Logic and Set Theory

Jan Geuenich

Universität Bielefeld – Sommersemester 2019

Contents

1 Logic

1.1 Boolean rings

The field $\mathbb{F}_2 = \{0,1\}$ with two elements is a natural choice to encode logical values. Here 0 should be interpreted as "false" and 1 as "true". More generally, the following class of rings will turn out to be useful for the proof of independence results:

Definition 1.1.1. *A* boolean ring *is a unitary ring R such that all its elements are* idempotent, *i.e. $x^2 = x$ holds for all $x \in R$.*

Every boolean ring R can be equipped with a unary operation $\neg$ (NOT), six binary operations $\wedge, \vee, \rightarrow, \leftrightarrow, \uparrow, \downarrow$ (AND, OR, THEN, IFF, NAND, NOR), and a relation $\leq$ given for all $x, y \in R$ as follows:

$$\begin{array}{rcl} \neg x & = & x+1 \\ x \wedge y & = & xy \\ x \vee y & = & x+y+xy \\ x \rightarrow y & = & \neg x \vee y \\ x \leftrightarrow y & = & (x \rightarrow y) \wedge (y \rightarrow x) \\ x \uparrow y & = & \neg(x \wedge y) \\ x \downarrow y & = & \neg(x \vee y) \\ x \leq y & \Leftrightarrow & x \rightarrow y = 1 \end{array}$$

Exercise 1.1.2. Each of the operations $\neg, +, \cdot, \wedge, \vee, \rightarrow, \leftrightarrow, \downarrow$ on a boolean ring can be expressed using just the *NAND operation* $\uparrow$ (e.g. $\neg x = x \uparrow x$).

Lemma 1.1.3. *Every boolean ring $R \neq 0$ is commutative and has characteristic 2.*

Proof. Let $x, y \in R$. Using $x^2 = x$, $y^2 = y$ and $x^2 + xy + yx + y^2 = (x+y)^2 = x + y$ we get $2x = 0$ with $x = y$ and then in general $xy = -yx = yx$. $\square$

Ring-theoretic ideals in boolean rings are order-theoretic ideals:

Lemma 1.1.4. *A subset I of a boolean ring R is an ideal if and only if $0 \in I$ and $x, y \in I \Rightarrow x \vee y \in I$ and $x \leq y \in I \Rightarrow x \in I$ for all $x, y \in R$.*

All prime ideals in a boolean ring R are maximal and an ideal I in R is maximal if and only if for all $x \in R$ we have the equivalence $x \notin I \Leftrightarrow \neg x \in I$.

Proof. For every ideal I we have $0 \in I$ and $x, y \in I$ implies $x \vee y = x + y + xy \in I$ and $x \leq y \in I$ implies $x = x(x \rightarrow y) = x(\neg x \vee y) = x(\neg x + \neg xy + y) = xy \in I$.

Conversely, let $I \subseteq R$ with $0 \in I$ and $x, y \in I \Rightarrow x \vee y \in I$ and $x \leq y \in I \Rightarrow x \in I$ for all $x, y \in R$. Using Lemma 1.1.3 we see $x + y = (x\neg y) \vee (\neg xy)$ so that it suffices to check $xy \leq y$ for all $x, y \in R$ in order to conclude that I is an ideal:

$$(xy) \rightarrow y \;=\; \neg(xy) \vee y \;=\; (xy+1) + y + (xy+1)y \;=\; 1$$

This proves the first statement.

If I is a prime ideal and $x \notin I$, then $\neg x \in I$ because of $x \neg x = 0 \in I$. If I is a proper ideal and $x \in I$, then $\neg x \notin I$ because of $x + \neg x = 1 \notin I$. But ideals I with the property $x \notin I \Leftrightarrow \neg x \in I$ must be maximal, since then $1 = \neg x + x \in I + Rx$ for every $x \notin I$. This proves the second statement, since maximal ideals are prime. $\square$

We will often use the following rules to calculate in boolean rings (see Problem Set 1, Exercise 1 for the next two exercises):

Exercise 1.1.5. For every boolean ring R and $u, v, x, y, z \in R$ we have:

(i) $\wedge, \vee, +$ are associative.

(ii) $\wedge, \vee$ are idempotent.

(iii) $\wedge, \vee, +, \leftrightarrow, \uparrow, \downarrow$ are commutative.

(iv) $x \wedge 0 = 0$, $x \wedge 1 = x$ and $x \vee 0 = x$, $x \vee 1 = 1$.

(v) $\wedge, \vee$ are *mutually absorptive*: $x \wedge (x \vee y) = x$ and $x \vee (x \wedge y) = x$.

(vi) *De Morgan's laws hold*: $\neg(x \wedge y) = \neg x \vee \neg y$ and $\neg(x \vee y) = \neg x \wedge \neg y$.

(vii) $x \to y = \neg y \to \neg x$.

(viii) $x \leftrightarrow y = 1 \Leftrightarrow x = y$.

(ix) $(x \wedge y) \to z = x \to (y \to z)$.

(x) $(x \wedge \neg y) \vee (\neg x \wedge y) = x + y = (x \vee y) \wedge \neg(x \wedge y)$.

(xi) $x \leq y \;\Leftrightarrow\; x \wedge y = x \;\Leftrightarrow\; x \vee y = y \;\Leftrightarrow\; x \wedge \neg y = 0 \;\Leftrightarrow\; y \mid x$.

(xii) $(u \leq x$ and $v \leq y) \;\Rightarrow\; (u \wedge v \leq x \wedge y$ and $u \vee v \leq x \vee y)$.

(xiii) $(x \wedge y = 0$ and $x \vee y = 1) \;\Rightarrow\; y = \neg x$.

Exercise 1.1.6. Every boolean ring R is a *boolean algebra* (= *complemented distributive lattice*) w.r.t. $(\leq, 0, 1, \neg, \wedge, \vee)$, i.e. for $x, y, z \in R$:

(I) $\leq$ is a partial order on R with least element 0 and greatest element 1.

(II) $x \wedge y$ is the infimum and $x \vee y$ the supremum of $\{x, y\}$ w.r.t. $\leq$.

(III) $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ and $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$.

(IV) $x \wedge \neg x = 0$ and $x \vee \neg x = 1$.

Conversely, every boolean algebra $(R, \leq, 0, 1, \neg, \wedge, \vee)$ defines a boolean ring R with addition and multiplication given for $x, y \in R$ by:

$$\begin{aligned} x + y &= (x \wedge \neg y) \vee (\neg x \wedge y) \\ x \cdot y &= x \wedge y \end{aligned}$$

Exercise 1.1.7. Using Lemma 1.1.4 and Exercise 1.1.6, the principal ideal generated by an element y of a boolean ring R is easily seen to be $\{x \in R : x \leq y\}$.

Beside $\mathbb{F}_2$ the following is the prototypical example of a boolean ring:

Example 1.1.8. For every set X the power set $\mathcal{P}(X)$ is a boolean ring with

$$\begin{aligned} A + B &= (A \setminus B) \cup (B \setminus A)\ , \\ A \cdot B &= A \cap B \end{aligned}$$

for all $A, B \in \mathcal{P}(X)$. In this case $(0, 1, \neg, \wedge, \vee, \leq) = (\emptyset, X, X\setminus, \cap, \cup, \subseteq)$.

As a subring of $\mathcal{P}(X)$, every σ-algebra on a set X is a boolean ring.

Remark 1.1.9. The function $\mathcal{P}(X) \to {}^X\mathbb{F}_2$ that associates with subsets of X their characteristic functions is a ring isomorphism.

The `NOT` operation $\neg$ for boolean rings induces a duality in the following sense:

Lemma 1.1.10. *If a set R is a boolean algebra w.r.t. $(\leq, 0, 1, \neg, \wedge, \vee)$, then it is also a boolean algebra w.r.t. $(\leq^{\mathrm{op}}, 1, 0, \neg, \vee, \wedge)$ and $\neg$ defines an involutive isomorphism from $(R, \leq, 0, 1, \neg, \wedge, \vee, \leftrightarrow, +, \uparrow, \downarrow)$ to $(R, \leq^{\mathrm{op}}, 1, 0, \neg, \vee, \wedge, +, \leftrightarrow, \downarrow, \uparrow)$.*

Proof. The first part is obvious. The last part easily follows with Exercise 1.1.5. $\square$

Definition 1.1.11. *The* dual $R^\neg$ *of a boolean ring R is the boolean ring induced on the set R by the boolean-algebra structure $(\leq^{\mathrm{op}}, 1, 0, \neg, \vee, \wedge)$.*

A subset $F \subseteq R$ is said to be a filter *resp. an* ultrafilter *in the boolean ring R if F is an ideal resp. a maximal ideal in $R^\neg$.*

Lemma 1.1.12. *A subset F of a boolean ring R is a filter if and only if $1 \in F$ and $x, y \in F \Rightarrow x \wedge y \in F$ and $x \geq y \in F \Rightarrow x \in F$ for all $x, y \in R$.*

A filter F in a boolean ring R is an ultrafilter if and only if for all $x \in R$ we have the equivalence $x \notin F \Leftrightarrow \neg x \in F$.

Proof. This is Lemma 1.1.4 for $R^\neg$. $\square$

Lemma 1.1.13 (Ultrafilter Lemma)**.** *Every proper filter in a boolean ring R is contained in some ultrafilter.*

Proof. This is a reformulation of the fact that every proper ideal in $R^\neg$ is contained in a maximal ideal. (The proof uses Zorn's lemma.) $\square$

To be able to assign boolean values to first-order formulas with quantifiers we will mostly work with boolean rings that are complete in the following sense:

Definition 1.1.14. *A boolean ring R is said to be* complete w.r.t. $U \subseteq \mathcal{P}(R)$ *if every $Y \in U$ has an infimum $\bigwedge Y$ and a supremum $\bigvee Y$ w.r.t. $\leq$.*

A boolean ring R is complete *if R is complete w.r.t. $\mathcal{P}(R)$.*

Example 1.1.15. The power set $\mathcal{P}(X)$ of any set X is a complete boolean ring with infimum $\bigwedge Y = \bigcap Y$ and supremum $\bigvee Y = \bigcup Y$ for every $Y \subseteq \mathcal{P}(X)$.

Every σ-algebra is as a boolean ring complete w.r.t. the set of its countable subsets.

Example 1.1.16. In view of Exercise 1.1.6 every boolean ring R is complete w.r.t. the set of its finite subsets $Y = \{y_1, \ldots, y_n\} \subseteq R$ with

$$\bigwedge Y \;=\; y_1 \wedge \cdots \wedge y_n \quad \text{and} \quad \bigvee Y \;=\; y_1 \vee \cdots \vee y_n \,.$$

As a consequence, every finite boolean ring is complete, in particular so is $\mathbb{F}_2$.

1.2 Formulas

We will now introduce a formal language that we will use later on to talk about "sets" and, more generally, about elements of other mathematical structures.

Definition 1.2.1. *Let $S = (S_L, S_{R,r}, S_{F,r}, S_C)_{r\in\mathbb{N}}$ be a* vocabulary *consisting of sets*

$$\begin{array}{lcll} S_L & = & \{\perp, \rightarrow, \bigwedge\} & \textit{of } \text{logical symbols,} \\ S_{R,r} & & & \textit{of } r\text{-ary relation symbols,} \\ S_{F,r} & & & \textit{of } r\text{-ary function symbols,} \end{array}$$

that are pairwise disjoint and with $S_C \subseteq S_{F,0}$.

Part of this vocabulary S are the sets

$$\begin{array}{lcll} S_T & = & S_{R,0} & \textit{of } \text{truth symbols,} \\ S_I & = & S_{F,0} & \textit{of } \text{individual symbols,} \\ S_V & = & S_I \setminus S_C & \textit{of } \text{variable symbols,} \\ S_C & & & \textit{of } \text{constant symbols.} \end{array}$$

*The set $\mathcal{T} = \bigcup_{n\in\mathbb{N}} \mathcal{T}_n$ of S-*terms *is with $\mathcal{T}_{-1} = \emptyset$ defined by*

$$\begin{aligned} \mathcal{T}_0 \;&=\; \big\{(x) : x \in S_I\big\}\,, \\ \mathcal{T}_n \;&=\; \mathcal{T}_{n-1} \\ &\cup\; \big\{(\sharp, t_1, \ldots, t_r) : r \in \mathbb{N}, \sharp \in S_{F,r}, (t_1, \ldots, t_r) \in \mathcal{T}^r_{n-1} \setminus \mathcal{T}^r_{n-2}\big\}\,. \end{aligned}$$

*The set $\mathcal{F} = \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ of S-*formulas *with $\mathcal{F}_{-1} = \emptyset$ is defined by*

$$\begin{aligned} \mathcal{F}_0 \;&=\; \big\{(\perp)\big\} \cup \big\{(\sharp, t_1, \ldots, t_r) : r \in \mathbb{N}, \sharp \in S_{R,r}, (t_1, \ldots, t_r) \in \mathcal{T}^r\big\}\,, \\ \mathcal{F}_n \;&=\; \mathcal{F}_{n-1} \\ &\cup\; \big\{(\rightarrow, \varphi, \psi) : (\varphi, \psi) \in \mathcal{F}^2_{n-1} \setminus \mathcal{F}^2_{n-2}\big\} \\ &\cup\; \big\{(\bigwedge, x, \varphi) : x \in S_V, \varphi \in \mathcal{F}_{n-1} \setminus \mathcal{F}_{n-2}\big\}\,. \end{aligned}$$

S-terms t in $\mathcal{T}_n \setminus \mathcal{T}_{n-1}$ and S-formulas π in $\mathcal{F}_n \setminus \mathcal{F}_{n-1}$ have rank $\mathrm{rk}(t) = n$ *and* $\mathrm{rk}(\pi) = n$*, respectively. They are called* atomic *if they have rank* 0.

