|}\kappa\right| \overset{bij.}{\longleftrightarrow} \left|{}^{A}\kappa\right|$ by setting for each $g \in {}^{|A|}\kappa$ and each $a \in A$,

$$h(g)(a) = g\big(f(a)\big).$$

Second, one has

$$\kappa^{\lambda+\mu} = \left|{}^{(\lambda+\mu)}\kappa\right| = \left|{}^{\left|\lambda\times\{0\}\cup\mu\times\{1\}\right|}\kappa\right| = \left|{}^{\lambda\times\{0\}\cup\mu\times\{1\}}\kappa\right|$$

and since $\big(\lambda \times \{0\}\big) \cap \big(\mu \times \{1\}\big) = \varnothing$, it follows that

$$^{\lambda\times\{0\}\cup\mu\times\{1\}}\kappa \simeq {}^{\lambda\times\{0\}}\kappa \times {}^{\mu\times\{1\}}\kappa \simeq {}^{\lambda}\kappa \times {}^{\mu}\kappa.$$

This leads to

$$\kappa^{\lambda+\mu} = \left|{}^{\lambda\times\{0\}\cup\mu\times\{1\}}\kappa\right| = \left|{}^{\lambda}\kappa \times {}^{\mu}\kappa\right| = \left|\left|{}^{\lambda}\kappa\right| \times \left|{}^{\mu}\kappa\right|\right| = \left|\kappa^{\lambda} \times \kappa^{\mu}\right| = \kappa^{\lambda} \cdot \kappa^{\mu}.$$

(2) One has

$$\left(\kappa^{\lambda}\right)^{\mu} = \left|{}^{\mu}\left(\kappa^{\lambda}\right)\right| = \left|{}^{\mu}\left|{}^{\lambda}\kappa\right|\right|$$

and also

$$^{\mu}\left|{}^{\lambda}\kappa\right| \simeq {}^{\mu}\left({}^{\lambda}\kappa\right) \simeq {}^{(\mu\times\lambda)}\kappa \simeq {}^{|\mu\times\lambda|}\kappa \simeq {}^{\mu\cdot\lambda}\kappa$$

which immediately gives the result.

□ 102

Lemma 103 (**ZFC**)**.** *Given any sets A and B,*

(1) if there exists $f : A \xrightarrow{1-1} B$, then $|A| \leqslant |B|$.

(2) If there exists $f : A \xrightarrow{onto} B$, then $|A| \geqslant |B|$.

Proof of Lemma 103:

(1) One has $f : A \overset{bij.}{\longleftrightarrow} ran(f) \subseteq B$, hence $|A| = |ran(f)| \leqslant |B|$.

(2) By **AC**, there exists a choice function[6] $g : B \xrightarrow{1-1} A$, which shows $|B| = |ran(g)| \leqslant |A|$.

□ 103

We extend the well-known result that says *"every countable union of countable sets is countable"* to a more general one that reads *"every union of at most κ many sets of cardinality at most κ has cardinality κ"*. Notice however, that this result requires the Axiom of Choice, or at least a weaker version: the Axiom of Countable Choice (**CC**) for the countable version of it. Indeed, if **ZF** is consistent, then in the Feferman-Lévy model the sets $\mathscr{P}(\omega)$, $\mathbb{R}$ (the set of reals) and ω_1 are all countable unions of countable sets [16, 24, 8].

Lemma 104 (**ZFC**)**.** *Let $\kappa \geqslant \aleph_0$ be any infinite cardinal and, for each $\alpha < \kappa$, A_α be any set such that $|A_\alpha| \leqslant \kappa$. One has*

$$\left|\bigcup_{\alpha \leqslant \kappa} A_\alpha\right| \leqslant \kappa.$$

Proof of Lemma 104: We consider the set

$$\left\{\left\{f \in {}^{A_\alpha}\kappa \mid f : A_\alpha \xrightarrow{1-1} \kappa\right\} \mid \alpha < \kappa\right\}$$

which is a family of non-empty sets [7]. By **AC**, we obtain a choice function

$$c : \kappa \to \bigcup \left\{\left\{f \in {}^{A_\alpha}\kappa \mid f : A_\alpha \xrightarrow{1-1} \kappa\right\} \mid \alpha < \kappa\right\}$$

[6] For instance, a mapping $g : B \xrightarrow{1-1} A$ can be obtained by considering the set $\{f^{-1}(b) \in \mathscr{P}(A) \mid b \in B\}$ which is a set of non-empty sets — in fact it is a partition of A into non-empty equivalence classes, namely the ones induced by $a \sim a' \iff f(a) = f(a')$ — and then, by mean of **AC**, picking a single element in each of these non-empty sets.

[7] Notice that even in the case where $A_\alpha = \varnothing$, one has ${}^{A_\alpha}\kappa = {}^{\varnothing}\kappa = \{\varnothing\}$. Moreover, the empty function is injective.

such that

$$c(\alpha) = f_\alpha \in \left\{ f \in {}^{A_\alpha}\kappa \mid f : A_\alpha \xrightarrow{1-1} \kappa \right\}.$$

We define $g : \bigcup_{\alpha \leqslant \kappa} A_\alpha \xrightarrow{1-1} \kappa \times \kappa$ by

$$g(a) = \big(\alpha, f_\alpha(a)\big) \textit{ where } \alpha \textit{ is the least ordinal s.t. } a \in A_\alpha.$$

From Lemma 103 (1) it follows that

$$\Big|\bigcup_{\alpha \leqslant \kappa} A_\alpha\Big| \leqslant |\kappa \times \kappa| = \kappa \cdot \kappa = \kappa.$$

□ 104

4.5 Cofinality

Definition 105. *Given any $\alpha \in \mathbf{On}$ and $A \subseteq \alpha$,*

$$A \textit{ is unbounded in } \alpha \iff \forall \gamma \in \alpha\ \exists \delta \in A \quad \delta \geqslant \gamma.$$

Remark 106. Depending on whether α is a successor ordinal or a limit ordinal, there is a simple characterisation of "being unbounded in α":

- If $\alpha = 0$, then any $A =\subseteq \alpha \implies A = \alpha = \varnothing$; hence $A = \varnothing$ is unbounded.
- If $\alpha = \beta + 1$, then $A \subseteq \alpha$ is unbounded in $\alpha \iff \beta \in A$.
- If α is a limit ordinal, then $A \subseteq \alpha$ is unbounded in $\alpha \iff \sup A = \alpha$.

Definition 107. *Given any $\alpha, \beta \in \mathbf{On}$,*

$$f : \beta \to \alpha \textit{ is cofinal if } ran(f) \textit{ is unbounded in } \alpha.$$

Notation 108. *Given any $\alpha, \beta \in \mathbf{On}$,*

$$f : \beta \xrightarrow{cof.} \alpha \textit{ stands for } f : \beta \to \alpha \textit{ is cofinal.}$$

Remarks 109.

(1) If $f : \gamma \xleftrightarrow{bij.} \beta$ and $g : \beta \xrightarrow{cof.} \alpha$, then $g \circ f : \gamma \xrightarrow{cof.} \alpha$.

(2) In case α is a successor ordinal ($\alpha = \gamma + 1$):

$$f : \beta \to \alpha \text{ is cofinal } \iff \gamma \in f[\beta].$$

And in case α is a limit ordinal:

$$f : \beta \to \alpha \text{ is cofinal } \iff \sup f[\beta] = \alpha.$$

Definition 110. *Given any $\alpha \in \mathbf{On}$,*

$$\mathit{cof}(\alpha) \textit{ is the least } \beta \textit{ s.t. there exists some } f : \beta \xrightarrow{cof.} \alpha.$$

Remarks 111.

(1) One has $\mathit{cof}(\alpha) \leqslant \alpha$ and by Remark 109 (2):

- $\mathit{cof}(\alpha) = 1$ in case α is a successor ordinal, and
- $\mathit{cof}(\alpha)$ is a limit ordinal in case α is a limit ordinal.

(2) By Remark 109 (1) $\mathit{cof}(\alpha)$ is a cardinal.

Lemma 112. *Given any $\alpha \in \mathbf{On}$,*

$$\textit{there exists } f : \mathit{cof}(\alpha) \xrightarrow{cof.} \alpha \textit{ s.t. } \forall \xi, \xi' \in \mathit{cof}(\alpha) \quad \xi < \xi' \longrightarrow f(\xi) < f(\xi').$$

Proof of Lemma 112: If α is a successor ordinal, then $\alpha = \delta + 1$, $\mathit{cof}(\alpha) = 1$ and $g : 1 \to \alpha$ defined by $f(1) = \delta$ satisfies the requirement.

If α is a limit ordinal, from any $g : \mathit{cof}(\alpha) \xrightarrow{cof.} \alpha$ we define $f : \mathit{cof}(\alpha) \xrightarrow{cof.} \alpha$ strictly increasing by transfinite recursion:

- $f(0) = g(0)$
- $f(\xi+1) = \max\{g(\xi+1), f(\xi)+1\}$ $\big(f(\xi)+1 \in \alpha$ since α is limit$\big)$
- $f(\delta) = \sup\big(\{g(\delta)\} \cup \{f(\xi) \mid \xi < \delta\}\big)$ for δ limit $\big(f(\delta) \in \alpha$ since $f(\delta) \notin \alpha \to cof(\alpha) \leqslant \delta\big)$.

□ 112

Lemma 113. *If there exist $\omega \leqslant \alpha, \beta \in$ **On**, and $f : \beta \xrightarrow{cof.} \alpha$ strictly increasing, then*

$$cof(\alpha) = cof(\beta)\,.$$

Proof of Lemma 113:

***cof*($\boldsymbol{\alpha}$) $\leqslant$ *cof*($\boldsymbol{\beta}$):** By Lemma 112 there exists some strictly increasing $g : cof(\beta) \to \beta$. Then the mapping $g \circ f : cof(\beta) \to \alpha$ is cofinal which shows $cof(\alpha) \leqslant cof(\beta)$.