Whenever we want to stress the dependence on the vocabulary S, we will write $\mathcal{T}_S$ and $\mathcal{F}_S$ instead of $\mathcal{T}$ and $\mathcal{F}$.

Using induction on the rank of formulas and terms we have the following important statements about their unique decomposition:

Remark 1.2.2. Every S-term t has the form $(\sharp, t_1, \dots, t_r)$ with uniquely determined $r \in \mathbb{N}$, $\sharp \in S_{F,r}$ and S-terms $t_1, \dots, t_r$ of strictly smaller rank than t.

Similarly, every S-formula π occurs in exactly one of the forms $(\bot)$, $(\sharp, t_1, \dots, t_r)$, $(\to, \varphi, \psi)$, $(\bigwedge, x, \varphi)$ with uniquely determined entries in either case, where φ and ψ are S-formulas of strictly smaller rank than π.

The remark makes it possible to prove facts about terms and formulas by *structural induction* and to define functions with domain $\mathcal{T}$ or $\mathcal{F}$ by *structural recursion.*

Definition 1.2.3. *The* symbols *of an S-term t are recursively defined as*

$$\operatorname{sym}(t) \;=\; \{\sharp\} \cup \operatorname{sym}(t_1) \cup \dots \cup \operatorname{sym}(t_r) \quad \textit{for } t = (\sharp, t_1, \dots, t_r) \textit{ with } \sharp \in S_{F,r},$$

and the symbols *of an S-formula π as*

$$\operatorname{sym}(\pi) = \begin{cases} \{\bot\} & \textit{for } \pi = (\bot), \\ \{\sharp\} \cup \operatorname{sym}(t_1) \cup \dots \cup \operatorname{sym}(t_r) & \textit{for } \pi = (\sharp, t_1, \dots, t_r) \textit{ with } \sharp \in S_{R,r}, \\ \{\to\} \cup \operatorname{sym}(\varphi) \cup \operatorname{sym}(\psi) & \textit{for } \pi = (\to, \varphi, \psi), \\ \{\bigwedge, x\} \cup \operatorname{sym}(\varphi) & \textit{for } \pi = (\bigwedge, x, \varphi). \end{cases}$$

For $X \in \{L, (R,r), (F,r), T, I, V, C\}$ and $- \in \{t, \pi\}$ set $\operatorname{sym}_X(-) = \operatorname{sym}(-) \cap S_X$.

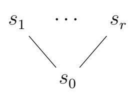
The free variables *of an S-formula π are the elements of*

$$\operatorname{fvar}(\pi) \;=\; \begin{cases} \operatorname{sym}_V(\pi) \cup \operatorname{sym}_T(\pi) & \textit{for } \pi \textit{ atomic}, \\ \operatorname{fvar}(\varphi) \cup \operatorname{fvar}(\psi) & \textit{for } \pi = (\to, \varphi, \psi), \\ \operatorname{fvar}(\varphi) \setminus \{x\} & \textit{for } \pi = (\bigwedge, x, \varphi). \end{cases}$$

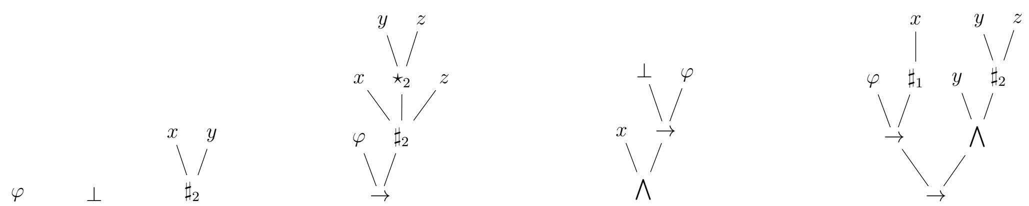
We call S-formulas in $\mathcal{F}_S^\bullet = \{\pi \in \mathcal{F}_S : \operatorname{fvar}(\pi) = \emptyset\}$ S-sentences *and S-terms in $\mathcal{T}_S^\bullet = \{t \in \mathcal{T}_S : \operatorname{sym}_V(t) = \emptyset\}$* variable-free S-terms.

A set of S-sentences is called an S-theory.

To improve readability we occasionally use for tuples $(s_0, s_1, \dots, s_r)$ the following alternative notation as tree:



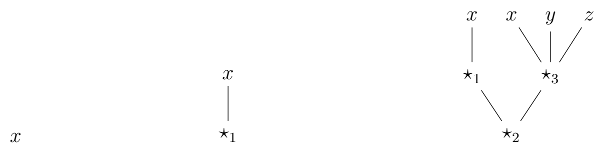
Example 1.2.4. With this notation the following are examples for S-formulas where $\varphi, \psi \in S_T$, $x, y, z \in S_V$, $\sharp_r \in S_{R,r}$, $\star_r \in S_{F,r}$:



$\varphi \qquad \bot \qquad x \sharp_2 y \qquad (\varphi \to \sharp_3 x (y \star_2 z) z) \qquad \bigwedge_x (\bot \to \varphi) \qquad ((\varphi \to \sharp_1 x) \to \bigwedge_y y \sharp_2 z)$

The S-formula π displayed rightmost has the set $\mathrm{sym}(\pi) = \{\to, \varphi, \sharp_1, x, \bigwedge, y, \sharp_2, z\}$ of symbols, $\mathrm{sym}_V(\pi) = \{x, y, z\}$ of variables and $\mathrm{fvar}(\pi) = \{\varphi, x, z\}$ of free variables.

Here are some examples of S-terms:



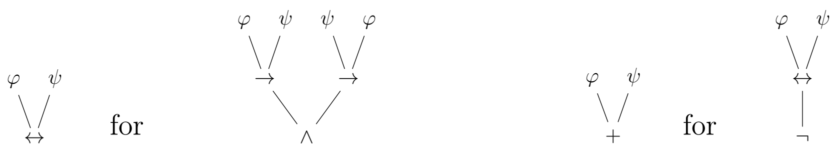
$$x \qquad \star_1 x \qquad (\star_1 x \ \star_2 \ \star_3 xyz)$$

The rightmost S-term t has $\mathrm{sym}(t) = \{\star_2, \star_1, x, \star_3, y, z\}$ and $\mathrm{sym}_V(t) = \{x, y, z\}$.

Given that the set $S_L = \{\bot, \to, \bigwedge\}$ of logical symbols at our disposal is very limited, we introduce the following abbreviations, where $\varphi, \psi \in \mathcal{F}_S$ stand for formulas, $x \in S_V$ for a variable, $s, t \in \mathcal{T}_S$ for terms and $\sim \in S_{R,2}$ for a binary relation symbol:

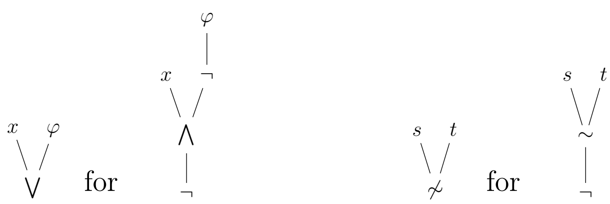
$$\neg\varphi \qquad (\varphi \to \bot) \qquad\qquad \top \qquad \neg\bot$$

$$(\varphi \wedge \psi) \qquad \neg(\varphi \to \neg\psi) \qquad\qquad (\varphi \vee \psi) \qquad (\neg\varphi \to \psi)$$



$$(\varphi \leftrightarrow \psi) \qquad ((\varphi \to \psi) \wedge (\psi \to \varphi)) \qquad\qquad (\varphi + \psi) \qquad \neg(\varphi \leftrightarrow \psi)$$

$$(\varphi \uparrow \psi) \qquad \neg(\varphi \wedge \psi) \qquad\qquad (\varphi \downarrow \psi) \qquad \neg(\varphi \vee \psi)$$



$$\bigvee_x \varphi \qquad \neg\bigwedge_x \neg\varphi \qquad\qquad s \not\sim t \qquad \neg\, s \sim t$$

Definition 1.2.5. *Let S and S' be two vocabularies with $S_V \subseteq S'_V$ and $S_{F,r} \subseteq S'_{F,r}$ and $S_{R,r} \subseteq S'_{R,r}$ for all $r > 0$.*

*An S-S'-*substitution *is a function $f\colon S_I \cup S_T \to \mathcal{T}_{S'} \cup \mathcal{F}_{S'}$ with $f(S_I) \subseteq \mathcal{T}_{S'}$ and $f(S_T) \subseteq \mathcal{F}_{S'}$. For such f and each S-term t we define the S'-term*

$$f_*(t) = \begin{cases} f(x) & \text{for } t = (x) \text{ with } x \in S_I, \\ (\sharp, f_*(t_1), \dots, f_*(t_r)) & \text{for } t = (\sharp, t_1, \dots, t_r) \text{ with } \sharp \in S_{F,r} \text{ and } r > 0, \end{cases}$$

and for each S-formula π the S'-formula

$$f_*(\pi) = \begin{cases} (\bot) & \text{for } \pi = (\bot), \\ f(\varphi) & \text{for } \pi = (\varphi) \text{ with } \varphi \in S_T, \\ (\sharp, f_*(t_1), \dots, f_*(t_r)) & \text{for } \pi = (\sharp, t_1, \dots, t_r) \text{ with } \sharp \in S_{R,r} \text{ with } r > 0, \\ (\to, f_*(\varphi), f_*(\psi)) & \text{for } \pi = (\to, \varphi, \psi), \\ (\bigwedge, x, f_*^x(\varphi)) & \text{for } \pi = (\bigwedge, x, \varphi), \end{cases}$$

where f^x denotes the S-S'-substitution with $f^x(x) = (x)$ and $f^x(v) = f(v)$ for $v \neq x$.

If there exists an S-S'-substitution f with $\{v \in S_I \cup S_T : f(v) \neq (v)\} \subseteq \{v_1, \dots, v_s\}$ and π is an S-formula, then we introduce the following alternative notation for $f_*(\pi)$:

$$\pi(v_1/f(v_1), \dots, v_s/f(v_s))$$

Example 1.2.6. Let $x \in S_V$, $w, y, z \in S_I$, $\varphi \in S_T$, $\equiv, \epsilon \in S_{R,2}$, $\sharp \in S_{F,1}$ and

$$\pi \;=\; (x \equiv y \to (\varphi \to \textstyle\bigvee_x x \equiv \sharp y))\,.$$

Assuming $x \neq y$, we get by substitution the S-formula

$$\pi(x/w, y/\sharp x, \varphi/w \;\epsilon\; z) = (w \equiv \sharp x \to (w \;\epsilon\; z \to \textstyle\bigvee_x x \equiv \sharp\sharp x))\,.$$

The following notion of safe substitutability will later be used in the substitution axioms of our formal proof system:

Definition 1.2.7. *An S-Term t is said to be* safely substitutable *for an individual symbol $y \in S_I$ in an S-formula π if*

$$\begin{array}{ll} & \pi \textit{ is atomic} \\ \textit{or} & \pi = (\to, \varphi, \psi) \textit{ and } t \textit{ is safely substitutable for } y \textit{ in } \varphi \textit{ and } \psi \\ \textit{or} & \pi = (\bigwedge, x, \varphi) \textit{ and} \\ & \quad \textit{either} \;\; y \notin \mathrm{fvar}(\pi) \\ & \quad\quad\;\; \textit{or} \;\; x \notin \mathrm{sym}_V(t) \textit{ and } t \textit{ is safely substitutable for } y \textit{ in } \varphi. \end{array}$$

Example 1.2.8. Let π be the formula

$$(\textstyle\bigwedge_y \bigvee_x x \not\equiv y \to \bigvee_x x \not\equiv y)$$

with $x \neq y$. Then x is not safely substitutable for y in π. Note that

$$\pi(y/x) \;=\; (\textstyle\bigwedge_y \bigvee_x x \not\equiv y \to \bigvee_x x \not\equiv x)\,.$$

1.3 Structures

So far formulas are nothing but compilations of symbols according to certain rules. To give them a meaning we want to interpret them in so-called structures:

Definition 1.3.1. *Let S be a vocabulary.*

A boolean-valued S-structure $\mathcal{M}$ *consists of*

$\underline{\mathcal{M}}$	*an* underlying *non-empty set,*
$R^{\mathcal{M}}$	*an* underlying *non-trivial boolean ring,*
$S_V^{\mathcal{M}}$	*a subset of S_V of* assigned variables*,*
$\sharp^{\mathcal{M}}$	*a function $\underline{\mathcal{M}}^r \to \underline{\mathcal{M}}$ for every $\sharp \in S_{F,r}^{\mathcal{M}} = S_{F,r} \setminus \big(S_V \setminus S_V^{\mathcal{M}}\big)$,*
$\sharp^{\mathcal{M}}$	*a function $\underline{\mathcal{M}}^r \to R^{\mathcal{M}}$ for every $\sharp \in S_{R,r}$,*

such that the definition of the value $\pi^{\mathcal{M}} \in R^{\mathcal{M}}$ for S-formulas π with $\mathrm{fvar}(\pi) \subseteq S_V^{\mathcal{M}}$ *given below makes sense, i.e. the boolean ring $R^{\mathcal{M}}$ must be complete w.r.t.*

$$\left\{\left\{\varphi^{\mathcal{M}_a^x} : a \in \underline{\mathcal{M}}\right\} : \pi = (\textstyle\bigwedge, x, \varphi) \text{ is an } S\text{-formula with } \mathrm{fvar}(\pi) \subseteq S_V^{\mathcal{M}}\right\}.$$

We call $\mathcal{M}$ fully assigned *or* unassigned *if $S_V^{\mathcal{M}} = S_V$ or $S_V^{\mathcal{M}} = \emptyset$, respectively.*

The value $t^{\mathcal{M}} \in \underline{\mathcal{M}}$ in $\mathcal{M}$ *of an S-term t with* $\mathrm{sym}_V(t) \subseteq S_V^{\mathcal{M}}$ *is defined as*

$$t^{\mathcal{M}} = \sharp^{\mathcal{M}}\big(t_1^{\mathcal{M}}, \ldots, t_r^{\mathcal{M}}\big) \quad \text{for } t = (\sharp, t_1, \ldots, t_r) \text{ with } \sharp \in S_{F,r}.$$

The value $\pi^{\mathcal{M}} \in R^{\mathcal{M}}$ in $\mathcal{M}$ *of an S-formula π with* $\mathrm{fvar}(\pi) \subseteq S_V^{\mathcal{M}}$ *is*

$$\pi^{\mathcal{M}} = \begin{cases} 0 & \text{for } \pi = (\bot), \\ \sharp^{\mathcal{M}}\big(t_1^{\mathcal{M}}, \ldots, t_r^{\mathcal{M}}\big) & \text{for } \pi = (\sharp, t_1, \ldots, t_r) \text{ with } \sharp \in S_{R,r}, \\ \varphi^{\mathcal{M}} \to \psi^{\mathcal{M}} & \text{for } \pi = (\to, \varphi, \psi), \\ \bigwedge_{a \in \underline{\mathcal{M}}} \varphi^{\mathcal{M}_a^x} & \text{for } \pi = (\bigwedge, x, \varphi), \end{cases}$$

where $\mathcal{M}_a^x$ is the boolean-valued S-structure with the same underlying set and boolean ring as $\mathcal{M}$, $S_V^{\mathcal{M}_a^x} = S_V^{\mathcal{M}} \cup \{x\}$ and with $x^{\mathcal{M}_a^x}(\emptyset) = a$ and $\sharp^{\mathcal{M}_a^x} = \sharp^{\mathcal{M}}$ for all $\sharp \neq x$.

We write $\mathcal{M} \vDash \pi$ synonymously for $\pi^{\mathcal{M}} = 1$. For sets of S-formulas Π with free variables in $S_V^{\mathcal{M}}$ we write $\mathcal{M} \vDash \Pi$ in case $\mathcal{M} \vDash \pi$ for all $\pi \in \Pi$.

Given that $\mathcal{M} \vDash \Pi$, we call $\mathcal{M}$ a boolean-valued model of Π.

We write $\mathcal{M} \vDash \pi[\vec{x}/\vec{a}]$ for tuples $\vec{x} = (x_1, \ldots, x_r)$ of pairwise distinct symbols in S_V and $\vec{a} = (a_1, \ldots, a_r) \in \underline{\mathcal{M}}^r$ in case $\mathcal{M}_{\vec{a}}^{\vec{x}} = (\cdots(\mathcal{M}_{a_1}^{x_1})\cdots)_{a_r}^{x_r} \vDash \pi$.

We say that $\mathcal{M}$ has witnesses *if for all S-formulas π of the form $\bigwedge_x \varphi$ and tuples $\vec{x} = (x_1, \ldots, x_r)$ of pairwise distinct symbols in $S_V \setminus \{x\}$ and $\vec{a} = (a_1, \ldots, a_r) \in \underline{\mathcal{M}}^r$ there exists $a \in \underline{\mathcal{M}}$ such that $\pi^{\mathcal{M}_{\vec{a}}^{\vec{x}}} = \varphi^{\mathcal{M}_{(\vec{a},a)}^{(\vec{x},x)}}$.*

*Boolean-valued S-structures $\mathcal{M}$ with $R^{\mathcal{M}} = \mathbb{F}_2$ are called S-*structures*. Accordingly, S-structures $\mathcal{M}$ with $\mathcal{M} \vDash \Pi$ are called* models *of* Π.

For (sets of) S-formulas Π, Π' we write $\Pi \vDash \Pi'$ (resp. $\Pi \vDash' \Pi'$) if every fully assigned model (resp. boolean-valued model) $\mathcal{M}$ of Π is a model (resp. boolean-valued model) of Π', too. We call Π satisfiable *(resp.* boolean-valued satisfiable*) if there is a model (resp. boolean-valued model) of Π. Finally, Π is said to be* tautological *(resp.* boolean-valued tautological*) if $\emptyset \vDash \Pi$ (resp. $\emptyset \vDash' \Pi$).*

S-formulas π and σ are semantically equivalent *(resp.* boolean-valued semantically equivalent*) if $(\pi \leftrightarrow \sigma)$ is tautological (resp. boolean-valued tautological).*

An r-ary relation \$ on $\underline{\mathcal{M}}$ is definable in $\mathcal{M}$ *if there are finitely many elements $b_1, \ldots, b_s \in \underline{\mathcal{M}}$ and an S-formula π with $\operatorname{fvar}(\pi) \subseteq \{x_1, \ldots, x_r, y_1, \ldots, y_s\}$ such that:*

$$\$ = \Big\{(a_1, \ldots, a_r) \in \underline{\mathcal{M}}^r : \mathcal{M} \vDash \pi\big[(x_1, \ldots, x_r, y_1, \ldots, y_s)/(a_1, \ldots, a_r, b_1, \ldots, b_s)\big]\Big\}$$

Remark 1.3.2. For every set M the r-ary relations $\sharp$ on M correspond to functions $\hat{\sharp}\colon M^r \to \mathbb{F}_2$ via $\hat{\sharp}(x) = 1 \Leftrightarrow x \in \sharp$. We will say $\hat{\sharp}$ is *given by* $\sharp$, and vice versa.

With this terminology, for S-structures $\mathcal{M}$ and every $\sharp \in S_{R,r}$ the function $\sharp^{\mathcal{M}}$ is given by an r-ary relation on $\underline{\mathcal{M}}$.

Lemma 1.3.3. *Every S-structure has witnesses.*

Proof. This holds because for every non-empty family $(x_i)_{i \in I}$ of elements in $\mathbb{F}_2$ there exists $j \in I$ with $\bigwedge_{i \in I} x_i = x_j$. □

Example 1.3.4. The S-formula $\bot$ is (boolean-valued) unsatisfiable and $\top$ (boolean-valued) tautological. The S-formulas φ, $\neg\neg\varphi$, $(\varphi \vee \bot)$, $(\top \to \varphi)$ are (boolean-valued) semantically equivalent.

Remark 1.3.5. In every boolean-valued S-structure $\mathcal{M}$, for $\diamond \in \{\vee, \wedge, \leftrightarrow, +, \uparrow, \downarrow\}$ and S-formulas φ, ψ with $\operatorname{fvar}(\varphi) \cup \operatorname{fvar}(\psi) \subseteq S_V^{\mathcal{M}}$, we have:

$$\begin{aligned} (\neg\varphi)^{\mathcal{M}} &= \neg\varphi^{\mathcal{M}} \\ (\varphi \diamond \psi)^{\mathcal{M}} &= \varphi^{\mathcal{M}} \diamond \psi^{\mathcal{M}} \\ (\textstyle\bigvee_x \varphi)^{\mathcal{M}} &= \textstyle\bigvee_{a \in \underline{\mathcal{M}}} \varphi^{\mathcal{M}^x_a} \end{aligned}$$

This easily follows with the calculation rules for boolean rings, e.g.:

$$(\neg\varphi)^{\mathcal{M}} = (\varphi \to \bot)^{\mathcal{M}} = \varphi^{\mathcal{M}} \to \bot^{\mathcal{M}} = \neg\varphi^{\mathcal{M}} \vee 0 = \neg\varphi^{\mathcal{M}}$$
$$(\varphi \vee \psi)^{\mathcal{M}} = (\neg\varphi \to \psi)^{\mathcal{M}} = (\neg\varphi)^{\mathcal{M}} \to \psi^{\mathcal{M}} = \neg\varphi^{\mathcal{M}} \to \psi^{\mathcal{M}} = \varphi^{\mathcal{M}} \vee \psi^{\mathcal{M}}$$
$$(\textstyle\bigvee_x \varphi)^{\mathcal{M}} = (\neg\textstyle\bigwedge_x \neg\varphi)^{\mathcal{M}} = \neg \textstyle\bigwedge_{a \in \underline{\mathcal{M}}} \neg\varphi^{\mathcal{M}^x_a} = \textstyle\bigvee_{a \in \underline{\mathcal{M}}} \varphi^{\mathcal{M}^x_a}$$

Remark 1.3.6. S-formulas φ and ψ are (boolean-valued) semantically equivalent iff $\varphi^{\mathcal{M}} = \psi^{\mathcal{M}}$ for every (boolean-valued) S-structure $\mathcal{M}$ by Exercise 1.1.5 (viii).

Lemma 1.3.7. *The set R_S of the equivalence classes of S-formulas under semantic equivalence is a boolean ring under the operations*

$$\neg[\varphi] = [\neg\varphi] \quad \textit{and} \quad [\varphi] \diamond [\psi] = [\varphi \diamond \psi] \quad \textit{for} \quad \diamond \in \{\wedge, \vee, \to, \leftrightarrow, +, \uparrow, \downarrow\}.$$

Proof. The well-definedness follows from Remarks 1.3.5 and 1.3.6. $\square$

Remark 1.3.8. More generally, let Π be a set of S-formulas and let R_Π be the set of the equivalence classes of S-formulas under the equivalence relation that identifies φ and ψ iff $\Pi \vDash (\varphi \leftrightarrow \psi)$. Similarly as in Lemma 1.3.7, R_Π naturally becomes a boolean algebra, the so-called *Lindenbaum-Tarski algebra* of Π.

Lemma 1.3.9. *Let $\mathcal{M}$ be a boolean-valued S-structure and $\mathfrak{m}$ a maximal ideal in $R^{\mathcal{M}}$. If $\mathcal{M}$ has witnesses, then the S-structure $\mathcal{M}/\mathfrak{m}$ with $\underline{\mathcal{M}/\mathfrak{m}} = \underline{\mathcal{M}}$ and*

$$\begin{aligned} \sharp^{\mathcal{M}/\mathfrak{m}}(m_1,\dots,m_r) = 1 &\Leftrightarrow \sharp^{\mathcal{M}}(m_1,\dots,m_r) \notin \mathfrak{m} && \text{for } \sharp \in S_{R,r}, \\ \sharp^{\mathcal{M}/\mathfrak{m}}(m_1,\dots,m_r) &= \sharp^{\mathcal{M}}(m_1,\dots,m_r) && \text{for } \sharp \in S_{F,r}^{\mathcal{M}/\mathfrak{m}} = S_{F,r}^{\mathcal{M}} \end{aligned}$$

satisfies $\mathcal{M}/\mathfrak{m} \vDash \pi \Leftrightarrow \pi^{\mathcal{M}} \notin \mathfrak{m}$ for all S-formulas π with $\mathrm{fvar}(\pi) \subseteq S_V^{\mathcal{M}/\mathfrak{m}} = S_V^{\mathcal{M}}$.

Proof. Note that $\mathcal{R}^{\mathcal{M}}/\mathfrak{m} \cong \mathbb{F}_2$ canonically. We will show $\pi^{(\mathcal{M}/\mathfrak{m})^{\vec{x}}_{\vec{a}}} = 1 \Leftrightarrow \pi^{\mathcal{M}^{\vec{x}}_{\vec{a}}} \notin \mathfrak{m}$ for all S-formulas π, tuples $\vec{x} = (x_1,\dots,x_n)$ of distinct variable symbols in S_V and $\vec{a} = (a_1,\dots,a_n)$ of elements in $\underline{\mathcal{M}}$. To do this, we use structural induction on π.

In case $\pi = \perp$ or $\pi = \sharp t_1 \cdots t_r$ or $\pi = (\varphi \to \psi)$ the claim is obvious by induction and because the canonical map $R \to R^{\mathcal{M}}/\mathfrak{m} \to \mathbb{F}_2$ is a morphism of boolean rings.

It remains to consider the case $\pi = \bigwedge_x \varphi$, where we may assume $x_j \neq x$ for all j. Since $\mathcal{M}$ has witnesses, there exists $b \in \underline{\mathcal{M}}$ such that

$$\begin{aligned} \pi^{(\mathcal{M}/\mathfrak{m})^{\vec{x}}_{\vec{a}}} = 1 &\Leftrightarrow \varphi^{(\mathcal{M}/\mathfrak{m})^{(\vec{x},x)}_{(\vec{a},a)}} = 1 \text{ for all } a \in \underline{\mathcal{M}} \\ &\Leftrightarrow \varphi^{\mathcal{M}^{(\vec{x},x)}_{(\vec{a},a)}} \notin \mathfrak{m} \quad \text{for all } a \in \underline{\mathcal{M}} \Leftrightarrow \pi^{\mathcal{M}^{\vec{x}}_{\vec{a}}} = \varphi^{\mathcal{M}^{(\vec{x},x)}_{(\vec{a},b)}} \notin \mathfrak{m}\,. \end{aligned}$$

$\square$

The following result will be proved later as a consequence of Gödel's Completeness Theorem. It also has a short direct proof (see Problem Set 2, Exercise 1):

Theorem 1.3.10 (Compactness Theorem)**.** *An S-theory T is satisfiable if and only if every finite subset of T is satisfiable.*

Another corollary of the Completeness Theorem is the following:

Theorem 1.3.11. *$T \vDash \varphi \Leftrightarrow T \vDash' \varphi$ for every S-theory T and S-formula φ.*

In particular, S-formulas are tautological iff they are boolean-valued tautological.

Propositional logic

Definition 1.3.12. *Fix a vocabulary $S = S^{\mathrm{PL}}$ consisting of a set S_T of countably many truth symbols and with $S_{R,r} = \emptyset$ for all $r \neq 0$ and $S_{F,r} = \emptyset$ for all r.*

Remark 1.3.13. An S^{PL}-structure $\mathcal{M}$ is determined by a set $\underline{\mathcal{M}}$ together with the function $S_T^{\mathrm{PL}} \to \mathbb{F}_2$ given by the rule $\varphi \mapsto \mathcal{M}(\varphi) = \varphi^{\mathcal{M}}(\emptyset)$.