***cof*($\boldsymbol{\beta}$) $\leqslant$ *cof*($\boldsymbol{\alpha}$):** By Lemma 112 there exists some strictly increasing $g : cof(\alpha) \to \alpha$. We define $h : cof(\alpha) \xrightarrow{cof.} \beta$ by

$$h(\xi) = \textit{least } \zeta \textit{ s.t. } f(\zeta) \geqslant g(\xi).$$

which shows $cof(\beta) \leqslant cof(\alpha)$.

□ 113

Corollary 114. *Given any $\alpha \in$ **On**,*

(1) $cof(cof(\alpha)) = cof(\alpha)$.

(2) If α is a limit ordinal, then $cof(\aleph_\alpha) = cof(\alpha)$.

Proof of Corollary 114: Immediate.

□ 114

Definition 115. *Given any infinite cardinal κ,*

- *κ is regular if $cof(\kappa) = \kappa$,*
- *κ is singular if $cof(\kappa) < \kappa$.*

Example 116.

- $\aleph_0$ *is regular since* $cof(\aleph_0) = \omega$
- $\aleph_\omega$ *is singular since* $cof(\aleph_\omega) = cof(\omega) = \omega$
- $\aleph_{\omega_1}$ *is singular since* $cof(\aleph_{\omega_1}) = cof(\omega_1) \leqslant \omega_1 < \aleph_{\omega_1}$
- $\aleph_{\varepsilon_0}$ *where* $\varepsilon_0 = \sup_{n<\omega} \underbrace{\omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega_0}}}}}}_{n}$ *is singular since* $cof(\aleph_{\varepsilon_0}) = cof(\varepsilon_0) = \omega$
- $\sup_{n<\omega} \underbrace{\aleph_{\aleph_{\aleph_{\cdot_{\cdot_{\cdot_{\aleph_0}}}}}}}_{n}$ *is singular since* $cof\left(\sup_{n<\omega} \underbrace{\aleph_{\aleph_{\aleph_{\cdot_{\cdot_{\cdot_{\aleph_0}}}}}}}_{n}\right) = \omega$

It seems from these examples that every infinite limit cardinal we come up with is singular. A question arises: "can we find an infinite limit cardinal that is regular?" The answer to this question — as very often in Set Theory — is "it depends". Indeed, this notion of an infinite limit regular cardinal will be the first one of a so-called "large cardinal" that we will encounter. For the moment, we notice that if we work with **AC**, then every infinite successor cardinal is regular.

Lemma 117 (**ZFC**)**.** *Let κ be any infinite cardinal.*

$$\kappa^+ \text{ is regular.}$$

Proof of Lemma 117: Towards a contradiction, we assume that $cof(\kappa^+) < \kappa^+$, so that $cof(\kappa^+) \leqslant \kappa$ holds and there exists some mapping $f : \kappa \xrightarrow{cof.} \kappa^+$. It follows

$$\kappa^+ = \sup f[\kappa] = \bigcup_{\xi<\kappa} f(\xi).$$

Since for each $\xi < \kappa$, $f(\xi) < \kappa^+$ holds, one has $|f(\xi)| \leqslant \kappa$ which — by Lemma 104 — leads to $\left|\bigcup_{\xi<\kappa} f(\xi)\right| = \kappa$, i.e., $|\kappa^+| = \kappa$, a contradiction.

$\square$ 117

König's Lemma (ZFC). *Let κ be any infinite cardinal.*

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

Proof of König's Lemma: We show that there is no $g : \kappa \xleftrightarrow{bij.} \kappa^{cof(\kappa)}$ for the reason that there is no $g : \kappa \xrightarrow{onto} {}^{cof(\kappa)}\kappa$. Given any mapping $g : \kappa \to {}^{cof(\kappa)}\kappa$ and any strictly increasing $f : cof(\kappa) \xrightarrow{cof.} \kappa$, we construct by a diagonal argument some mapping $h : cof(\kappa) \to \kappa$ that does not belong to $ran(g)$:

$$\forall \alpha < cof(\kappa) \quad h(\alpha) = \textit{least in } \{\beta \in \kappa \mid \forall \xi \leqslant f(\alpha) \quad g(\xi)(\alpha) \neq \beta\}$$

Notice that $h(\alpha)$ is well-defined since $cof(\kappa)$ being an infinite cardinal, it follows $\alpha < cof(\kappa)$ implies $|f(\alpha)| < \kappa$. From this, we obtain $\big|\{g(\xi)(\alpha) \mid \xi \leqslant f(\alpha)\}\big| < \kappa$ which shows that the set $\{\beta \in \kappa \mid \forall \xi \leqslant f(\alpha) \quad g(\xi)(\alpha) \neq \beta\}$ is non-empty.

$\square$ König's Lemma

Corollary 119 (ZFC). *Let κ be any infinite cardinal.*

$$cof(2^\kappa) > \kappa.$$

Proof of Corollary 119: The result is immediate from both

- $(2^\kappa)^\kappa = 2^\kappa$ $\quad\big($since $(2^\kappa)^\kappa = 2^{\kappa\cdot\kappa} = 2^\kappa\big)$
- $(2^\kappa)^{cof(2^\kappa)} > 2^\kappa$. $\quad\big($by König's Lemma, page 73$\big)$

$\square$ 119

In particular, for $\kappa = \aleph_0$, we have $cof\big(2^{\aleph_0}\big) > \aleph_0$, which shows that — under **AC** — we cannot have $|\mathscr{P}(\omega)| = \aleph_\omega$, nor $|\mathscr{P}(\omega)| = \sup_{n<\omega} \underbrace{\aleph_{\aleph_{\aleph_{\cdot_{\cdot_{\cdot_{\aleph_0}}}}}}}_{n}$.

4.6 First Encounter with CH, GCH, and Inaccessible Cardinals

On August 8 1900, at the Paris conference of the International Congress of Mathematicians which took place at the Sorbonne, Hilbert presented a list of 10 problems which he later extended to a list of 23 problems. These were thought out as the most challenging mathematical problems for the century to come. The first such problem — known as Hilbert's first problem — was the *Continuum Hypothesis*: *"there is no set of reals $A \subseteq \mathbb{R}$ such that $\omega \lnsim A \lnsim \mathbb{R}$"*. Or in other words, *every uncountable set $A \subseteq \mathbb{R}$ satisfies $A \simeq \mathbb{R}$.*

Definition 120 (**ZFC**).

The Continuum Hypothesis (**CH**) *is the closed formula*

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

A stronger version of the *Continuum Hypothesis* consists in generalising it to all infinite cardinals.

Definition 121 (**ZFC**).

The Generalized Continuum Hypothesis (**GCH**) *is the closed formula*

$$\forall \alpha \in \mathbf{On} \ \ 2^{\aleph_\alpha} = \aleph_{\alpha+1}.$$

Definition 122 (**ZFC**). *the functional* $\beth : \mathbf{On} \to \mathbf{On}$ *is defined by transfinite recursion:*

- $\beth_0 = \aleph_0$
- $\beth_{\alpha+1} = 2^{\beth_\alpha}$
- $\beth_\beta = \sup_{\alpha<\beta} 2^{\beth_\alpha}$ *(for β limit).*

Remark 123. The Generalized Continuum Hypothesis (**GCH**) sates that the aleph sequence and the beth sequence coincide:

$$\forall \alpha \in \mathbf{On} \ \ \aleph_\alpha = \beth_\alpha.$$

By **ZFC**+**CH** or **ZFC**+**GCH** we denote the extended theories **ZFC**$\cup\{$**CH**$\}$ or **ZFC**$\cup\{$**GCH**$\}$. We will show that one cannot prove nor disprove neither the *Generalized Continuum Hypothesis* nor the *Continuum Hypothesis*. Indeed, we will show the following

(1)
$$\mathbf{ZFC} + \mathbf{CH} \vdash_c \bot \implies \mathbf{ZFC} \vdash_c \bot$$

which, by contraposition, yields

$$\mathbf{ZFC} \nvdash \bot \implies \mathbf{ZFC} + \mathbf{CH} \nvdash \bot$$

and by the Completeness Theorem of first order logic:

$$\mathbf{ZFC} \not\models \bot \implies \mathbf{ZFC} + \mathbf{CH} \not\models \bot$$

which reads: *if there exists some model that satisfies* **ZFC**, *then there is also some model that satisfies* **ZFC+CH**.

(2)

$$\mathbf{ZFC} + \neg\mathbf{CH} \vdash_c \bot \implies \mathbf{ZFC} \vdash_c \bot$$

which, by contraposition, yields

$$\mathbf{ZFC} \nvdash \bot \implies \mathbf{ZFC} + \neg\mathbf{CH} \nvdash \bot$$

and by the Completeness Theorem of first order logic:

$$\mathbf{ZFC} \not\models \bot \implies \mathbf{ZFC} + \neg\mathbf{CH} \not\models \bot$$

which reads: *if there exists some model that satisfies* **ZFC**, *then there is also some model that satisfies* **ZFC+¬CH**.

Independently of **CH**, we will show

$$\mathbf{ZFC} \vdash_c \bot \iff \mathbf{ZF} \vdash_c \bot \iff \mathbf{ZF} + \neg\mathbf{AC} \vdash_c \bot$$

which will bring that all theories **ZF**, **ZF + AC**, **ZF + ¬AC**, **ZF + CH**, **ZF + ¬CH** all have the same consistency strength.

Definition 124 (**ZFC**). *Let $\aleph_0 < \kappa$ be any cardinal,*

- *κ is weakly inaccessible iff κ is regular and limit,*
- *κ is strongly inaccessible iff κ is regular and $\forall\lambda < \kappa \;\; 2^\lambda < \kappa$.*

Remarks 125.

- Every strongly inaccessible cardinal is also weakly inaccessible.
- Under the Generalized Continuum Hypothesis (**GCH**), a cardinal is weakly inaccessible if and only if it is strongly inaccessible.

Unless **ZFC** is inconsistent, one cannot prove the existence of neither a strongly inaccessible cardinal nor a weakly inaccessible cardinal. The reason for this will be that Unless is inconsistent, **ZFC** cannot prove its own consistency. But, if there exists some strongly (or even weakly) inaccessible cardinal, then one can construct a model that satisfies **ZFC**, proving this way that **ZFC** is consistent.

Chapter 5

Well-Founded Sets and the Axiom of Foundation

5.1 Well-Founded Sets

Following Kunen's notation, we now work in $\mathbf{ZF} \setminus \{\mathbf{AF}\} = \mathbf{ZF} \setminus \{\mathbf{Foundation}\}$. We define of well-founded sets.

Definition 126 ($\mathbf{ZF} \setminus \{\mathbf{AF}\}$)**.** *By transfinite recursion, we define* $\mathbf{W} : \mathbf{On} \to \mathbf{V}$ *by*

- $\mathbf{W}(0) = \varnothing$
- $\mathbf{W}(\alpha + 1) = \mathscr{P}\big(\mathbf{W}(\alpha)\big)$
- $\mathbf{W}(\alpha) = \bigcup_{\xi < \alpha} \mathbf{W}(\xi)$ *(when α is limit).*

$$\mathbf{WF} = \bigcup_{\alpha \in \mathbf{On}} \mathbf{W}(\alpha).$$

Definition 127 ($\mathbf{ZF} \setminus \{\mathbf{AF}\}$)**.** *If $x \in \mathbf{WF}$, then*

$$rk(x) = \text{least } \alpha \in \mathbf{On} \ \text{ such that } \ x \in \mathbf{W}(\alpha + 1).$$

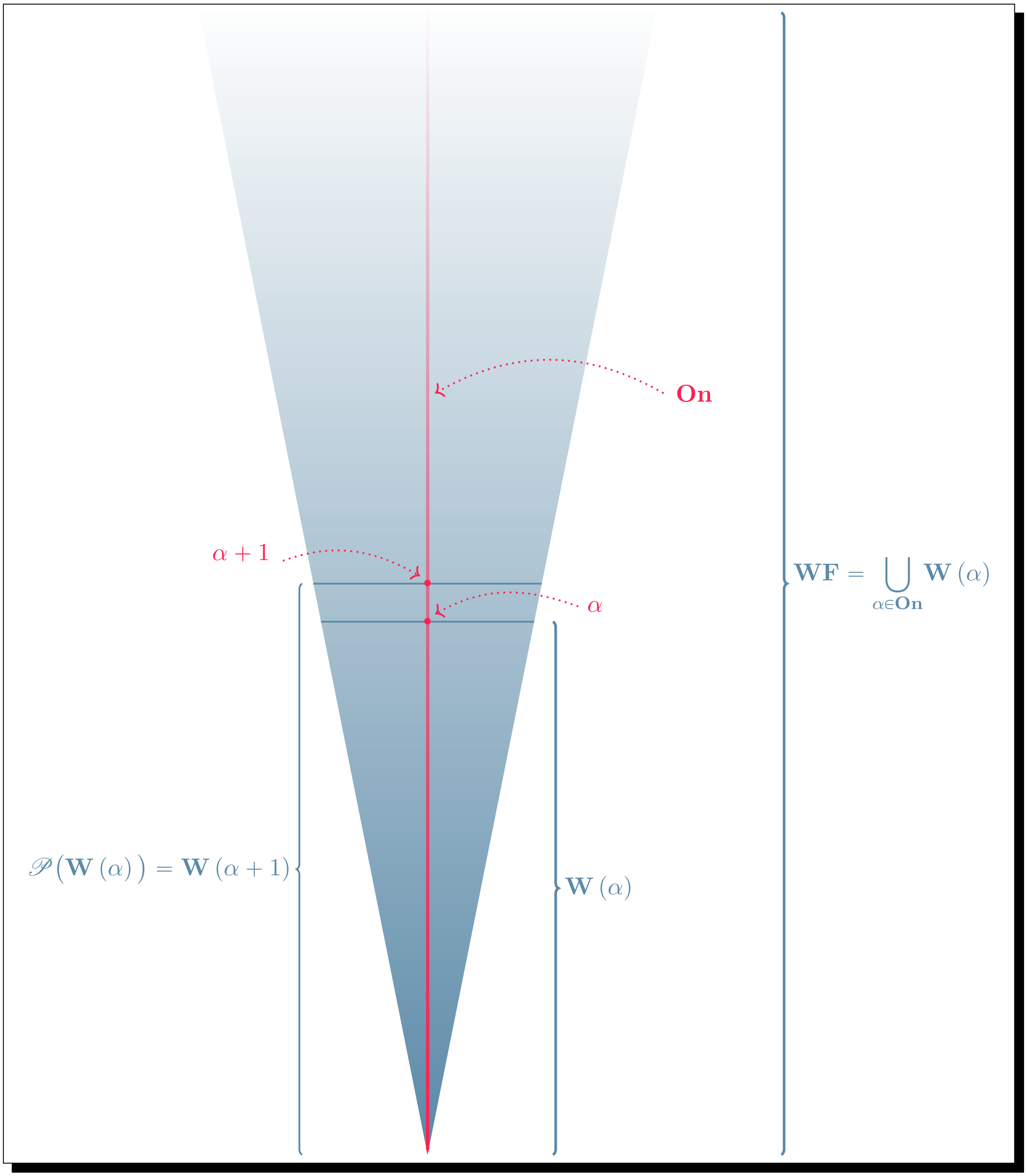


Figure 5.1: The Class of Well-Founded Sets $\mathbf{WF} = \bigcup_{\alpha \in \mathbf{On}} \mathbf{W}(\alpha)$.

Example 128 ($\mathbf{ZF} \smallsetminus \{\mathbf{AF}\}$)**.** *If* $x \in \mathbf{WF}$, *then*

- $rk(0) = rk(\varnothing) = 0$
- $rk(1) = rk(\{\varnothing\}) = 1$
- $rk(\{\varnothing, \{\varnothing, \{\varnothing\}\}\}) = 3$
- $rk(2) = rk(\{\varnothing, \{\varnothing\}\}) = 2$
- $rk(3) = rk(\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}) = 3$
- $rk(\omega) = \omega$.

Lemma 129 ($\mathbf{ZF} \smallsetminus \{\mathbf{AF}\}$)**.** *Given any* $\alpha \in \mathbf{On}$,

(1) $\mathbf{W}(\alpha)$ *is a transitive set*

(2) $\forall \xi \leqslant \alpha \quad \mathbf{W}(\xi) \subseteq \mathbf{W}(\alpha)$.

Proof of Lemma 129: We prove simultaneously (1) and (2) by induction on α.

$\boldsymbol{\alpha} = \mathbf{0}$**:** for (1) one has $\mathbf{W}(0) = \varnothing$ which is a transitive set, and for (2) there is nothing to do.

$\boldsymbol{\alpha} = \boldsymbol{\beta} + \mathbf{1}$**:** for (1) we consider

$$x \in y \in \mathscr{P}\big(\mathbf{W}(\beta)\big) = \mathbf{W}(\alpha)$$

i.e.,

$$x \in y \subseteq \mathbf{W}(\beta)$$

which comes down to

$$x \in \mathbf{W}(\beta).$$

By induction hypothesis, $\mathbf{W}(\beta)$ is transitive, so we obtain

$$x \subseteq \mathbf{W}(\beta).$$

i.e.,

$$x \in \mathscr{P}\big(\mathbf{W}(\beta)\big) = \mathbf{W}(\alpha).$$

For (2), we first notice that by induction hypothesis, for every $\xi \leqslant \beta$ one has

$$\mathbf{W}(\xi) \subseteq \mathbf{W}(\beta)$$

and one also has

$$\mathbf{W}(\beta) \subseteq \mathscr{P}\big(\mathbf{W}(\beta)\big) = \mathbf{W}(\alpha)$$

since, given any $x \in \mathbf{W}(\beta)$, one has $x \subseteq \mathbf{W}(\beta)$ by transitivity of $\mathbf{W}(\beta)$; hence $x \in \mathscr{P}\big(\mathbf{W}(\beta)\big)$. So, we obtain that for every $\xi \leqslant \alpha$

$$\mathbf{W}(\xi) \subseteq \mathbf{W}(\alpha).$$

α limit: for (1) we only need to show the following to get the result:

Claim 130. *The union of any transitive set is transitive:*

Proof of Claim 130: Let A be any transitive set. If $b \in \bigcup A$, then there exists $a \in A$, $b \in a \in A$, so that for every $c \in b$, one has $c \in b \in a \in A$, hence, by transitivity of A, $c \in b \in A$, i.e., $c \in \bigcup A$, which shows $b \subseteq \bigcup A$.

$\square$ 130

For (2), the result is immediate by definition.

$\square$ 129

Lemma 131 ($\mathbf{ZF} \smallsetminus \{\mathbf{AF}\}$)**.** *Given any $\alpha \in \mathbf{On}$,*

$$\mathbf{W}(\alpha) = \{x \in \mathbf{WF} \mid rk(x) < \alpha\}.$$

Proof of Lemma 131: For all $x \in \mathbf{WF}$, one has

- $rk(x) < \alpha$ if and only if for some $\beta < \alpha$, $x \in \mathbf{W}(\beta + 1)$.
- $x \in \mathbf{W}(\beta + 1)$ for some $\beta < \alpha$ if and only if[1] $x \in \mathbf{W}(\alpha)$.

$\square$ 131

Lemma 132 ($\mathbf{ZF} \smallsetminus \{\mathbf{AF}\}$)**.** *Given any $y \in \mathbf{WF}$,*

(1) $\forall x \in \mathbf{WF}\ \ \forall y \in \mathbf{WF}\ \ \big(x \in y \longrightarrow rk(x) < rk(y)\big)$

(2) $rk(y) = \sup\big\{rk(x) + 1 \mid x \in y\big\}$

Proof of Lemma 132:

[1] This is by transitivity of $\mathbf{W}(\xi)$ (any $\xi \in \mathbf{On}$.).

(1) Take $x \in y \in \mathbf{WF}$. We assume $rk(y) = \alpha$, which yields

$$y \in \mathbf{W}(\alpha + 1) = \mathscr{P}\big(\mathbf{W}(\alpha)\big).$$

This leads to $x \in \mathbf{W}(\alpha)$, hence $rk(x) < \alpha$.