The formal language of propositional logic is *functionally complete* in the sense that every *boolean function* $\mathbb{F}_2^n \to \mathbb{F}_2$ is given by an appropriate S^{PL}-formula:

Theorem 1.3.14 (Functional completeness)**.** *For every function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ with $n \in \mathbb{N}$ there exists an S^{PL}-formula π with an n-element set $\mathrm{sym}_T(\pi) = \{\varphi_1, \dots, \varphi_n\}$ of truth symbols such that for all S^{PL}-structures $\mathcal{M}$*

$$\pi^{\mathcal{M}} \;=\; f(\mathcal{M}(\varphi_1), \dots, \mathcal{M}(\varphi_n))\,.$$

Proof. By induction on its support every function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ is seen to be polynomial, i.e. there exist $\lambda_{i_1,\dots,i_k} \in \mathbb{F}_2$ with $f(\sum_i a_i e_i) = \sum_k \sum_{1\le i_1 \le \dots \le i_k \le n} \lambda_{i_1,\dots,i_k} a_{i_1} \cdots a_{i_k}$. Choose $\pi = \sum_k \sum_{\lambda_{i_1,\dots,i_k} \neq 0} \bigwedge_j \varphi_{i_j}$, where the "sums" are taken in some fixed order. □

Proof (alternative). Let $\{x_1, \dots, x_m\} = f^{-1}(1)$. Take for $\pi = \bigvee_{i=1}^m \bigwedge_{j=1}^n \varphi_j^{x_{ij}}$, where φ^0 has to be read as $\neg\varphi$ and φ^1 as φ and $\bigvee_\emptyset$ as $\bot$ and $\bigwedge_\emptyset$ as $\top$. □

Remark 1.3.15. The first proof of Theorem 1.3.14 indicates how one can show that $\{\top, +, \wedge\}$ is a so-called functionally complete set of logical connectives, while the second one indicates how this can be done for $\{\neg, \wedge\}$ and $\{\neg, \vee\}$.

Definition 1.3.16. *An S-formula π is called* propositionally tautological *if there is a tautological S^{PL}-formula φ and an S^{PL}-S-substitution f with $\pi = f_*(\varphi)$.*

Example 1.3.17. $(\bigwedge_x \varphi \vee \neg \bigwedge_x \varphi)$ with $x \in S_V$, $\varphi \in \mathcal{F}_S$ is propositionally tautological.

Remark 1.3.18. Whether a given formula is propositionally tautological can algorithmically be verified by the method of truth tables or with other SAT solvers.

1.4 Theories

Recall that an S-theory is by definition just a set of S-sentences.

Definition 1.4.1. *An S-*theory with equality $\equiv$ *is an S-theory over a vocabulary with $\equiv \in S_{R,2}$ that contains at least the following S-sentences:*

(R) *For all $x \in S_V$ an* axiom of reflexivity

$$\textstyle\bigwedge_x x \equiv x\,.$$

(S) *For all $x, y \in S_V$ an* axiom of symmetry

$$\textstyle\bigwedge_x \bigwedge_y (x \equiv y \to y \equiv x)\,.$$

(T) *For all $x, y, z \in S_V$ an* axiom of transitivity

$$\textstyle\bigwedge_x \bigwedge_y \bigwedge_z (x \equiv y \to (y \equiv z \to x \equiv z))\,.$$

(C) *For all* $r \in \mathbb{N}$, $x_1, y_1, \dots, x_r, y_r \in S_V$ *and* $(\sharp, \diamond) \in (S_{R,r} \times \{\to\}) \cup (S_{F,r} \times \{\equiv\})$ *an* axiom of congruence

$$\bigwedge_{x_1}\bigwedge_{y_1}\cdots\bigwedge_{x_r}\bigwedge_{y_r}(x_1 \equiv y_1 \to (\cdots \to (x_r \equiv y_r \to (\sharp x_1 \cdots x_r \diamond \sharp y_1 \cdots y_r))\cdots)).$$

A boolean-valued S*-structure* $\mathcal{M}$ *respects* $\equiv$ *if* $\equiv^{\mathcal{M}}(m_1, m_2) = 1 \Leftrightarrow m_1 = m_2$.

Lemma 1.4.2. *Let* $\mathcal{M}$ *be a (boolean-valued) model of an* S*-theory* T *with equality* $\equiv$*. The (boolean-valued)* S*-structure* $\mathcal{M}/_{\equiv}$ *with* $R^{\mathcal{M}/_{\equiv}} = R^{\mathcal{M}}$ *and* $\underline{\mathcal{M}/_{\equiv}} = \underline{\mathcal{M}}/_{\sim}$ *where* $\sim$ *is given by* $m_1 \sim m_2 \Leftrightarrow \equiv^{\mathcal{M}}(m_1, m_2) = 1$ *and with*

$$\begin{aligned} \sharp^{\mathcal{M}/_{\equiv}}([m_1], \dots, [m_r]) &= \sharp^{\mathcal{M}}(m_1, \dots, m_r) && \text{for } \sharp \in S_{R,r}, \\ \sharp^{\mathcal{M}/_{\equiv}}([m_1], \dots, [m_r]) &= [\sharp^{\mathcal{M}}(m_1, \dots, m_r)] && \text{for } \sharp \in S_{F,r}^{\mathcal{M}/_{\equiv}} = S_{F,r}^{\mathcal{M}} \end{aligned}$$

respects $\equiv$ *and satisfies* $\pi^{\mathcal{M}/_{\equiv}} = \pi^{\mathcal{M}}$ *for all* S*-formulas* π *with* $\mathrm{fvar}(\pi) \subseteq S_V^{\mathcal{M}/_{\equiv}} = S_V^{\mathcal{M}}$*.*

In particular, $\mathcal{M}/_{\equiv}$ *is a* $\equiv$*-respecting (boolean-valued) model of* T*.*

Proof. To see that $\sim$ is an equivalence relation use that $\mathcal{M}$ satisfies (R), (S), (T) and for the well-definedness of $\sharp^{\mathcal{M}/_{\equiv}}$ use that $\mathcal{M}$ satisfies (C). By structural induction we obtain the identities $\pi^{\mathcal{M}/_{\equiv}} = \pi^{\mathcal{M}}$. □

Example: Commutative ring theory

Definition 1.4.3. *Fix a vocabulary* $S = S^{\mathrm{Ring}}$ *consisting of sets*

$S_{R,2}$	$=$	$\{\equiv\}$	*of an equality symbol,*
S_V			*of countably many variable symbols,*
S_C	$=$	$\{\mathtt{0}, \mathtt{1}\}$	*of two constant symbols,*
$S_{F,2}$	$=$	$\{\oplus, \odot\}$	*of two binary function symbols,*

and with $S_{R,r} = \emptyset$ *for all* $r \neq 2$ *and* $S_{F,r} = \emptyset$ *for all* $r \notin \{0, 2\}$*.*

Definition 1.4.4. *Let* $S = S^{\mathrm{Ring}}$*.* CRT *("commutative ring theory") is the minimal* S*-theory with equality* $\equiv$ *that contains the following* S*-sentences:*

(1) For all distinct $x, y, z \in S_V$ *and* $\diamond \in \{\oplus, \odot\}$ *an* axiom of associativity

$$\bigwedge_x\bigwedge_y\bigwedge_z((x \diamond y) \diamond z) \equiv (x \diamond (y \diamond z)).$$

(2) For all distinct $x, y \in S_V$ *and* $\diamond \in \{\oplus, \odot\}$ *an* axiom of commutativity

$$\bigwedge_x\bigwedge_y(x \diamond y) \equiv (y \diamond x).$$

(3) For all $x \in S_V$ *and* $(e, \diamond) \in \{(\mathtt{0}, \oplus), (\mathtt{1}, \odot)\}$ *an* axiom of identity

$$\bigwedge_x(e \diamond x) \equiv x.$$

(4) For all distinct $x, y \in S_V$ *an* axiom of invertibility

$$\bigwedge_x\bigvee_y(x \oplus y) \equiv \mathtt{0}.$$

(5) For all distinct $x, y, z \in S_V$ *an* axiom of distributivity

$$\bigwedge_x\bigwedge_y\bigwedge_z(x \odot (y \oplus z)) \equiv ((x \odot y) \oplus (x \odot z)).$$

Example: Zermelo–Fraenkel set theory

Definition 1.4.5. *Fix a vocabulary* $S = S^{\mathrm{Set}}$ *consisting of sets*

$$\begin{array}{lcll} S_{R,2} & = & \{\equiv, \epsilon\} & \textit{of two binary relation symbols,} \\ S_V & & & \textit{of countably many variable symbols,} \end{array}$$

and with $S_{R,r} = \emptyset$ *for all* $r \neq 2$ *and* $S_C = \emptyset$ *and* $S_{F,r} = \emptyset$ *for all* $r \neq 0$.

Definition 1.4.6. *Let* $S = S^{\mathrm{Set}}$. ZFC *is the minimal* S*-theory with equality* $\equiv$ *that contains the following* S*-sentences:*

(`EXT`) *For all distinct* $x, y, v \in S_V$ *an* axiom of extensionality

$$\bigwedge_x \bigwedge_y (\bigwedge_v (v \,\epsilon\, x \leftrightarrow v \,\epsilon\, y) \rightarrow x \equiv y)\,.$$

(`EMP`) *For all distinct* $x, v \in S_V$ *an* axiom of empty set

$$\bigvee_x \bigwedge_v v \not\epsilon\, x\,.$$

(`PAI`) *For all distinct* $x, y, z, v \in S_V$ *an* axiom of pairing

$$\bigwedge_x \bigwedge_y \bigvee_z \bigwedge_v ((v \equiv x \vee v \equiv y) \leftrightarrow v \,\epsilon\, z)\,.$$

(`UNI`) *For all distinct* $x, y, v, w \in S_V$ *an* axiom of union

$$\bigwedge_x \bigvee_y \bigwedge_v (\bigvee_w (v \,\epsilon\, w \wedge w \,\epsilon\, x) \leftrightarrow v \,\epsilon\, y)\,.$$

(`REP`) *For all* $n \in \mathbb{N}$ *and distinct* $x, y, u, v, \hat{v}, w_1, \ldots, w_n \in S_V$ *and each* S*-formula* φ *with* $\mathrm{fvar}(\varphi) \subseteq \{x, u, v, w_1, \ldots, w_n\}$ *an* axiom of replacement

$$\begin{aligned} \bigwedge_{w_1} \cdots \bigwedge_{w_n} \bigwedge_x (\bigwedge_u (u \,\epsilon\, x \rightarrow \bigvee_{\hat{v}} \bigwedge_v (\varphi \rightarrow v \equiv \hat{v})) \rightarrow \\ \bigvee_y \bigwedge_v (v \,\epsilon\, y \leftrightarrow \bigvee_u (u \,\epsilon\, x \wedge \varphi)))\,. \end{aligned}$$

(`POW`) *For all distinct* $x, y, u, v \in S_V$ *an* axiom of power set

$$\bigwedge_x \bigvee_y \bigwedge_v (v \subseteq x \leftrightarrow v \,\epsilon\, y)$$

with short hand

$$v \subseteq x \quad \textit{for} \quad \bigwedge_u (u \,\epsilon\, v \rightarrow u \,\epsilon\, x).$$

(`INF`) *For all distinct* $x, y, u, v, w \in S_V$ *an* axiom of infinity

$$\bigvee_x (\varnothing \,\epsilon\, x \wedge \bigwedge_v (v \,\epsilon\, x \rightarrow v \cup \{v\} \,\epsilon\, x))$$

with short hand

$$\begin{array}{rcl} v \cup \{v\} \,\epsilon\, x & \textit{for} & \bigvee_w (w \,\epsilon\, x \wedge \bigwedge_u (u \,\epsilon\, w \leftrightarrow (u \,\epsilon\, v \vee u \equiv v))), \\ \varnothing \,\epsilon\, x & \textit{for} & \bigvee_v (v \,\epsilon\, x \wedge \bigwedge_u u \not\epsilon\, v). \end{array}$$

(CHO) *For all distinct $x, y, \hat{u}, u, v, w \in S_V$ an* axiom of choice

$$\bigwedge_x((\varnothing \not\epsilon\, x \wedge \bigwedge_{v \not\equiv w \epsilon x} v \cap w \equiv \varnothing) \to \\ \bigvee_y \bigwedge_v (v \,\epsilon\, x \to \bigvee^!_u (u \,\epsilon\, v \wedge u \,\epsilon\, y)))$$

with short hand

$$\begin{aligned} \varnothing \not\epsilon\, x \quad &\text{for} \quad \bigwedge_v (v \,\epsilon\, x \to \bigvee_u u \,\epsilon\, v), \\ \bigwedge_{v \not\equiv w \epsilon x} v \cap w \equiv \varnothing \quad &\text{for} \quad \bigwedge_v \bigwedge_w ((v \,\epsilon\, x \wedge w \,\epsilon\, x) \to \\ &\qquad (v \equiv w \vee \bigwedge_u (u \,\epsilon\, v \to u \not\epsilon\, w))), \\ \bigvee^!_u \varphi \quad &\text{for} \quad \bigvee_{\hat{u}} \bigwedge_u (\varphi \leftrightarrow u \equiv \hat{u}) \quad \text{where} \quad \varphi = (u \,\epsilon\, v \wedge u \,\epsilon\, y). \end{aligned}$$

(REG) *For all distinct $x, u, v \in S_V$ an* axiom of regularity

$$\bigwedge_x (x \not\equiv \varnothing \to \bigvee_v (v \,\epsilon\, x \wedge x \cap v \equiv \varnothing))$$

with short hand

$$\begin{aligned} x \not\equiv \varnothing \quad &\text{for} \quad \bigvee_v v \,\epsilon\, x, \\ x \cap v \equiv \varnothing \quad &\text{for} \quad \bigwedge_u (u \,\epsilon\, x \to u \not\epsilon\, v). \end{aligned}$$

We will denote by XXX *the minimal S-theory with equality $\equiv$ that contains all the axioms of type* XXX*, so*

$$\mathrm{ZFC} = \mathtt{EXT} \cup \mathtt{EMP} \cup \mathtt{PAI} \cup \mathtt{UNI} \cup \mathtt{REP} \cup \mathtt{POW} \cup \mathtt{INF} \cup \mathtt{CHO} \cup \mathtt{REG}.$$

Example: Peano arithmetic

Definition 1.4.7. *Fix a vocabulary $S = S^{\text{Peano}}$ consisting of sets*

$S_{R,2}$	$=$	$\{\equiv\}$	*of an equality symbol,*
S_V			*of countably many variable symbols,*
S_C	$=$	$\{\mathtt{0}\}$	*of a* zero symbol,
$S_{F,1}$	$=$	$\{\mathtt{S}\}$	*of a* successor symbol,

and with $S_{R,r} = \emptyset$ for all $r \neq 2$ and $S_{F,r} = \emptyset$ for all $r > 1$.