(2) We set $\alpha = \sup\big\{rk(x) + 1 \mid x \in y\big\}$.

- By 132 (1), we have $rk(x) < rk(y)$ which leads to $\alpha \leqslant rk(y)$.
- Since for every $x \in y$, one has $rk(x) < \alpha$, it follows that $x \in \mathbf{W}(\alpha)$, hence $y \subseteq \mathbf{W}(\alpha)$, i.e., $y \in \mathscr{P}\big(\mathbf{W}(\alpha)\big) = \mathbf{W}(\alpha + 1)$. Finally, we obtain $rk(y) \leqslant \alpha$.

□ 132

Definition 133. *A sequence is a function s from any ordinal α to any set A.*

$$s : \alpha \to A$$

- *$dom(s) = \alpha$ is called the length of s and is also denoted by $lh(s)$.*
- *If $s(\xi) = a_\xi$ holds for each ordinal $\xi < \alpha$, we denote s by $(a_\xi)_{\xi<\alpha}$ or by $\langle a_\xi \mid \xi < \alpha\rangle$.*

Remarks 134.

- Given any set x, the only sequence $s : \varnothing \to x$ is the empty function $\varnothing$ called the empty sequence, whose length is 0.
- Given any sets x, y, the couple (x, y) is *not the same* as the sequence $\langle x, y\rangle$ since
  - $(x, y) = \{\{x\}, \{x, y\}\}$
  - $$\begin{aligned}\langle x, y\rangle &= \{(0, x), (1, y)\}\\ &= \Big\{\big\{\{0\}, \{0, x\}\big\}, \big\{\{1\}, \{1, y\}\big\}\Big\}\\ &= \Big\{\big\{\{\varnothing\}, \{\varnothing, x\}\big\}, \big\{\{\{\varnothing\}\}, \{\{\varnothing\}, y\}\big\}\Big\}.\end{aligned}$$

Definition 135. *Given any sequences $s : \alpha \to A$ and $t : \beta \to B$, the concatenation of s and t is the sequence*

$$s^\frown t : \alpha + \beta \to A \cup B$$

- $dom\left(s^\frown t\right) = \alpha + \beta$
- $ran\left(s^\frown t\right) \subseteq A \cup B$
- $\forall \xi < \alpha + \beta \quad \begin{cases} s^\frown t(\xi) & = & s(\xi) & \text{if} \quad \xi < \alpha \\ s^\frown t(\xi) & = & t(\zeta) & \text{if} \quad \exists \zeta \quad \xi = \alpha + \zeta. \end{cases}$

Definition 136. *Given any $\alpha \in \mathbf{On}$ and any set A,*

- $A^\alpha = {}^\alpha A = \{s : \alpha \to A\}$
- $A^{<\alpha} = \bigcup_{\beta<\alpha} A^\beta = \bigcup \{A^\beta \mid \beta < \alpha\}$
- $A^{\leqslant\alpha} = \bigcup_{\beta\leqslant\alpha} A^\beta = \bigcup \{A^\beta \mid \beta \leqslant \alpha\}.$

Definition 137. *Given any set A, a tree on A is a set $T \subseteq A^{<\omega}$ closed under prefix. i.e.,*

$$\forall s \in T \; \forall n < lh(s) \quad s \restriction n \in T.$$

- *If $T \neq \varnothing$, then $\varnothing$ (the empty sequence) belongs to T and is called the* root *of T.*
- *If $s \in T$, any $t \in T$ s.t. $lh(t) = lh(s) + 1$ and $t \restriction lh(s) = s$ is called a* child *of s.*
- *Any $s \in T$ without children is called a* terminal node *or a* leaf*.*
- *If $\mathfrak{b} \in A^\omega$ and $\forall n \in \omega \;\; \mathfrak{b} \restriction n \in T$, then $\mathfrak{b}$ is called an* infinite branch *of T.*

Example 138 ($\mathbf{ZF} \setminus \{\mathbf{AF}\}$)**.**

- *$\varnothing$ is a tree on ω (the empty tree)*
- *$\{\varnothing\}$ is a tree on ω (the tree reduced to its root)*
- *$\omega^{<\omega}$ is a tree on ω*
- $\left\{s \in \omega^{<\omega} \mid \forall n \; \forall m \; \left(n < m < lh(s) \longrightarrow s(n) < s(m)\right)\right\}$ *is a tree on ω*
- $\left\{s \in \omega^{<\omega} \mid \forall n \; \forall m \; \left(n < m < lh(s) \longrightarrow s(n) > s(m)\right)\right\}$ *is a tree on ω.*

Definition 139. *Given any set A, T is a non-empty pruned (n.e.p.) tree on A if T has no terminal node. i.e.,*

$$\forall s \in T \ \exists a \in A \quad s^\frown \langle a \rangle \in T.$$

Example 140 ($\mathbf{ZF} \setminus \{\mathbf{AF}\}$)**.**

- $\varnothing$ *is not a n.e.p. tree on* ω
- $\{\varnothing\}$ *is not a n.e.p. tree on* ω
- $\omega^{<\omega}$ *is a n.e.p. tree on* ω
- $2^{<\omega}$ *is a n.e.p. tree on* ω
- $\Big\{s \in \omega^{<\omega} \mid \forall n \ \forall m \ \big(n < m < lh(s) \longrightarrow s(n) < s(m)\big)\Big\}$ *is a n.e.p. tree on* ω
- $\Big\{s \in \omega^{<\omega} \mid \forall n \ \forall m \ \big(n < m < lh(s) \longrightarrow s(n) > s(m)\big)\Big\}$ *is not a n.e.p. tree* [2] *on* ω

Definition 141 (**ZF**)**.** *The Axiom of Dependent Choice (*$\mathbf{DC}$*) is the statement*

"every non-empty pruned tree admits an infinite branch".

i.e.,

$$\forall A \ \forall T \subseteq A^{<\omega} \quad \Big(\text{"}T \text{ is a n.e.p. tree"} \longrightarrow \exists \mathfrak{b} \in A^\omega \ \forall n \in \omega \quad \mathfrak{b} \restriction n \in T \Big).$$

Remarks 142.

- The Axiom of Dependent Choice is a weak version of the Axiom of Choice. One has

  $$\mathbf{ZFC} \vdash_c \mathbf{DC}.$$

  But, assuming that **ZF** is consistent, one has

  $$\mathbf{ZF} + \mathbf{DC} \nvdash \mathbf{AC}.$$

[2] Indeed, $\langle 0 \rangle$ belongs to $\Big\{s \in \omega^{<\omega} \mid \forall n \ \forall m \ \big(n < m < lh(s) \longrightarrow s(n) > s(m)\big)\Big\}$ but does not have any children, hence it is a terminal node of this tree.

- The Axiom of Dependent Choice is a stronger version of the Axiom of Countable Choice. One has
$$\mathbf{ZF} + \mathbf{DC} \vdash_c \mathbf{CC}.$$
But, assuming that **ZF** is consistent, one has
$$\mathbf{ZF} + \mathbf{CC} \nvdash \mathbf{DC}.$$
- Under **ZF**, **DC** is equivalent to the Baire Category Theorem for complete metric spaces which states that

Every complete metric space is a Baire space;

where a topological space is a Baire space if for each family $(O_n)_{n\in\omega}$ of dense open subsets, their intersection $\bigcap_{n\in\omega} O_n$ is dense.

Corollary 143 (**ZFC**)**.** *If $x_0 \in \mathbf{WF}$, then there is no infinite $\ni$-descending sequence*
$$x_0 \ni x_1 \ni x_2 \ni \ldots \ni x_n \ni x_{n+1} \ldots\ldots$$

Proof of Corollary 143: Towards a contradiction, we assume there exists such an $\ni$-descending sequence $(x_i)_{i\in\omega}$ and consider
$$\{rk(x_i) \in \mathbf{On} \mid i \in \omega\}.$$
This non-empty set of ordinals has no minimal element, since by Lemma 132 (1) one has $rk(x_i) > rk(x_{i+1})$ holds for each integer i. This leads to
$$x_0 \ni x_1 \ni x_2 \ni \ldots \ni x_n \ni x_{n+1} \ldots\ldots$$
$$\Longrightarrow$$
$$rk(x_0) > rk(x_1) > rk(x_2) > \ldots > rk(x_n) > rk(x_{n+1}) \ldots\ldots$$
contradicting Theorem 34 (5).

□ 143

Lemma 144 (**ZF** ∖ {**AF**})**.** *Given any $\alpha \in \mathbf{On}$,*

(1) $\mathbf{On} \subseteq \mathbf{WF}$

(2) $\forall \alpha \in \mathbf{On} \quad rk(\alpha) = \alpha$

(3) $\forall \alpha \in \mathbf{On} \quad \mathbf{W}(\alpha) \cap \mathbf{On} = \alpha$.

Proof of Lemma 144:

- We prove (1) and (2) simultaneously by induction on α

  $\boldsymbol{\alpha} = 0$: $\varnothing \in \{\varnothing\} = \mathscr{P}(\varnothing) = \mathbf{W}(1)$ and $rk(\varnothing) = 0$

  $\boldsymbol{\alpha} = \boldsymbol{\beta} + 1$: $\alpha = \beta \cup \{\beta\}$. By induction hypothesis, since $rk(\beta) = \beta$, on has $\beta \in \mathbf{W}(\beta + 1)$. Since $\mathbf{W}(\beta + 1)$ is a transitive set by Lemma 129 (1), it follows that $\beta \subseteq \mathbf{W}(\beta + 1)$, hence $\beta \cup \{\beta\} \subseteq \mathbf{W}(\beta + 1)$ which leads to

  $$\alpha = \beta \cup \{\beta\} \in \mathscr{P}\big(\mathbf{W}(\beta + 1)\big) = \mathbf{W}(\alpha + 1).$$

  By Lemma 132 (2)

  $$rk(\alpha) = \sup\big\{rk(x) + 1 \mid x \in \alpha\big\} = \sup\big\{rk(x) + 1 \mid x \in \beta \ \vee \ x = \beta\big\}$$

  Since by Lemma 132 (1) for each $x \in \beta$, one has $rk(x) < rk(\beta)$, we obtain

  $$rk(\alpha) = \sup\big\{rk(x) + 1 \mid x \in \beta \ \vee \ x = \beta\big\} = rk(\beta) + 1 = \beta + 1 = \alpha.$$

  $\boldsymbol{\alpha}$ limit: $\alpha = \sup\big\{\beta \mid \beta < \alpha\big\}$. By induction hypothesis, for each $\beta < \alpha$ one has

  $$\beta \in \mathbf{W}(\beta + 1) \ \text{ and } rk(\beta) = \beta$$

  so that

  $$\alpha = \big\{\beta \mid \beta < \alpha\big\} \subseteq \bigcup_{\beta < \alpha} \mathbf{W}(\beta) = \mathbf{W}(\alpha)$$

  hence

  $$\alpha = \big\{\beta \mid \beta < \alpha\big\} \in \mathscr{P}\left(\bigcup_{\beta < \alpha} \mathbf{W}(\beta)\right) = \mathscr{P}\big(\mathbf{W}(\alpha)\big) = \mathbf{W}(\alpha + 1).$$

  By Lemma 132 (2) one has

  $$rk(\alpha) = \sup\big\{rk(\beta) + 1 \mid \beta \in \alpha\big\} = \sup\big\{\beta + 1 \mid \beta \in \alpha\big\} = \alpha.$$

- For (3) it is enough to notice that

  $$\mathbf{W}(\alpha) \cap \mathbf{On} = \big\{\beta \in \mathbf{On} \mid rk(\beta) < \alpha\big\} = \big\{\beta \in \mathbf{On} \mid \beta < \alpha\big\} = \alpha.$$

$\square$ 144

Lemma 145 ($\mathbf{ZF} \setminus \{\mathbf{AF}\}$).

$$\forall y \ \ (y \in \mathbf{WF} \longleftrightarrow y \subseteq \mathbf{WF}).$$

Proof of Lemma 145:

(⇒) $y \in \mathbf{WF}$ if and only if $y \in \mathbf{W}(\alpha)$ for some ordinal α. By Lemma 129 (1) $\mathbf{W}(\alpha)$ is transitive, which gives $y \subseteq \mathbf{W}(\alpha) \subseteq \mathbf{WF}$.

(⇐) Let $\alpha = \sup\{rk(x) \mid x \in y\}$, one has $\forall x \in y \ \ x \in \mathbf{W}(\alpha+1)$, hence $y \subseteq \mathbf{W}(\alpha+1)$, i.e., $y \in \mathbf{W}(\alpha+2) \subseteq \mathbf{WF}$.

$\square$ 145

Remarks 146.

- If $x \in \mathbf{WF}$, then the following sets belong to $\mathbf{WF}$:

$$\begin{array}{ccrcl}
\{x\} & \rightsquigarrow & rk(\{x\}) & = & rk(x)+1 \\
\bigcup x & \rightsquigarrow & rk(\bigcup x) & \leqslant & rk(x) \\
\mathscr{P}(x) & \rightsquigarrow & rk\big(\mathscr{P}(x)\big) & = & rk(x)+1 \\
Sx & \rightsquigarrow & rk\big(Sx\big) & = & rk(x)+1
\end{array}$$

- If $x, y \in \mathbf{WF}$, then the following sets belong to $\mathbf{WF}$:

$$\begin{array}{ccrcl}
\{x,y\} & \rightsquigarrow & rk(\{x,y\}) & = & \max\{rk(x), rk(y)\}+1 \\
x \times y & \rightsquigarrow & rk(x \times y) & = & \max\{rk(x), rk(y)\}+3 \\
x \cup y & \rightsquigarrow & rk\big(x \cup y\big) & = & \max\{rk(x), rk(y)\} \\
x \cap y & \rightsquigarrow & rk\big(\mathscr{P}(x)\big) & \leqslant & \max\{rk(x), rk(y)\} \\
(x,y) & \rightsquigarrow & rk\big((x,y)\big) & = & \max\{rk(x), rk(y)\}+2 \\
\langle x,y\rangle & \rightsquigarrow & rk\big(\langle x,y\rangle\big) & = & \max\{rk(x), rk(y)\}+3 \\
{}^{x}y & \rightsquigarrow & rk\big({}^{x}y\big) & \leqslant & \max\{rk(x), rk(y)\}+4.
\end{array}$$

Definition 147 (**ZF** $\smallsetminus$ {**AF**})**.** *A relation* $<_A$ *is well-founded on a set* A *iff every non-empty subset of* A *admits some* $<_A$*-minimal element. i.e.,*

$$\forall B \subseteq A \ \left(B \neq \varnothing \longrightarrow \exists y \in B \ \forall x \in B \ \ x \nless_A y\right).$$

A relation that is not well-founded *is called* ill-founded*.*

Lemma 148 (**ZF** $\smallsetminus$ {**AF**}+**DC**)**.** *Let* $<_A$ *be some relation on* A*.*

$$<_A \ \textit{is well-founded on } A$$

$$\Longleftrightarrow$$

$$\neg\exists (a_i)_{i\in\omega} \ \ \forall i \in \omega \ \ a_{i+1} <_A a_i.$$

Proof of Lemma 148:

(⇒) By contraposition, we assume

$$\exists (a_i)_{i\in\omega} \ \ \forall i \in \omega \ \ a_{i+1} <_A a_i.$$

and form $B = \{a_i \mid i \in \omega\} \neq \varnothing$ which satisfies

$$\forall y \in B \ \exists x \in B \ \ x <_A y.$$

(⇐) By contraposition, we assume $<_A$ is ill-founded on A. So, there exists $B \subseteq A$, $B \neq \varnothing$ such that

$$\forall y \in B \ \exists x \in B \ \ x <_A y.$$

One forms

$$T = \left\{s \in B^{<\omega} \mid \forall n \ \forall m \ \left(n < m < lh(s) \longrightarrow s(m) <_A s(n)\right)\right\}.$$

Notice that T is a n.e.p. tree on B. By **DC**, there exists some infinite branch $\mathtt{b} \in B^\omega$ which is an infinitely $<_A$-decreasing sequence.

$\square$ 148

Lemma 149 (**ZF** $\smallsetminus$ {**AF**})**.** *Let* A *be any set.*

$$\textit{If } A \in \mathbf{WF}, \ \textit{then} \in \textit{is well-founded on } A.$$

Proof of Lemma 149:

- If $A = \varnothing$, then the result is obvious since $\in_A = \varnothing$.
- If $A \neq \varnothing$, then consider

$$\alpha = \textit{ least ordinal in } \{rk(x) \mid x \in A\}.$$

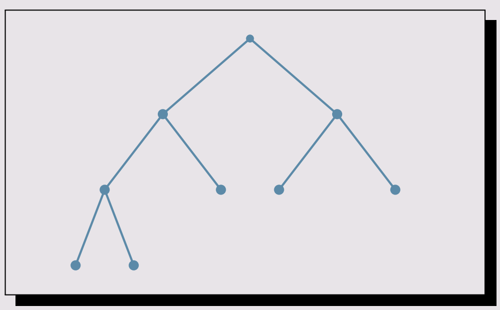
Now, every element $a \in A$ such that $rk(a) = \alpha$ is a $\in$-least element in A since

$$\forall x \in \mathbf{WF}\ \ \forall y \in \mathbf{WF}\ \ \big(x \in y \longrightarrow rk(x) < rk(y)\big).$$

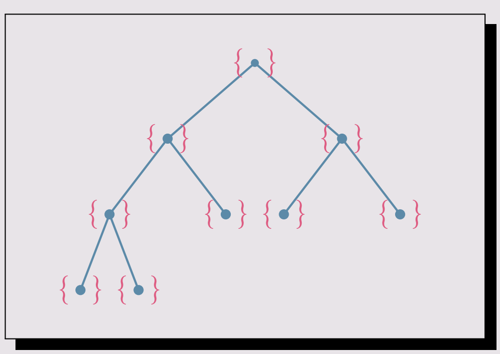
holds by Lemma 132(1).

□ 149

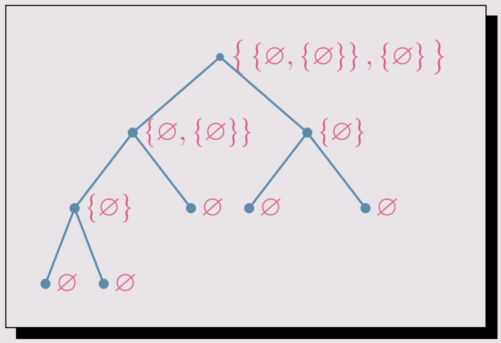
Remark 150. There is a nice way to represent well-founded sets (the sets in **WF**) by well-founded trees (unordered trees[3]).
Imagine you are given a well-founded tree such as the one below:



Then, you can associate to each node of the tree, the set of the sets associated with each of its children. So that the leaves of the tree are associated with the empty set as shown in the following picture.



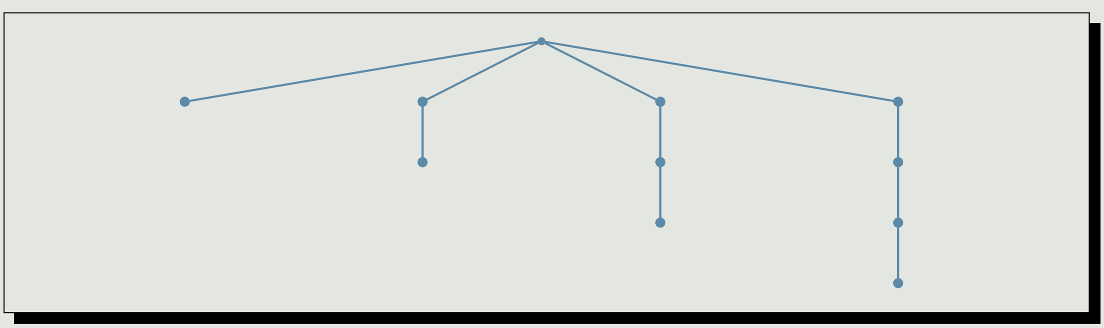
Precisely, to each node $\mathtt{n}$ we associate $\widehat{\mathtt{n}} = \{\widehat{\mathtt{c}} \mid \mathtt{c} \text{ is a child of } \mathtt{n}\}$ which gives the result as described in the picture below:



A well-founded tree represents the set that is collected at its root. For instance, the tree above represents the set $\big\{\{\varnothing,\{\varnothing\}\},\{\varnothing\}\big\}$.