Definition 1.4.8. *Let $S = S^{\text{Peano}}$.* PA *("Peano arithmetic") is the minimal S-theory with equality $\equiv$ that contains the following S-sentences:*

(1) For all $x \in S_V$ an axiom of non-circularity

$$\bigwedge_x \mathtt{S}x \not\equiv \mathtt{0}.$$

(2) For all distinct $x, y \in S_V$ an axiom of unique successor

$$\bigwedge_x \bigwedge_y (\mathtt{S}x \equiv \mathtt{S}y \to x \equiv y).$$

(3) For all $n \in \mathbb{N}$ and all distinct $x, w_1, \ldots, w_n \in S_V$ and each S-formula φ with $\mathrm{fvar}(\varphi) \subseteq \{x, w_1, \ldots, w_n\}$ *an* axiom of induction

$$\bigwedge_{w_1} \cdots \bigwedge_{w_n} ((\varphi(x/\mathtt{0}) \wedge \bigwedge_x (\varphi \to \varphi(x/Sx))) \to \bigwedge_x \varphi).$$

Example 1.4.9. The natural numbers can be regarded as a model $\check{\mathbb{N}}$ of PA, the so-called *standard model*, with $\underline{\check{\mathbb{N}}} = \mathbb{N}$, $\mathtt{0}^{\check{\mathbb{N}}} = 0$, and $\mathtt{S}^{\check{\mathbb{N}}}(n) = n + 1$ for all $n \in \mathbb{N}$.

1.5 Formal proofs

One common approach to formalize the concept of a mathematical proof is to devise rules specifying how formulas can be manipulated. After selecting several tautological formulas as axioms, these rules are then used to infer new sentences from the axioms and from the sentences of a given theory.

The rules and axioms should be *sound* in the sense that all inferred sentences are satisfied in all models of the theory. Vice versa, *completeness* is desirable in the sense that it be possible to infer all sentences that are satisfied in all models of the theory.

The proof system below – a Hilbert-style deduction system – will turn out to be both sound and complete. It was chosen to resemble the Metamath Proof Explorer:

Definition 1.5.1. *An S-*rule (of inference) *is a pair* $(\mathrm{Pre}, \mathrm{Con})$ *of sets of S-formulas. In case* $\mathrm{Pre} = \{\varphi_1, \ldots, \varphi_s\}$ *and* $\mathrm{Con} = \{\psi_1, \ldots, \psi_t\}$ *we write this rule as*

$$\frac{\varphi_1, \ldots, \varphi_s}{\psi_1, \ldots, \psi_t}.$$

*The set $\mathcal{R}_S$ of S-*rules of predicate calculus *consists of the following S-rules:*

(`MP`) *For all $\varphi, \psi \in \mathcal{F}_S$ a rule of modus ponens:*

$$\frac{\varphi, (\varphi \to \psi)}{\psi}$$

(`GE`) *For all $\varphi \in \mathcal{F}_S$ and $x \in S_V$ a rule of generalization:*

$$\frac{\varphi}{\bigwedge_x \varphi}$$

*The set $\mathcal{A}_S$ of S-*axioms of predicate calculus *consists of the following S-formulas:*

($\Box$) *For every propositionally tautological $\varphi \in \mathcal{F}_S$ an* axiom of tautology*:*

$$\varphi$$

($\bigwedge$) *For all $x \in S_V$, $\varphi \in \mathcal{F}_S$ with $x \notin \mathrm{fvar}(\varphi)$ an axiom of universality:*

$$(\varphi \to \bigwedge_x \varphi)$$

(Q) *For all $x \in S_V$, $\varphi, \psi \in \mathcal{F}_S$ an axiom of quantified implication:*

$$(\bigwedge_x(\varphi \to \psi) \to (\bigwedge_x \varphi \to \bigwedge_x \psi))$$

($/$) *For all $\varphi \in \mathcal{F}_S$, $x \in S_V$, $t \in \mathcal{T}_S$ with the property that t is safely substitutable for x in φ an* axiom of substitution*:*

$$(\bigwedge_x \varphi \to \varphi(x/t))$$

For every S-theory T we recursively define the set $T_\infty = \bigcup_{n\in\mathbb{N}} T_n$ of S-formulas provable from T *by*

$$\begin{aligned} T_0 &= \mathcal{A}_S \cup T\,, \\ T_n &= T_{n-1} \cup \bigcup\{\text{Con} : (\text{Pre}, \text{Con}) \in \mathcal{R}_S \text{ with } \text{Pre} \subseteq T_{n-1}\} \quad \textit{for } n > 0\,. \end{aligned}$$

Instead of $\varphi \in T_\infty$ we usually write $T \vdash \varphi$. Whenever it seems necessary to stress the dependence on S, we will use the notation $\vdash_S$ for $\vdash$.

*A T-*proof *of φ is a finite sequence $(\varphi_1, \ldots, \varphi_{s-1}, \varphi_s = \varphi)$ of S-formulas such that for every $k \in \{1, \ldots, s\}$ either $\varphi_k \in T_0$ or there exist $i, j \in \{1, \ldots, k-1\}$ with $\varphi_j = (\varphi_i \to \varphi_k)$ or there exists $i \in \{1, \ldots, k-1\}$ and $x \in S_V$ with $\varphi_k = \bigwedge_x \varphi_i$.*

Using induction it is evident that an S-formula has a T-proof if and only if it is provable from T. Provable formulas are semantically correct in the following sense:

Lemma 1.5.2 (Soundness)**.** *For every S-theory and every S-formula φ:*

$$T \vdash \varphi \;\Rightarrow\; T \vDash' \varphi$$

Proof. This follows from $T \vDash' \text{Pre} \Rightarrow T \vDash' \text{Con}$ for all $(\text{Pre}, \text{Con}) \in \mathcal{R}_S$ together with the fact that all the S-formulas in ($\square$), ($\bigwedge$), (Q), (/) are tautological. $\square$

Definition 1.5.3. *A vocabulary $\overline{S}$ is an* extension by constants *of the vocabulary S if $S_C \subseteq \overline{S}_C$, $S_V = \overline{S}_V$, $S_{R,r} = \overline{S}_{R,r}$ for all r and $S_{F,r} = \overline{S}_{F,r}$ for all $r > 0$.*

The next lemma collects the important properties of the provability relation $\vdash$.

Lemma 1.5.4. *For all S-theories $T, \overline{T}$ and $\varphi, \psi \in \mathcal{F}_S$, $x \in S_V$ and every extension by constants $\overline{S}$ of S we have the following properties for $\vdash = \vdash_S$:*

- (`AX`) *If $\varphi \in \mathcal{A}_S$ or $\varphi \in T$, then $T \vdash \varphi$.*
- (`FM`) *If $T \vdash \varphi$ and $\overline{T} \supseteq T$, then $\overline{T} \vdash \varphi$.*
- (`MP`) *If $T \vdash \varphi$ and $T \vdash (\varphi \to \psi)$, then $T \vdash \psi$.*
- (`GE`) *If $T \vdash \varphi$, then $T \vdash \bigwedge_x \varphi$.*
- (`CP`) *If $T \vdash \varphi$, then there is a finite subset $\underline{T}$ of T with $\underline{T} \vdash \varphi$.*
- (`FC`) *If $T \vdash \varphi$, then $T \vdash_{\overline{S}} \varphi$.*

Proof. (`AX`), (`FM`), (`MP`), (`GE`), (`FC`) are clear.

To prove (`CP`), assume $T \vdash \varphi$. Then we can choose a T-proof $(\varphi_1, \ldots, \varphi_s)$ of φ and observe that $\underline{T} \vdash \varphi$ with $\underline{T} = \{\pi \in T : \pi = \varphi_k \text{ for some } k \in \{1, \ldots, s\}\}$. $\square$

Lemma 1.5.5 (Generalization and specialization)**.** *For every S-theory T, all $\varphi \in \mathcal{F}_S$ and all $x, y \in S_V$ with $y \notin \text{fvar}(\bigwedge_x \varphi)$:*

$$T \vdash \varphi(x/y) \;\Leftrightarrow\; T \vdash \bigwedge\nolimits_x \varphi$$

Moreover, $T \vdash \bigwedge_x \varphi \Rightarrow T \vdash \varphi(x/t)$ whenever $t \in \mathcal{T}_S$ is safely substitutable for x in φ.

Proof. Apply $(\mathtt{AX})$ with $(\bigwedge_x \varphi \to \varphi(x/t)) \in \mathcal{A}_S$ and $(\mathtt{MP})$ to obtain the last statement. The implication $\Leftarrow$ in the first statement is the special case $t = y$.

For $\Rightarrow$ we get $T \vdash \bigwedge_y \varphi(x/y)$ from $T \vdash \varphi(x/y)$ with $(\mathtt{GE})$, so $T \vdash \varphi(x/y)(y/x) = \varphi$ with the already shown $\Leftarrow$, and then finally $T \vdash \bigwedge_x \varphi$ using $(\mathtt{GE})$ again. $\square$

Lemma 1.5.6. *Given an extension by constants $\overline{S}$ of S with infinitely many variable symbols, we have for every S-theory T, all $\varphi \in \mathcal{F}_{\overline{S}}$ with m-element set $\mathrm{sym}_C(\varphi) \setminus S_C = \{c_1, \ldots, c_m\}$ of constant symbols and every m-element set $\{x_1, \ldots, x_m\} \subseteq S_V \setminus \mathrm{fvar}(\varphi)$:*

$$T \vdash_{\overline{S}} \varphi \;\Leftrightarrow\; T \vdash_S \bigwedge_{x_1} \cdots \bigwedge_{x_m} \varphi(c_1/x_1, \ldots, c_m/x_m)$$

Proof. $\Leftarrow$ holds due to $(\mathtt{FC})$ and Lemma 1.5.5.

To verify $\Rightarrow$ let $(\varphi_1, \ldots, \varphi_s)$ be a T-proof of φ over the vocabulary $\overline{S}$ and let $\{c_1, \ldots, c_n\} = (\mathrm{sym}_C(\varphi_1) \cup \cdots \cup \mathrm{sym}_C(\varphi_s)) \setminus S_C$ be an n-element set. Given that S_V is infinite, it is possible to choose an n-element set $\{y_1, \ldots, y_n\} \subseteq S_V$ of variable symbols containing none of the symbols in $\mathrm{sym}_V(\varphi_1) \cup \cdots \cup \mathrm{sym}_V(\varphi_s) \cup \{x_1, \ldots, x_m\}$. Let $\psi_i = \varphi_i(c_1/y_1, \ldots, c_n/y_n)$. Then $(\psi_1, \ldots, \psi_s)$ is a T-proof of $\varphi(c_1/y_1, \ldots, c_m/y_m)$ over S. Lemma 1.5.5 now yields $T \vdash_S \bigwedge_{x_1} \cdots \bigwedge_{x_m} \varphi(c_1/x_1, \ldots, c_m/x_m)$. $\square$

Lemma 1.5.7 (Deduction)**.** *For every S-theory T and $\varphi \in \mathcal{F}_S^\bullet$ and $\psi \in \mathcal{F}_S$:*

$$T \vdash (\varphi \to \psi) \;\Leftrightarrow\; T \cup \{\varphi\} \vdash \psi$$

Proof. If $T \vdash (\varphi \to \psi)$, then $T \cup \{\varphi\} \vdash (\varphi \to \psi)$ by $(\mathtt{FM})$ and the following tree diagram illustrates how ψ is provable from $T \cup \{\varphi\}$:

$$\begin{array}{ccc} \varphi & & (\varphi \to \psi) \\ & \diagdown\;\diagup & \\ & \psi & \end{array}$$

Conversely, assume $T \cup \{\varphi\} \vdash \psi$. It is then enough to show $T \vdash (\varphi \to \psi_k)$ for all $k \in \{1, \ldots, s\}$ in a $T \cup \{\varphi\}$-proof $(\psi_1, \ldots, \psi_s)$ of ψ. We fix k and let $\pi = (\varphi \to \psi_k)$.

In case $\psi_k = \varphi$ we have $T \vdash \pi$ according to $(\square)$ and $(\mathtt{AX})$.