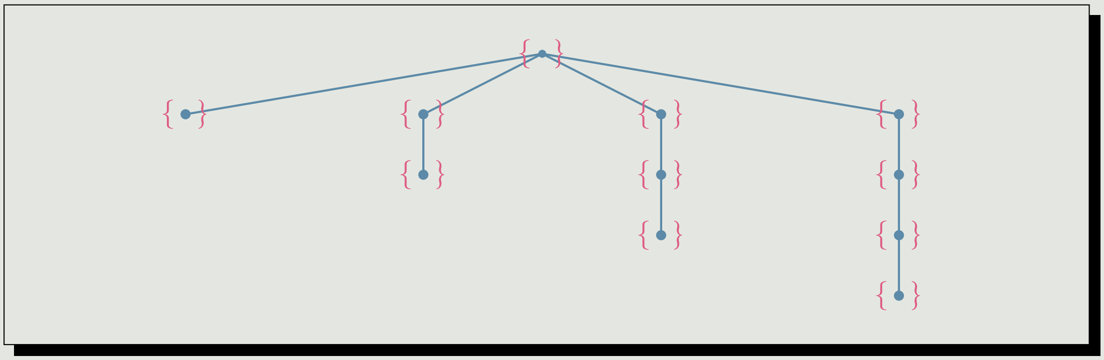
Example 151. *A tree that represents the set*

$$\Big\{\varnothing,\{\varnothing\},\{\{\varnothing\}\},\big\{\{\{\varnothing\}\}\big\}\Big\}$$

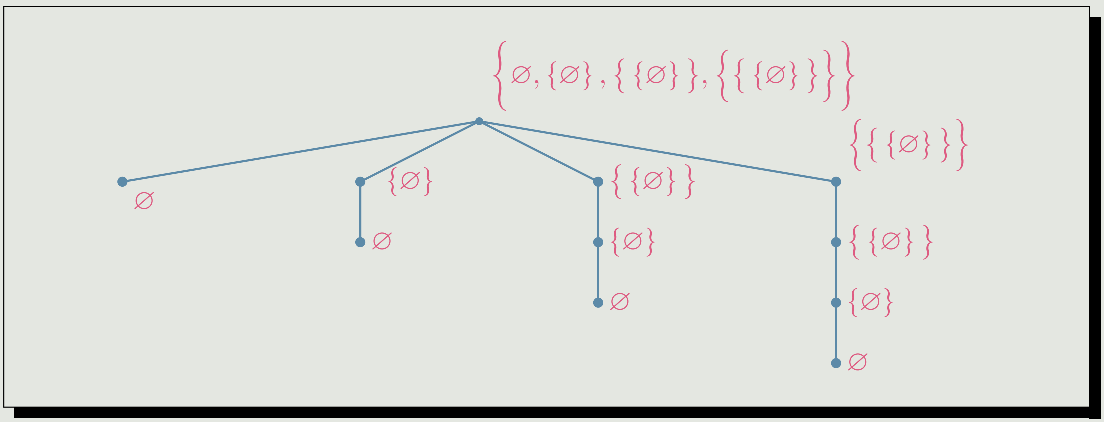


By taking into account that at each node $\mathtt{n}$ we associate $\widehat{\mathtt{n}} = \{\widehat{\mathtt{c}} \mid \mathtt{c}$ is a child of $\mathtt{n}\}$ we aim at what is described in the next pictures:

[3]A tree is ordered if for each node there is a an ordering on its children. A tree is unordered if there is no such ordering, which means that on a given picture, the ordering of the children of a node from left to right does not matter.

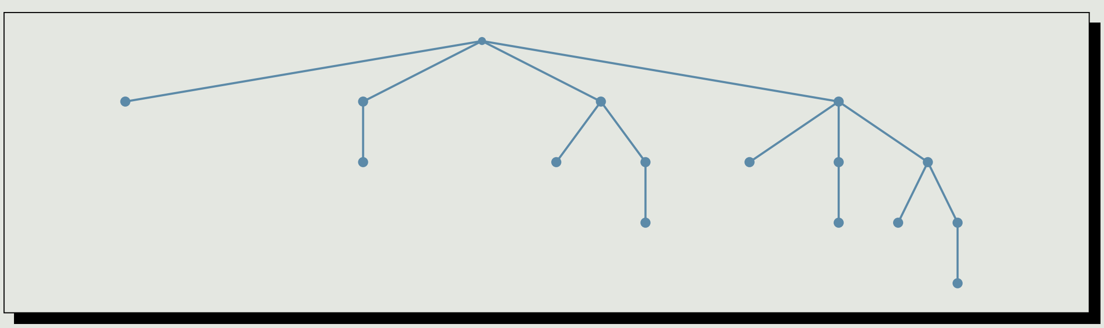


The set $\widehat{\mathfrak{n}}$ *associated to the node* $\mathfrak{n}$ *is written at the right of* $\mathfrak{n}$*.*

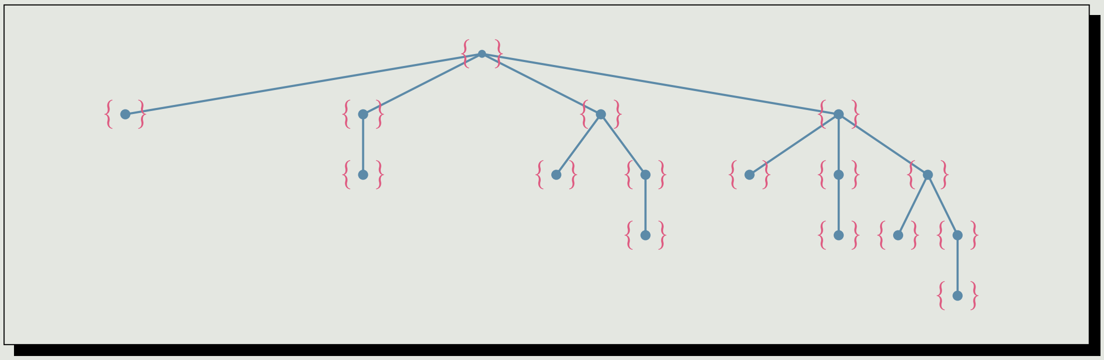


Example 152. *A tree that represents the ordinal 4 where*

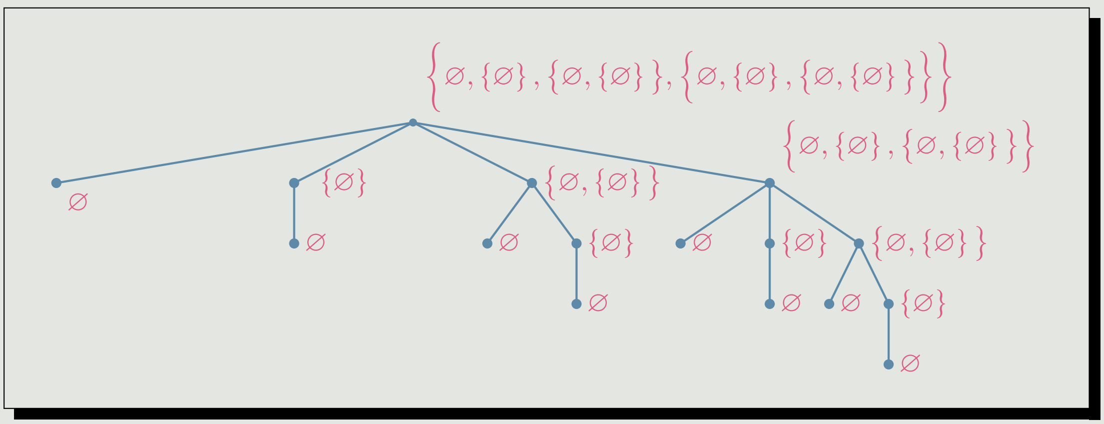
$$4 = \Big\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \Big\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\Big\}\Big\}.$$



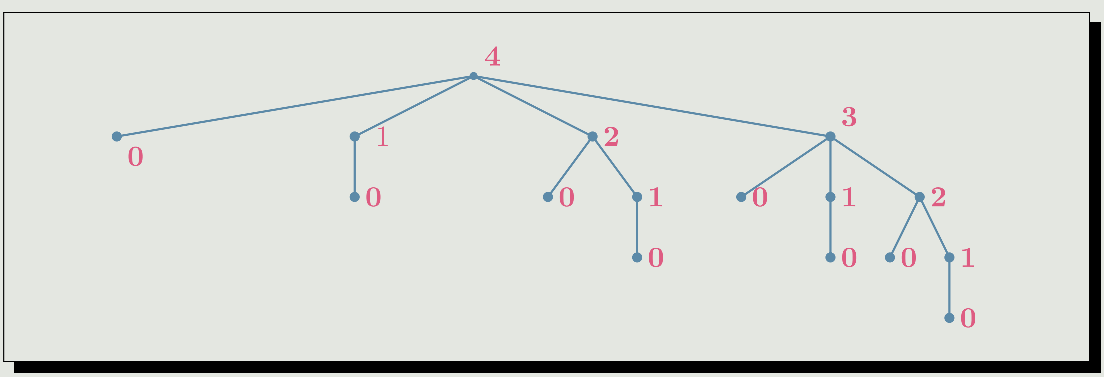
By taking into account that at each node $\mathfrak{n}$ we associate $\widehat{\mathfrak{n}} = \{\widehat{\mathfrak{c}} \mid \mathfrak{c}$ is a child of $\mathfrak{n}\}$ we aim at what is described in the next pictures:



The set $\widehat{\mathfrak{n}}$ associated to the node $\mathfrak{n}$ is written at the right of $\mathfrak{n}$.



For a better visualisation, we replace the ordinals by their usual descriptions in base 10.



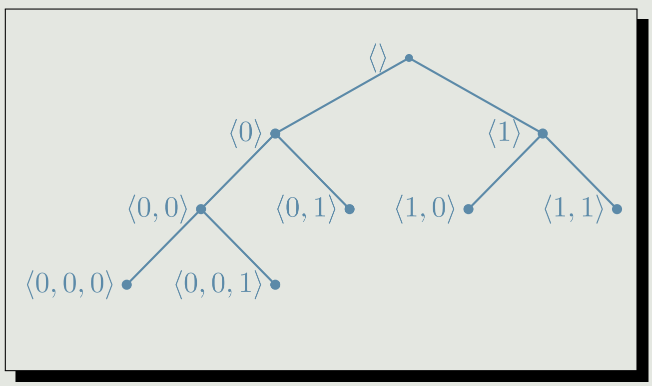
Definition 153 (Height of a non-empty well-founded tree)**.** *Let A be any set and $T \subseteq A^{<\omega}$ be any non-empty well-founded tree on A. We define $ht : T \to \mathbf{On}$ by*

- *if s is a leaf*[4]*: $ht(s) = 0$,*
- *if s is not a leaf: $ht(s) = \sup\left\{ht\big(s^\frown \langle a\rangle\big) + 1 \in tc(A)^{<\omega} \mid s^\frown \langle a\rangle \in T\right\}$.*

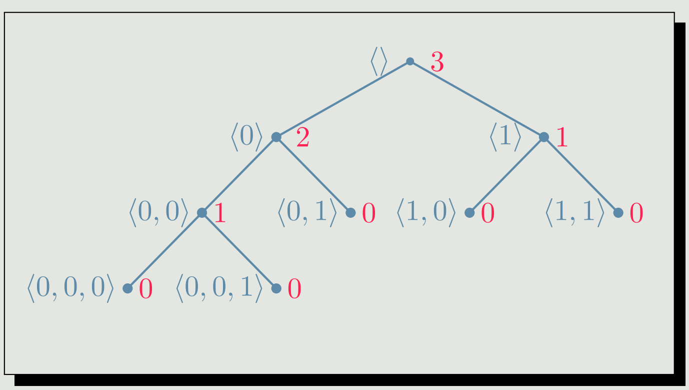
The height of the tree is the height of its root: $ht(T) = ht(\langle\rangle)$.

Clearly, the height of a node which is not a leaf is the least ordinal which is strictly greater than the height of every one of its children.

Example 154. *A tree $T = \Big\{ \langle\rangle, \langle 0\rangle, \langle 1\rangle, \langle 0,0\rangle, \langle 1,0\rangle, \langle 0,1\rangle, \langle 1,1\rangle, \langle 0,0,0\rangle, \langle 0,0,1\rangle \Big\}$.*



The heights of each node of this tree:



So, the height of this tree is 3.

[4] The sequence s is a leaf if for all $a \in A$, $s^\frown \langle a\rangle \notin T$.

We saw that from any well-founded tree we could obtain a set which will necessarily be in **WF**. Now, we focus on the same kind of operation, but the other way round: from any well-founded set A, obtain in a canonical way, a well-founded tree $\mathbb{T}_A$ that represents A.

Definition 155 (The canonical tree representation of a well-founded set)**.** *Let $A \in \mathbf{WF}$ be any set. The canonical tree $\mathbb{T}_A$ that represents A is defined as*

- *if $A = \varnothing$, then $\mathbb{T}_A = \{\langle\rangle\} = \{\varnothing\}$;*
- *if $A \neq \varnothing$, then*

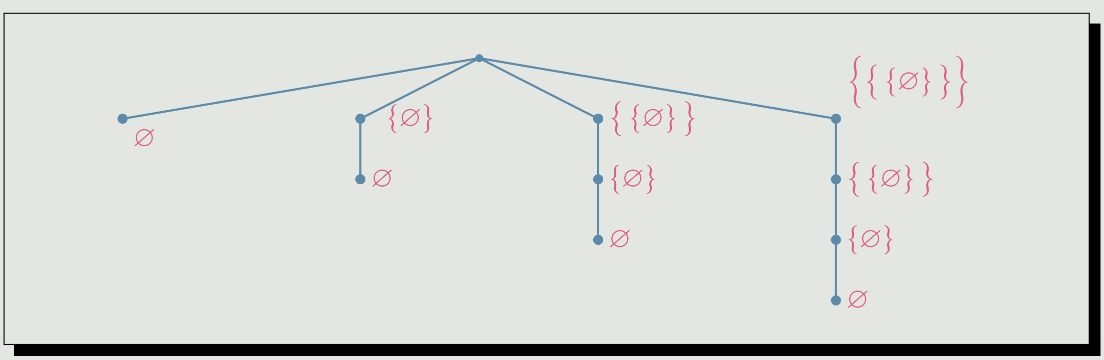
$$\begin{aligned}\mathbb{T}_A &= \{\langle\rangle\} \cup \left\{s \in tc(A)^{<\omega} \mid lh(s) > 0 \wedge s(0) \in A \wedge \forall i < lh(s) - 1 \;\; s(i) \ni s(i+1)\right\}\\ &= \{\langle\rangle\} \cup \left\{\langle x_0, x_1, x_2, \ldots, x_k\rangle \mid k \in \omega \text{ and } A \ni x_0 \ni x_1 \ni x_2 \ni \ldots \ni x_k\right\}.\end{aligned}$$

(For the definition of $tc(A)$, see Definition 159.)

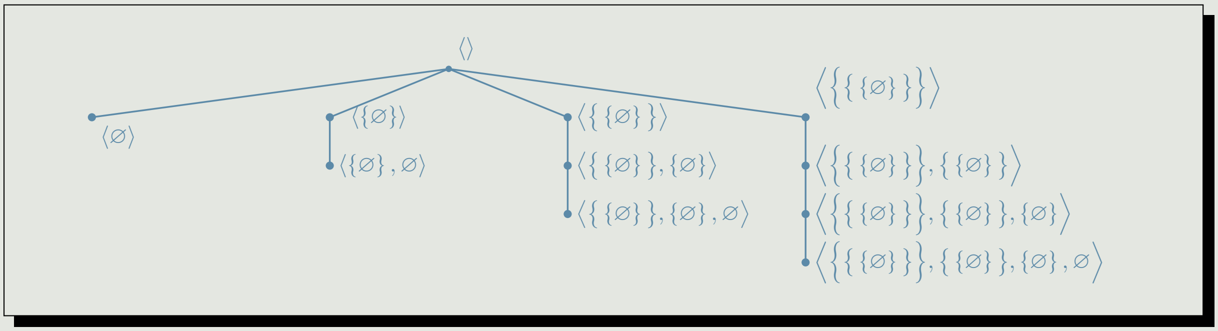
Notice that $\mathbb{T}_A$ has no infinite branch, since $A \in \mathbf{WF}$.

Example 156. *The canonical tree representation $\mathbb{T}_A$ of the well-founded set*

$$A = \left\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \left\{\{\{\varnothing\}\}\right\}\right\}$$



*As a **set** this tree is the set of all the 11 sequences that are shown below:*



Proposition 157. *Let A be any well-founded set and $\mathbb{T}_A$ its canonical tree representation.*

$$ht(\mathbb{T}_A) = rk(A)$$

Proof of Proposition 157: By induction on $rk(A)$:

- $rk(A) = 0 \iff A = \varnothing \iff \mathbb{T}_A$ is a single node[5] $\iff ht(\mathbb{T}_A) = 0$.
- $rk(A) = \alpha > 0 \iff rk(A) = \sup\left\{rk(x) + 1 \mid x \in A\right\} \iff rk(A) = \sup\left\{ht(\mathbb{T}_x) + 1 \mid x \in A\right\} = ht(\mathbb{T}_A)$.

$\square$ 157

Lemma 158 ($\mathbf{ZF} \smallsetminus \{\mathbf{AF}\}$)**.** *Let A be any set.*

*If A is transitive and $\in$ is well-founded on A, then $A \in$ **WF**.*

Proof of Lemma 158: By Lemma 145 it is enough to show that $A \subseteq \mathbf{WF}$. So, towards a contradiction, we assume $A \smallsetminus \mathbf{WF} \neq \varnothing$ and consider some $\in$-minimal element $a \in A \smallsetminus \mathbf{WF}$. For each $x \in a$, one has $x \in A$ since A is transitive, and $x \notin A \smallsetminus \mathbf{WF}$, hence $x \in \mathbf{WF}$. This leads to $a \subseteq \mathbf{WF}$ and once again, by Lemma 145, to $a \in \mathbf{WF}$, a contradiction.

$$\alpha = \textit{ least ordinal in } \{rk(x) \mid x \in B\}.$$

Now every element $a \in B$ such that $rk(a) = \alpha$ is a $\in$-least element in B since by Lemma 132(1) one has $\forall y \in \mathbf{WF}\ \forall z \in y\ \big(z \in \mathbf{WF}\ rk(z) < rk(y)\big)$.

[5] Both root and leaf at the same time.

$\square$ 158

Given any set A, if we intend to turn A into some transitive set by adding the minimal amount of elements to A, what we need to do is to add the elements of the elements of A, then the elements of the elements of the elements of A, and the elements of the elements of the elements of the elements of A,... etc.
This is what is called the "*transitive closure of A*" and is defined by

Definition 159 ($\mathbf{ZF} \setminus \{\mathbf{AF}\}$)**.** *Let A be any set.*
The transitive closure of A — denoted by $tc(A)$ — is defined by recursion on the integers:

$$\begin{aligned} \bigcup^0 A &= A \\ \bigcup^{n+1} A &= \bigcup \left(\bigcup^n A \right) \\ tc(A) &= \bigcup \left\{ \bigcup^n A \mid n \in \omega \right\}. \end{aligned}$$

Lemma 160 ($\mathbf{ZF} \setminus \{\mathbf{AF}\}$)**.** *Let A be any set.*

$$tc(A) \text{ is the } \subseteq\text{-least transitive set such that } A \subseteq tc(A).$$

Proof of Lemma 160: $A \subseteq tc(A)$ is by definition. To show that $tc(A)$ is transitive, it is enough to notice that for every integer n, one has that if $x \in \bigcup^n A$, then $x \subseteq \bigcup^{n+1} A$. To show that $tc(A)$ is the $\subseteq$-least such set, we prove by induction on $n \in \omega$ that any set B that is both transitive and contains A contains each $\bigcup^n A$.