In case $\psi_k \in T_0$ the following tree diagram shows how π is provable from T:

$$\begin{array}{ccc} \psi_k & & (\psi_k \to \pi) \\ & \diagdown\;\diagup & \\ & \pi & \end{array}$$

In case there are $i, j \in \{1, \ldots, k-1\}$ with $\psi_j = (\psi_i \to \psi_k)$ we may inductively assume $T \vdash \pi_{10} = (\varphi \to \psi_i), \pi_0 = (\varphi \to \psi_j)$ and then see that π is provable from T:

$$\begin{array}{ccccc} & & \pi_{10} & & (\pi_{10} \to \pi_1) \\ & & & \diagdown\;\diagup & \\ \pi_0 & & & \pi_1 = (\pi_0 \to \pi) & \\ & \diagdown & & \diagup & \\ & & \pi & & \end{array}$$

In case there is $i \in \{1, \ldots, k-1\}$ with $\psi_k = \bigwedge_x \psi_i$ we now may inductively assume $T \vdash \pi_{000} = (\varphi \to \psi_i)$ and then see that π is provable from T as follows:

$$
\begin{tikzcd}[column sep=tiny]
\pi_{000} \arrow[d, dash] & & & & & & \\
\pi_{00} = \bigwedge_x \pi_{000} \arrow[dr, dash] & & (\pi_{00} \to \pi_0) \arrow[dl, dash] & & \pi_{01} = (\varphi \to \bigwedge_x \varphi) \arrow[dr, dash] & & (\pi_{01} \to \pi_1) \arrow[dl, dash] \\
& \pi_0 = (\bigwedge_x \varphi \to \psi_k) \arrow[drr, dash] & & & & \pi_1 = (\pi_0 \to \pi) \arrow[dll, dash] & \\
& & & \pi & & &
\end{tikzcd}
$$

□

Lemma 1.5.8 (Double negation)**.** *For every S-theory, all $\varphi \in \mathcal{F}_{\overline{S}}$:*

$$T \vdash \varphi \ \Leftrightarrow \ T \vdash \neg\neg\varphi$$

Proof. Let $(\psi, \psi') \in \{(\varphi, \neg\neg\varphi), (\neg\neg\varphi, \varphi)\}$. Then we have $T \vdash (\psi \to \psi')$ using (AX) such that (MP) yields $T \vdash \psi \Rightarrow T \vdash \psi'$. □

1.6 Completeness theorem

In this section we prove Gödel's Completeness Theorem according to which a theory (of predicate calculus with infinitely many variable symbols) is satisfiable if and only if it is consistent. The idea of the proof we present here is due to Henkin.

Definition 1.6.1. *An S-theory T is* inconsistent *if $T \vdash \bot$. It is* consistent *otherwise.*

Lemma 1.6.2. *For inconsistent S-theories T we have $T \vdash \varphi$ for all S-formulas φ.*

Proof. By (AX) we have $T \vdash (\bot \to \varphi) \in \mathcal{A}_S$ such that (MP) gives $T \vdash \varphi$. □

Lemma 1.6.3. *For every S-theory T and all S-sentences φ we have:*

(1) T is inconsistent iff $T \vdash \varphi$ and $T \vdash \neg\varphi$.

(2) $T \cup \{\varphi\}$ is inconsistent iff $T \vdash \neg\varphi$.

(3) If T is consistent, then so is $T \cup \{\varphi\}$ or $T \cup \{\neg\varphi\}$.

Proof. In (1) we have $\Rightarrow$ by Lemma 1.6.2 and $\Leftarrow$ follows from (MP).

(2) is the special case of Lemma 1.5.7 for $\psi = \bot$.

(3) follows from (1), (2) and Lemma 1.5.8. □

Lemma 1.6.4. *For every inclusion-wise totally ordered set P of consistent S-theories their union $T = \bigcup P$ is again a consistent S-theory.*

Proof. By (CP) it is enough to check that every finite subset $\underline{T}$ of T is consistent. Since $\underline{T} \subseteq \bigcup P$ is finite and P is totally ordered, there is some $T' \in P$ with $\underline{T} \subseteq T'$. The consistency of T' then implies the consistency of $\underline{T}$ in view of (FM). □

Corollary 1.6.5. *Every consistent S-theory is contained in a* maximal consistent *(i.e. inclusion-wise maximal consistent) S-theory.*

Proof. This follows immediately from Lemma 1.6.4 with Zorn's lemma. □

Lemma 1.6.6. *For every maximal consistent S-theory T we have:*

(1) $T \vdash \varphi \Leftrightarrow \varphi \in T \Leftrightarrow \neg\varphi \notin T \Leftrightarrow T \nvdash \neg\varphi$ *for all S-sentences φ.*

(2) $(\varphi \to \psi) \in T \Leftrightarrow (\varphi \notin T$ *or* $\psi \in T)$ *for all S-sentences φ, ψ.*

(3) $\bigwedge_x \varphi \notin T \Leftrightarrow \bigvee_x \neg\varphi \in T$ *for all S-sentences $\bigwedge_x \varphi$.*

(4) $\varphi(x/t) \in T$ *for all* $t \in \mathcal{T}_S^\bullet$ *and S-sentences* $\bigwedge_x \varphi \in T$.

Proof. (1) The first and third implication $\Leftarrow$ hold thanks to ($\mathtt{AX}$), the second implication $\Leftarrow$ is due to the maximality of T in view of Lemma 1.6.3 (3) and the implication $T \vdash \varphi \Rightarrow T \nvdash \neg\varphi$ holds because of Lemma 1.6.3 (1).

(2) If $\varphi, (\varphi \to \psi) \in T$, then $\psi \in T$ by ($\mathtt{MP}$) and (1). This shows $\Rightarrow$. To check $\Leftarrow$ let $\varphi \notin T$ or $\psi \in T$. By (1) there is $\pi_0 \in \{\neg\varphi, \psi\}$ with $\pi_0 \in T$, so $\pi = (\varphi \to \psi) \in T$ by ($\mathtt{MP}$) because of $(\pi_0 \to \pi) \in \mathcal{A}_S$.

(3) This follows from (1) and Lemmas 1.5.5 and 1.5.8.

(4) $T \vdash \varphi(x/t)$ by ($\mathtt{MP}$) because of $(\bigwedge_x \varphi \to \varphi(x/t)) \in \mathcal{A}_S$. Now use (1). □

Definition 1.6.7. *An S-theory T* has witnesses *if for every S-sentence $\bigwedge_x \varphi \notin T$ there exists some $t \in \mathcal{T}_S^\bullet$ with $\varphi(x/t) \notin T$.*

*An S-*ultratheory *is a maximal consistent S-theory with witnesses.*

Lemma 1.6.8. *For every S-ultratheory T and every S-sentence of the form $\bigwedge_x \varphi$ we have $\bigwedge_x \varphi \in T$ if and only if $\varphi(x/t) \in T$ for all $t \in \mathcal{T}_S^\bullet$.*

Proof. $\Leftarrow$ holds, since T has witnesses, and $\Rightarrow$ follows from Lemma 1.6.6 (4). □

The following key result implies that every S-ultratheory is satisfiable:

Theorem 1.6.9. *Let T be an S-ultratheory. Then the unassigned S-structure $\mathcal{M}_T$ with $\underline{\mathcal{M}_T} = \mathcal{T}_S^\bullet$ and*

$$\begin{aligned} \sharp^{\mathcal{M}_T}(m_1, \dots, m_r) = 1 \;&\Leftrightarrow\; \sharp m_1 \cdots m_r \in T \qquad \textit{for } \sharp \in S_{R,r}, \\ \sharp^{\mathcal{M}_T}(m_1, \dots, m_r) \;&=\; \sharp m_1 \cdots m_r \qquad\quad \textit{for } \sharp \in S_{F,r}^{\mathcal{M}_T} \end{aligned}$$

satisfies $\mathcal{M}_T \vDash \pi \Leftrightarrow \pi \in T$ for all S-sentences π.

Proof. Abbreviate $\mathcal{M} = \mathcal{M}_T$ and let π be an S-sentence.

In case $\pi = \bot$ we have $\mathcal{M} \nvDash \pi$ and, since T is consistent, $\pi \notin T$.

In case $\pi = \sharp t_1 \cdots t_r$ we have $\mathcal{M} \vDash \pi \Leftrightarrow \pi \in T$ by definition.

In case $\pi = (\varphi \to \psi)$ we have by structural induction and Lemma 1.6.6 (2)

$$\mathcal{M} \vDash \pi \;\Leftrightarrow\; (\mathcal{M} \nvDash \varphi \text{ or } \mathcal{M} \vDash \psi) \;\Leftrightarrow\; (\varphi \notin T \text{ or } \psi \in T) \;\Leftrightarrow\; \pi \in T\,.$$

In case $\pi = \bigwedge_x \varphi$ with induction, $\varphi^{\mathcal{M}_t^x} = \varphi(x/t)^{\mathcal{M}}$ and Lemma 1.6.8 we see

$$\mathcal{M} \vDash \pi \;\Leftrightarrow\; \mathcal{M} \vDash \varphi(x/t) \text{ for all } t \in \mathcal{T}_S^\bullet \;\Leftrightarrow\; \varphi(x/t) \in T \text{ for all } t \in \mathcal{T}_S^\bullet \;\Leftrightarrow\; \pi \in T\,.$$

□

Lemma 1.6.10. *For consistent S-theories T and extensions by constants $\overline{S}$ of S with infinite S_V and $\overline{S}_C \setminus S_C = \{c_\varphi : \bigvee_x \varphi \in T\}$ and pairwise distinct c_φ*

$$\overline{T} \;=\; T \cup \{\varphi(x/c_\varphi) : \bigvee_x \varphi \in T\}$$

is a consistent $\overline{S}$-theory.

Proof. To arrive at a contradiction assume $\overline{T} \vdash_{\overline{S}} \bot$. In view of (CP) and Lemma 1.5.6 there are $\bigvee_{x_1} \varphi_1, \ldots, \bigvee_{x_n} \varphi_n \in T$ such that with $\psi_i = \varphi_i(x_i/c_{\varphi_i})$ we have

$$T^{(m)} \;=\; T \cup \{\psi_i : i \in \{1, \ldots, m\}\} \;\vdash_{S^{(m)}}\; \bot \tag{$\star$}$$

for $m = n$ with $S^{(m)}$ the extension by constants of S with $S^{(m)}_C \setminus S_C = \{c_{\varphi_1}, \ldots, c_{\varphi_m}\}$. Assuming that ($\star$) holds for some $m \in \{1, \ldots, n\}$ we get by Lemma 1.5.7

$$T^{(m-1)} \;\vdash_{S^{(m)}}\; \neg\psi_m\,.$$

Lemma 1.5.6 yields $T^{(m-1)} \vdash_{S^{(m-1)}} \bigwedge_{x_m} \neg\varphi_m$ and therefore $T^{(m-1)} \vdash_{S^{(m-1)}} \neg\bigvee_{x_m} \varphi_m$ by Lemma 1.5.8. Another application of Lemma 1.5.7 then shows

$$T^{(m-1)} \;=\; T^{(m-1)} \cup \{\bigvee_{x_m} \varphi_m\} \;\vdash_{S^{(m-1)}}\; \bot\,.$$

By induction this finally yields the contradiction $T \vdash_S \bot$. $\square$

We make the following definition so that the Completeness Theorem can be stated in such a way that it applies both to propositional logic and to predicate logic:

Definition 1.6.11. *A vocabulary S* has enough variables *if S_V is empty or infinite.*

Theorem 1.6.12. *Assume S has enough variables. For every consistent S-theory T there is an extension by constants $\overline{S}$ of S and an $\overline{S}$-ultratheory $\overline{T}$ with $T \subseteq \overline{T}$.*

Proof. We can recursively define vocabularies $S^{(n)}$ and $S^{(n)}$-theories $T^{(n)} \subseteq \overline{T}^{(n)}$ with

$$\begin{aligned} S^{(0)} &= S\,, & S^{(n)}_C &= \Big\{ \quad c_\varphi \quad : \bigvee_x \varphi \in \overline{T}^{(n-1)} \Big\}\,,\\ T^{(0)} &= T\,, & T^{(n)} &= \overline{T}^{(n-1)} \cup \Big\{ \varphi(x/c_\varphi) : \bigvee_x \varphi \in \overline{T}^{(n-1)} \Big\}\,, \end{aligned}$$

such that $S^{(n)}$ is an extension by constants of $S^{(n-1)}$ with pairwise distinct symbols c_φ. Using Lemma 1.6.10 (in case S_V is infinite) and Corollary 1.6.5 we can assume by induction that every $\overline{T}^{(n)}$ is a maximal consistent $S^{(n)}$-theory.

Let $\overline{S}$ be the extension by constants of S with $\overline{S}_C = \bigcup_{n\in\mathbb{N}} S^{(n)}_C$ and $\overline{T} = \bigcup_{n\in\mathbb{N}} \overline{T}^{(n)}$. It is enough to verify that $\overline{T}$ is an $\overline{S}$-ultratheory.

By Lemmas 1.5.6 and 1.6.4 it follows that $\overline{T}$ is consistent. To prove the maximal consistency, we must show that $\overline{T} \cup \{\varphi\}$ is inconsistent for every $\overline{S}$-sentences $\varphi \notin \overline{T}$. Choosing $n \in \mathbb{N}$ such that $\varphi \in \mathcal{F}_{S^{(n)}}$, the maximal consistency of $\overline{T}^{(n)}$ shows that the $S^{(n)}$-theory $\overline{T}^{(n)} \cup \{\varphi\}$ and so by (FM), (FC) the $\overline{S}$-theory $\overline{T} \cup \{\varphi\}$ is inconsistent.