($\mathbf{n := 0}$): $\bigcup^0 A = A \subseteq B$ by definition.

($\mathbf{n := k+1}$): assuming $\bigcup^k A \subseteq B$ and B is transitive yields that for every $x \in \bigcup^k$, one has $x \subseteq B$. i.e., $\bigcup \left(\bigcup^k A \right) \subseteq B$ which is $\bigcup^{k+1} A \subseteq B$.

$\square$ 160

Theorem 161 ($\mathbf{ZF} \setminus \{\mathbf{AF}\}$)**.** *Let A be any set. The following are equivalent:*

(1) $\in$ *is well-founded on* $tc(A)$

(2) $tc(A) \in \mathbf{WF}$

(3) $A \in \mathbf{WF}$.

Proof of Lemma 161:

(1) ⇒ (2): since $tc(A)$ is both transitive and well-founded by $\in$, we obtain by Lemma 158 that $tc(A) \in \mathbf{WF}$.

(2) ⇒ (3): by applying Lemma 145 twice, we obtain

$$A \subseteq tc(A) \in \mathbf{WF} \Longrightarrow A \subseteq tc(A) \subseteq \mathbf{WF} \Longrightarrow A \subseteq \mathbf{WF} \Longrightarrow A \in \mathbf{WF}.$$

(3) ⇒ (1): by Remarks 146, $\mathbf{WF}$ is closed under $\bigcup : \mathbf{V} \to \mathbf{V}$. Hence, for each integer n we have both $\bigcup^n A \in \mathbf{WF}$, and by Lemma 145 $\bigcup^n A \subseteq \mathbf{WF}$. This leads to

$$tc(A) = \bigcup \{ \bigcup^n A \mid n \in \omega \} \subseteq \mathbf{WF},$$

and by Lemma 145 once again to $tc(A) \in \mathbf{WF}$.

□ 161

5.2 Foundation

In this section we will see that the Axiom of **Foundation** essentially restricts the universe of all sets to the ones that are $\in$-well-founded. However, the Axiom of **Foundation** is not stated in terms of the well-foundedness of the membership relation. It says that every non-emptyset has an element with which it has no element in common:

$$\forall x \left(\exists y\, y \in x \to \exists y \left(y \in x \ \wedge\ \neg \exists z \left(z \in x \ \wedge\ z \in y \right) \right) \right).$$

This forbids, for instance, sets of the form $x = \{x\}$ or even any set that satisfies $x \in x$, since any such set leads to $x \in \{x\}$ and $x \in x$, hence $x \in \left(\{x\} \cap x \right)$ which shows that $\{x\}$ does not contain any element with which it has no element in common.

Theorem 162 (**ZF** $\smallsetminus$ {**AF**})**.** *The following are equivalent:*

(1) Axiom of ***Foundation***

(2) the membership relation is well-founded on each set

(3) $\mathbf{V} = \mathbf{WF}$.

Proof of Lemma 162:

(1) ⇒ (2): given any non-empty set A, by the Axiom of **Foundation**, there exists some element $a \in A$ such that $a \cap A = \varnothing$ holds. So, a is some $\in$-minimal element in A.

(2) ⇒ (3): given any set A, one has that the membership relation is well-founded on $tc(A)$, and by Theorem 161, $A \in \mathbf{WF}$.

(3) ⇒ (1): given any non-empty set A, since $A \in \mathbf{WF}$ we consider

$$\alpha = \text{least ordinal inside } \{rk(x) \in \mathbf{On} \mid x \in A\}$$

and pick any $a \in A$ with $rk(a) = \alpha$. It follows that either $a = \varnothing$ or $\forall x \in a \;\; x \notin A$.

□ 162

So, taking into account the Axiom of **Foundation** forbids the models of **ZF** or **ZFC** that contain non-well-founded sets. We may think this is too restrictive. But on the other hand, when all sets are well-founded by the membership relation, we can make use of recursion everywhere for constructing many objects on the basis of any sets.

The question we should ask however is:

"do we risk to render set theory inconsistent while adding the Axiom of ***Foundation****?"*

We will see that the answer is: *"no!"*. Essentially, what we will show is that if there exists some model of **ZF** or **ZFC**, then we can restrict this model to the sole well-founded sets and obtain another model that still satisfies **ZF** or **ZFC**. So, we will prove:

$$\mathbf{ZF} \setminus \{\mathbf{AF}\} \nvdash \bot \implies \mathbf{ZF} \nvdash \bot$$

which is logically equivalent to

$$\mathbf{ZF} \vdash_c \bot \implies \mathbf{ZF} \setminus \{\mathbf{AF}\} \vdash_c \bot$$

which says that if **ZF** is inconsistent, then **ZF** without the Axiom of **Foundation** is already inconsistent as well. To do this, we will need to define — within the language of set theory — an operation that consists in restricting ourselves to some given class (in particular to **WF**). This operation is called relativization.
Working with the Axiom of **Foundation**, since $\mathbf{V} = \mathbf{WF}$ holds one usually uses the notation

Definition 163 (**ZF**)**.** *When working with the Axiom of* ***Foundation****,* $\mathbf{V} = \mathbf{WF}$ *holds. Hence, one usually uses the notation* $\mathbf{V}(\alpha)$ *instead of* $\mathbf{W}(\alpha)$*. This means that by transfinite recursion, one defines*

- $\mathbf{V}(0) = \varnothing$
- $\mathbf{V}(\alpha + 1) = \mathscr{P}\big(\mathbf{V}(\alpha)\big)$
- $\mathbf{V}(\alpha) = \bigcup_{\xi<\alpha} \mathbf{V}(\xi)$ *(when α is limit).*

$$\mathbf{V} = \bigcup_{\alpha\in\mathbf{On}} \mathbf{V}(\alpha).$$

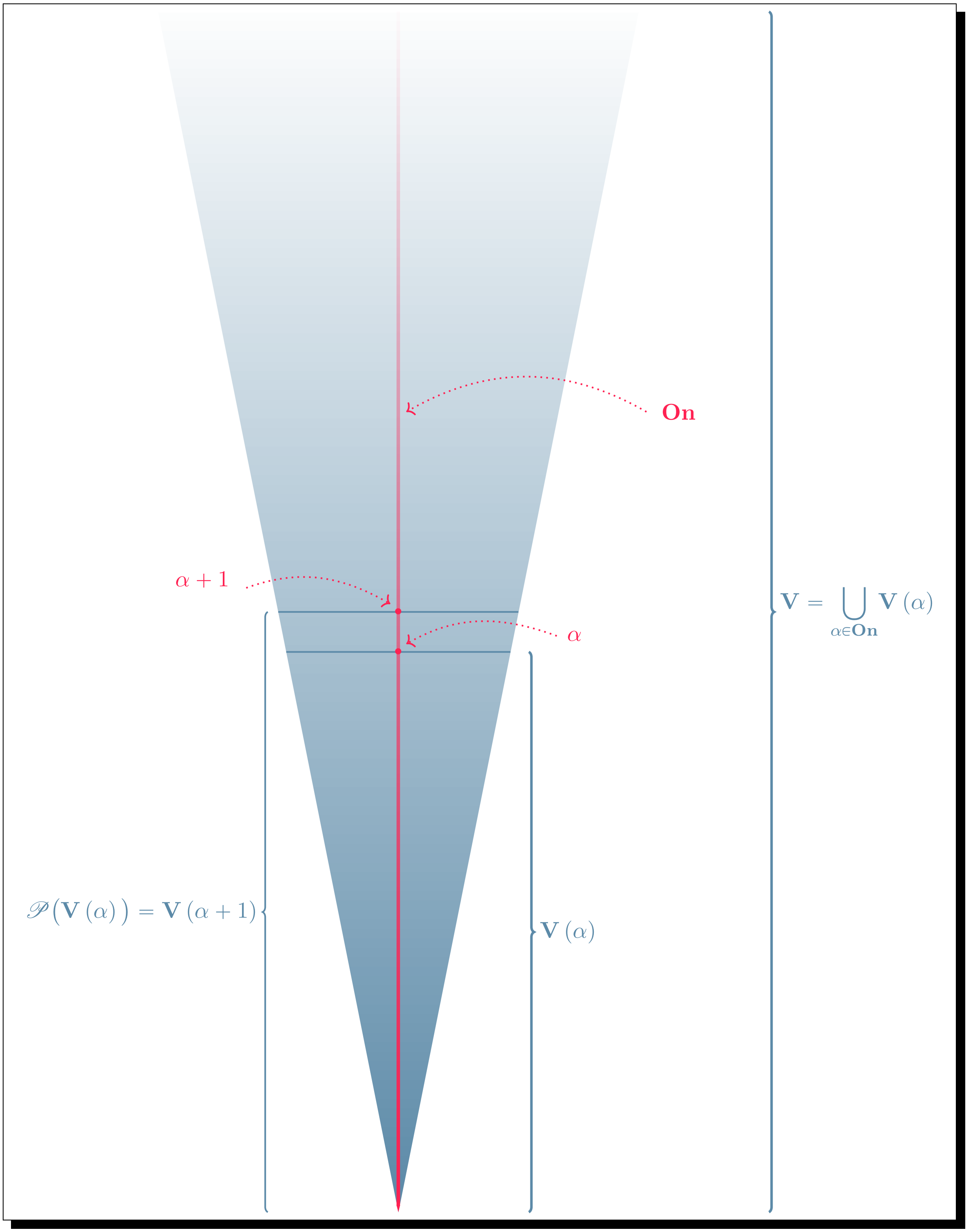


Figure 5.2: The Universe $\mathbf{V} = \bigcup_{\alpha \in \mathbf{On}} \mathbf{V}(\alpha)$.