Finally, $\overline{T}$ has witnesses, since for all $\overline{S}$-sentences $\bigwedge_x \varphi \notin \overline{T}$ we have $\bigvee_x \neg\varphi \in \overline{T}$ by Lemma 1.6.6 (3), so $\neg\varphi(x/c_{\neg\varphi}) \in \overline{T}$ and thus $\varphi(x/c_{\neg\varphi}) \notin \overline{T}$ by Lemma 1.6.6 (1). $\square$

Corollary 1.6.13. *If S has enough variables, every consistent S-theory is satisfiable.*

Proof. Combine Theorems 1.6.9 and 1.6.12. □

Theorem 1.6.14 (Completeness Theorem)**.** *Assuming that S has enough variables, for every S-theory T and all S-formulas φ:*

$$T \vdash \varphi \;\Leftrightarrow\; T \vDash' \varphi \;\Leftrightarrow\; T \vDash \varphi$$

Proof. By Lemma 1.5.5 we can assume that φ is an S-sentence.

The first implication $\Rightarrow$ is Lemma 1.5.2 and the second implication $\Rightarrow$ is clear.

To verify the missing implication $T \vDash \varphi \Rightarrow T \vdash \varphi$ assume that $T \nvdash \varphi$. Then $T \nvdash \neg\neg\varphi$ by Lemma 1.5.8, so by Lemma 1.6.3 (2) $T \cup \{\neg\varphi\}$ is consistent and thus by Corollary 1.6.13 satisfiable, i.e. $T \nvDash \varphi$. □

As a corollary, we can now prove the Compactness Theorem:

Proof of Theorem 1.3.10. We may assume that S has enough variables, since an S-theory is satisfiable iff it is satisfiable as an $\overline{S}$-theory for all $\overline{S}$ with $\overline{S}_X \supseteq S_X$.

Now, according to ($\mathtt{CP}$), T is consistent iff every finite subset of T is consistent. But by Theorem 1.6.14 a theory is consistent iff it is satisfiable. □

Proof of Theorem 1.3.11. Similarly as in the previous proof, we may assume that S has enough variables. But then Theorem 1.3.11 is part of Theorem 1.6.14. □

Theorem 1.6.15. *Assume S_V is infinite and S has only countably many symbols. Then every consistent S-theory T admits a model whose underlying set is countable.*

Proof. Let $\overline{T}$ be an $\overline{S}$-ultratheory constructed as in the proof of Theorem 1.6.12. Then the underlying set $\mathcal{T}^{\bullet}_{\overline{S}}$ of the model $\mathcal{M}_{\overline{T}}$ of $\overline{T} \supseteq T$ is countable. □

Example 1.6.16. Let $S = S^{\mathrm{Peano}}$. Theorem 1.6.15 and the Compactness Theorem (Theorem 1.3.10) imply the existence of countable S-structures that satisfy all the S-sentences satisfied by the natural numbers but contain "infinite" elements:

Let $\mathrm{Th}(\check{\mathbb{N}})$ be the set of all S-sentences satisfied in the standard model $\check{\mathbb{N}}$ of PA (see Example 1.4.9). We consider an extension by constants $\overline{S}$ of S with $\overline{S}_C \setminus S_C = \{c\}$ and $\overline{T} = \mathrm{Th}(\check{\mathbb{N}}) \cup \{\varphi_n : n \in \mathbb{N}\}$ where $\varphi_n = c \not\equiv t_n$ with $t_0 = \mathtt{0}$ and $t_n = \mathtt{S}t_{n-1}$.

Clearly, for every finite subset $\underline{T}$ of $\overline{T}$, the standard model $\check{\mathbb{N}}$ of PA becomes a model of $\underline{T}$ by choosing $c^{\check{\mathbb{N}}} = \max\{n \in \mathbb{N} : \varphi_n \in \underline{T}\} + 1$. So, according to Theorems 1.3.10 and 1.6.15 there is a model $\mathcal{N}$ of $\overline{T}$ whose underlying set is countable. In particular, $\mathcal{N}$ is a model of $\mathrm{Th}(\check{\mathbb{N}})$ with an element $c^{\mathcal{N}}$, which cannot be obtained from $\mathtt{0}^{\mathcal{N}}$ by a finite number of applications of the successor operation $\mathtt{S}^{\mathcal{N}}$.

2 Set Theory

2.1 Universes

Just as commutative ring theory studies $\equiv$-respecting models of CRT (Definition 1.4.4) we regard set theory as the study of models of ZFC (Definition 1.4.6).

Definition 2.1.1. *Let $T \subseteq \mathrm{ZFC}$ be an S^{Set}-theory with equality $\equiv$.*

A T-universe *is an unassigned $\equiv$-respecting model $\mathcal{M}$ of T, where the elements of $\underline{\mathcal{M}}$ are called* $\mathcal{M}$-sets *and the subsets of $\underline{\mathcal{M}}$ definable in $\mathcal{M}$ are called* $\mathcal{M}$-classes.

To begin with, let us unravel what the satisfaction of the cryptic axioms of ZFC means for T-universes in common mathematical language:

Remark 2.1.2. A T-universe $\mathcal{M}$ consists of a set $M = \underline{\mathcal{M}}$ together with a binary relation $\sqsubseteq$ given by $\epsilon^{\mathcal{M}}$, which in case $\mathrm{ZF}^{\circ} = \mathtt{EXT} \cup \mathtt{EMP} \cup \mathtt{PAI} \cup \mathtt{UNI} \cup \mathtt{REP} \subseteq T$ has the following properties:

(1) Extensionality. The following function $M \to \mathcal{P}(M)$ is injective:

$$X \ \mapsto\ \sqsubseteq^{-1}(X) \ =\ \{V \in M : V \sqsubseteq X\}$$

(2) Empty set. There exists $\varnothing \in M$ with

$$\sqsubseteq^{-1}(\varnothing) \ =\ \emptyset\,.$$

(3) Pairing. For all $X, Y \in M$ there exists $[X, Y] \in M$ with

$$\sqsubseteq^{-1}\big([X, Y]\big) \ =\ \{X, Y\}\,.$$

(4) Union. For all $X \in M$ there exists $\bigsqcup X \in M$ with

$$\sqsubseteq^{-1}\big(\bigsqcup X\big) \ =\ \bigsqcup \sqsubseteq^{-1}(X)$$

where $\bigsqcup C = \{V \in M : V \sqsubseteq W \text{ for some } W \in C\}$ for $\mathcal{M}$-classes C.

(5) Replacement. For every $X \in M$ and all partial functions $f \colon \sqsubseteq^{-1}(X) \nrightarrow M$ definable in $\mathcal{M}$ there exists $f[X] \in M$ with

$$\sqsubseteq^{-1}(f[X]) \ =\ f(\sqsubseteq^{-1}(X))\,.$$

In case $\mathrm{ZF}^{\circ} \cup \mathtt{POW} \subseteq T$ the binary relation $\sqsubseteq$ on M will additionally satisfy:

(6) Power set. For all $X \in M$ there exists $\mathsf{P}(X) \in M$ with

$$\sqsubseteq^{-1}(\mathsf{P}(X)) \ =\ \mathsf{P}(\sqsubseteq^{-1}(X))$$

where $\mathsf{P}(C) = \{V \in M : U \sqsubseteq V \Rightarrow U \in C \text{ for all } U \in M\}$ for $\mathcal{M}$-classes C.

In case $\mathrm{ZF}^\circ \cup \mathtt{INF} \subseteq T$:

(7) Infinity. There exists an *inductive* $X \in M$, i.e. with the properties $\underline{0} = \varnothing \sqsubseteq X$ and $V + \underline{1} = \bigsqcup[V, [V, V]] \sqsubseteq X$ for all $V \in \sqsubseteq^{-1}(X)$.

In case $\mathrm{ZF}^\circ \cup \mathtt{CHO} \subseteq T$:

(8) Choice. For all $X \in M$ such that $\sqsubseteq^{-1}(V)$ with $V \in \sqsubseteq^{-1}(X)$ are non-empty and pairwise disjoint there exists $Y \in M$ with

$$|\sqsubseteq^{-1}(V) \cap \sqsubseteq^{-1}(Y)| = 1 \text{ for all } V \in \sqsubseteq^{-1}(X)\,.$$

And finally, in case $\mathrm{ZF}^\circ \cup \mathtt{REG} \subseteq T$:

(9) Regularity. For all $X \in M$ with $X \neq \varnothing$ there exists $V \in \sqsubseteq^{-1}(X)$ with

$$\sqsubseteq^{-1}(V) \cap \sqsubseteq^{-1}(X) \;=\; \emptyset\,.$$

For $X \in M$ the elements of $\sqsubseteq^{-1}(X)$ are called *$\mathcal{M}$-elements of X*.

We will say that an $\mathcal{M}$-class C *forms an $\mathcal{M}$-set* if there exists an $\mathcal{M}$-set X such that $C = \sqsubseteq^{-1}(X)$. In this case we write $[C]$ for X.

We use the common abbreviations

$$\begin{aligned} \mathrm{ZF}^- &= \mathrm{ZF}^\circ \cup \mathtt{POW} \cup \mathtt{INF}\,, & \mathrm{ZFC}^- &= \mathrm{ZF}^- \cup \mathtt{CHO}\,,\\ \mathrm{ZF} &= \mathrm{ZF}^- \cup \mathtt{REG}\,, & \mathrm{ZFC} &= \mathrm{ZFC}^- \cup \mathtt{REG}\,. \end{aligned}$$

Remark 2.1.3. For all $\mathcal{M}$-sets X we have $\left[\sqsubseteq^{-1}(X)\right] = X$.

Convention 2.1.4. From now on, fix a ZF°-universe $\mathcal{M}$, abbreviate $M = \underline{\mathcal{M}}$ and denote by $\sqsubseteq$ the relation on M given by $\epsilon^{\mathcal{M}}$.

2.2 Elementary sets

To construct many elementary sets, the assumption that $\mathcal{M}$ merely is a ZF°-universe, which does not necessarily satisfy Power set, Infinity, Choice and Regularity, is enough.

Definition 2.2.1. *Let X, Y, $V_1, \ldots, V_n$ be $\mathcal{M}$-sets and let C be an $\mathcal{M}$-class.*

We introduce the following notation for the $\mathcal{M}$-set Y listed in the right column of the table whenever $\sqsubseteq^{-1}(Y)$ equals the corresponding entry in the left column:

$\sqsubseteq^{-1}(Y)$	Y
$\{V \in M : \cdots\}$	$[V : \cdots]$
$\{V \in \sqsubseteq^{-1}(X) : \cdots\}$	$[V \sqsubseteq X : \cdots]$
$\{V_1, V_2, \ldots, V_n\}$	$[V_1, V_2, \ldots, V_n]$

Instead of $\bigsqcup[V_1, V_2, \ldots, V_n]$ we usually write $V_1 \sqcup V_2 \sqcup \cdots \sqcup V_n$.

*$X \sqsubseteq C$ means $\sqsubseteq^{-1}(X) \subseteq C$ and in this case X is called an $\mathcal{M}$-*subset *of C.*

*$X \sqsubseteq Y$ means $X \sqsubseteq \sqsubseteq^{-1}(Y)$ and in this X is called an $\mathcal{M}$-*subset *of Y.*

Lemma 2.2.2. *For $\mathcal{M}$-sets $X_1, \ldots, X_n$ there exists an $\mathcal{M}$-set $[X_1, \ldots, X_n]$.*

Proof. We get $\mathcal{M}$-sets $[X_i] = [X_i, X_i]$ using Pairing. Given $\mathcal{M}$-sets $[X_1, \ldots, X_{i-1}]$ and $[X_i]$ Pairing and Union yield the $\mathcal{M}$-set

$$\bigsqcup[[X_1, \ldots, X_{i-1}], [X_i]] \;=\; [X_1, \ldots, X_i]\,.$$

Now use induction. □

Lemma 2.2.3 (Separation). *For $\mathcal{M}$-sets X and $\mathcal{M}$-classes C there is an $\mathcal{M}$-set*

$$X \sqcap C \;=\; \big[V \sqsubseteq X : V \in C\big]\,.$$

Proof. Let π be an S^{Set}-formula, $\mathrm{fvar}(\pi) \subseteq \{v, w_1, \ldots, w_n\}$, $W_1, \ldots, W_n \in M$ with

$$C \;=\; \big\{V \in M : \mathcal{M} \vDash \pi[(v, w_1, \ldots, w_n)/(V, W_1, \ldots, W_n)]\big\}\,.$$

Take $\varphi = ((u \equiv v \wedge v \,\epsilon\, x) \wedge \pi)$ with distinct $u, x \in S^{\mathrm{Set}}_V \setminus \{v, w_1, \ldots, w_n\}$. Then

$$\begin{aligned} f \;&=\; \big\{(U, V) \in M^2 : \mathcal{M} \vDash \varphi[(u, v, x, w_1, \ldots, w_n)/(U, V, X, W_1, \ldots, W_n)]\big\} \\ &=\; \big\{(V, V) \in M^2 : V \sqsubseteq X \text{ and } V \in C\big\} \end{aligned}$$

is a partial function $\sqsubseteq^{-1}(X) \nrightarrow M$ definable in $\mathcal{M}$. Replacement now yields the $\mathcal{M}$-set $f[X]$, which has the desired property. □

Lemma 2.2.4 (Difference). *For $\mathcal{M}$-sets X and $\mathcal{M}$-classes C there is an $\mathcal{M}$-set*

$$X \smallsetminus C \;=\; \big[U \sqsubseteq X : U \notin C\big]\,.$$

Proof. The set $M \setminus C$ is an $\mathcal{M}$-class, since C is one. So take $X \smallsetminus C = X \sqcap (M \setminus C)$. □

Lemma 2.2.5 (Intersection). *For non-empty $\mathcal{M}$-classes C there is an $\mathcal{M}$-set*

$$\bigsqcap C \;=\; \big[V : V \sqsubseteq W \text{ for all } W \in C\big]\,.$$

Proof. The set $D = \{V \in M : V \sqsubseteq W \text{ for all } W \in C\}$ is an $\mathcal{M}$-class, since C is one. Take $\bigsqcap C = X \sqcap D$ for some arbitrarily chosen $X \in C$. □

Definition 2.2.6. *We will write $X \smallsetminus Y$ for $X \smallsetminus \sqsubseteq^{-1}(Y)$ and $\bigsqcap X$ for $\bigsqcap \sqsubseteq^{-1}(X)$.*

Instead of $\bigsqcap[V_1, V_2, \ldots, V_n]$ we usually write $V_1 \sqcap V_2 \sqcap \cdots \sqcap V_n$.

The (ordered) $\mathcal{M}$-pair *formed by $\mathcal{M}$-sets X and Y is defined as*

$$\langle X, Y\rangle \;=\; [X, [X, Y]]$$

and then recursively the $\mathcal{M}$-tuple *of $\mathcal{M}$-sets $X_1, \ldots X_n$ as*

$$\langle X_1, \ldots, X_n\rangle \;=\; \langle\langle X_1, \ldots, X_{n-1}\rangle, X_n\rangle\,.$$

For $\mathcal{M}$-classes C and D we define their cartesian product *as the $\mathcal{M}$-class*

$$C * D \;=\; \big\{\langle U, V\rangle : U \in C \text{ and } V \in D\big\}\,.$$

We also write $X * D$ for $\sqsubseteq^{-1}(X) * D$ and $C * Y$ for $C * \sqsubseteq^{-1}(Y)$.

Lemma 2.2.7 (Cartesian product). *For $\mathcal{M}$-sets X and Y there is an $\mathcal{M}$-set*

$$X * Y \;=\; \sqsubseteq^{-1}(X) * \sqsubseteq^{-1}(Y)\,.$$

Proof. Since the functions $f_U\colon \sqsubseteq^{-1}(Y) \to M$, $V \mapsto \langle U, V\rangle$, are definable in $\mathcal{M}$, we firstly get for each $U \sqsubseteq X$ an $\mathcal{M}$-set $f_U[Y] = [\langle U, V\rangle : V \sqsubseteq Y]$ by Replacement. Secondly, $g\colon \sqsubseteq^{-1}(X) \to M$, $U \mapsto f_U[Y]$, is also definable in $\mathcal{M}$ such that applying Replacement again yields the $\mathcal{M}$-set $g[X]$. Finally, take $X * Y = \bigsqcup g[X]$. $\square$

Exercise 2.2.8. Assuming $\mathcal{M}$ satisfies Power set, find another proof of Cartesian product that uses Union and Separation but not Replacement.

Definition 2.2.9. *The* disjoint union *of $\mathcal{M}$-sets X and Y is*

$$X \mathbin{\dot\sqcup} Y \;=\; \big([\underline{0}] * X\big) \sqcup \big([\underline{1}] * Y\big)$$

where $\underline{0} = \varnothing$ and $\underline{1} = [\underline{0}]$. We then define for $\mathcal{M}$-sets $X_1, \dots, X_n$ and $\boxtimes \in \{*, \dot\sqcup\}$

$$X_1 \boxtimes \cdots \boxtimes X_n \;=\; (X_1 \boxtimes \cdots \boxtimes X_{n-1}) \boxtimes X_n\,.$$

Relations

Definition 2.2.10. *We call a relation \$ between $\mathcal{M}$-classes $C_1, \dots, C_n$, i.e. a subset of $C_1 \times \cdots \times C_n$, an* $\mathcal{M}$-class relation *if it is definable in $\mathcal{M}$. Whenever the $\mathcal{M}$-class*

$$\langle \$ \rangle \;=\; \{\langle U_1, \dots, U_n\rangle : (U_1, \dots, U_n) \in \$\}$$

forms an $\mathcal{M}$-set we say that \$ forms an $\mathcal{M}$-set *and then write [\$] for [⟨\$⟩].*

As to be expected, an $\mathcal{M}$-class function *is an $\mathcal{M}$-class relation that is a function.*

Lemma 2.2.11. *An $\mathcal{M}$-class function $f\colon C \to D$ forms an $\mathcal{M}$-set if and only if its domain C forms an $\mathcal{M}$-set. In this case, its image $f(C)$ also forms an $\mathcal{M}$-set.*

In particular, for every $\mathcal{M}$-set X the restriction $f|_{\sqsubseteq^{-1}(X)}$ forms an $\mathcal{M}$-set $f|_X$.

Proof. If $[C] = X$ for some $\mathcal{M}$-set X, then $[f(C)] = f[X]$ by Replacement and then $[f] = (X * f[X]) \sqcap \langle f\rangle$ by Cartesian product and Separation.

For the converse note that $p\colon \langle f\rangle \to C$, $\langle U, V\rangle \mapsto U$, is definable in $\mathcal{M}$ and has image C. So Replacement yields $[C] = p[Z]$ if $[f] = Z$ for some $\mathcal{M}$-set Z. $\square$

Definition 2.2.12. *$\mathcal{M}$-subsets $\diamond$ of $X_1 * \cdots * X_n$ are called* n-ary $\mathcal{M}$-relations *between the $\mathcal{M}$-sets $X_1, \dots, X_n$. In case $X = X_1 = \cdots = X_n$ we speak of $\mathcal{M}$-relations on X.*

Observe that $\diamond$ gives rise to an $\mathcal{M}$-class relation

$$\hat\diamond \;=\; \{(U_1, \dots, U_n) : \langle U_1, \dots, U_n\rangle \sqsubseteq \diamond\}\,.$$

In case $\diamond$ is a binary $\mathcal{M}$-relation we write $U \diamond V$ instead of $\langle U, V\rangle \sqsubseteq \diamond$ and we say that $\diamond$ is injective, functional *etc. if $\hat\diamond$ has the respective property.*

Given that $\diamond$ is a binary $\mathcal{M}$-relation on X we call $\diamond$ reflexive, transitive, a partial order, an equivalence relation *etc. if the relation $\hat\diamond$ has the respective property.*

Accordingly, an $\mathcal{M}$-element of X is said to be a minimal element, least element, greatest element *etc. w.r.t. $\diamond$ if it has the respective property w.r.t. $\hat\diamond$.*

The main purpose of the next exercise is to introduce some notations:

Exercise 2.2.13. Let $f \sqsubseteq X * Y$ and $g \sqsubseteq Y * Z$ be $\mathcal{M}$-relations and $A \sqsubseteq X$.

(i) IDENTITY ON X. The identity on $\sqsubseteq^{-1}(X)$ forms an $\mathcal{M}$-set

$$\mathrm{id}_X \;=\; [\langle U, U\rangle : U \sqsubseteq X]\,.$$

(ii) DOMAIN OF f. The domain of $\hat{f}$ forms an $\mathcal{M}$-set

$$\mathrm{dom}(f) \;=\; [U : \langle U, V\rangle \sqsubseteq f]\,.$$

(iii) IMAGE OF f. The image of $\hat{f}$ forms an $\mathcal{M}$-set

$$\mathrm{img}(f) \;=\; [V : \langle U, V\rangle \sqsubseteq f]\,.$$

(iv) INVERSE OF f. The inverse relation of $\hat{f}$ forms an $\mathcal{M}$-set

$$f^{-1} \;=\; [\langle V, U\rangle : \langle U, V\rangle \sqsubseteq f]\,.$$

(v) RESTRICTION OF f. The restriction of $\hat{f}$ to $\sqsubseteq^{-1}(A)$ forms an $\mathcal{M}$-set

$$f|_A \;=\; [\langle U, V\rangle \sqsubseteq f : U \sqsubseteq A]\,.$$

(vi) IMAGE OF A UNDER f. The image of $\sqsubseteq^{-1}(A)$ under $\hat{f}$ forms an $\mathcal{M}$-set

$$f[A] \;=\; \mathrm{img}(f|_A)\,.$$

(vii) COMPOSITION OF f AND g. The composition $\hat{g} \circ \hat{f}$ forms an $\mathcal{M}$-set

$$g \circ f \;=\; [\langle U, W\rangle : \langle U, V\rangle \sqsubseteq f \text{ and } \langle V, W\rangle \sqsubseteq g]\,.$$

Definition 2.2.14. *We write $f\colon X \twoheadrightarrow Y$ to indicate that f is an $\mathcal{M}$-*function*, i.e. a functional $\mathcal{M}$-relation, with* $\mathrm{dom}(f) = X$ *and* $\mathrm{img}(f) \sqsubseteq Y$.

We say that $f\colon X \twoheadrightarrow Y$ is surjective*, if* $\mathrm{img}(f) = Y$.

We say that $f\colon X \twoheadrightarrow Y$ is bijective *if f is injective and surjective.*

For $\mathcal{M}$-functions we also write $f(U) = V$ instead of $U \diamond V$.

Lemma 2.2.15 (QUOTIENT SET)**.** *For each equivalence relation $\sim$ on an $\mathcal{M}$-set X there exists a surjective $\mathcal{M}$-function*

$$f_\sim\colon X \twoheadrightarrow X/_\sim, \quad U \mapsto {\sim}[U] = [V : U \sim V]\,.$$

Proof. Clearly, $F\colon \sqsubseteq^{-1}(X) \to M$, $U \mapsto {\sim}[U]$, is definable in $\mathcal{M}$ such that according to Lemma 2.2.11 we can take $f_\sim = [F]$. $\square$

Definition 2.2.16. *For $\mathcal{M}$-sets X and $\mathcal{M}$-classes D we define the $\mathcal{M}$-class*

$$\{X \rightharpoonup D\} \;=\; \{f \in M : f \text{ is an } \mathcal{M}\text{-function } X \rightharpoonup D\}\;,$$

where $f\colon X \rightharpoonup D$ means that $\mathrm{dom}(f) = X$ *and* $\mathrm{img}(\hat{f}) \subseteq D$.

$\mathcal{M}$-functions $X \in \{I \rightharpoonup M\}$ are called families of $\mathcal{M}$-sets. *Such families X will be often written as $\langle X_i\rangle_{i \sqsubseteq I}$ where $X_i = X(i)$.*

Lemma 2.2.17 (Union of families)**.** *For every family $\langle X_i\rangle_{i\sqsubseteq I}$ of $\mathcal{M}$-sets there exist $\mathcal{M}$-sets*

$$\bigsqcup_{i \sqsubseteq I} X_i \;=\; \bigsqcup\,[X_i : i \sqsubseteq I] \quad \text{and} \quad \dot{\bigsqcup_{i\sqsubseteq I}} X_i \;=\; \bigsqcup\,[[i] * X_i : i \sqsubseteq I]\,.$$

Proof. By Image we have $[X_i : i \sqsubseteq I]$, so the first statement follows with Union.

The function $\sqsubseteq^{-1}(I) \to M$ given by $i \mapsto [i] * X_i$ is definable in $\mathcal{M}$, hence forms an $\mathcal{M}$-set Y according to Lemma 2.2.11. Apply the first statement to the family $\langle Y_i\rangle_{i\sqsubseteq I}$ to get the second statement. □

Lemma 2.2.18 (Function set)**.** *Assume $\mathcal{M}$ satisfies* Power set. *For all $\mathcal{M}$-sets X and Y there exists an $\mathcal{M}$-set*

$$[X \rightharpoonup Y] \;=\; [f : f \text{ is an } \mathcal{M}\text{-function } X \rightharpoonup Y]\,.$$

Proof. It is not hard to see that $C = \{f : f \text{ is an } \mathcal{M}\text{-function } X \rightharpoonup Y\}$ is an $\mathcal{M}$-class. Using Power set and Separation we can take $[X \rightharpoonup Y] = \mathsf{P}(X * Y) \sqcap C$. □

Lemma 2.2.19 (Product of families)**.** *Assume $\mathcal{M}$ satisfies* Power set. *For every family $\langle X_i\rangle_{i\sqsubseteq I}$ of $\mathcal{M}$-sets there exists an $\mathcal{M}$-set*

$$\mathop{\ast}_{i\sqsubseteq I} X_i \;=\; \Big[f : \langle f_i\rangle_{i\sqsubseteq I} \text{ is a family of } \mathcal{M}\text{-sets with } f_i \sqsubseteq X_i \text{ for all } i \sqsubseteq I\Big]\,.$$

Proof. $C = \{f \in \{I \rightharpoonup M\} : f_i \sqsubseteq X_i \text{ for all } i \sqsubseteq I\}$ is an $\mathcal{M}$-class. With Union of families, Function set and Separation $\mathop{\ast}_{i\sqsubseteq I} X_i = \big[I \rightharpoonup \bigsqcup_{i\sqsubseteq I} X_i\big] \sqcap C$. □

Entering the universe

With all these constructions it now is straightforward to mimic "inside of $\mathcal{M}$" statements and proofs about sets to obtain corresponding results for $\mathcal{M}$-sets. For instance, we can formulate and prove an analog of the Knaster–Tarski Theorem for $\mathcal{M}$-sets:

Definition 2.2.20. *$\langle X, \leq\rangle$ is called a* complete lattice in $\mathcal{M}$ *if $\leq$ is a partial order on X such that every $\mathcal{M}$-subset $Y \sqsubseteq X$ has an infimum and a supremum w.r.t. $\leq$.*

Example 2.2.21. If $\mathsf{P}(X)$ exists, then $\langle \mathsf{P}(X), \sqsubseteq\rangle$ is a complete lattice in $\mathcal{M}$, where for each $Y \sqsubseteq \mathsf{P}(X)$ its infimum is $\bigsqcap Y$ and its supremum $\bigsqcup Y$.

Lemma 2.2.22 (Knaster–Tarski)**.** *Let $\langle X, \leq\rangle$ be a complete lattice in $\mathcal{M}$, $f\colon X \rightharpoonup X$ preserve $\leq$, and $X^f = [x \sqsubseteq X : f(x) = x]$. Then $\langle X^f, \leq\rangle$ is a complete lattice in $\mathcal{M}$.*

Proof. This is proved similarly as Problem Set 3, Exercise 4. □

2.3 Ordinal numbers

Natural numbers are useful for counting because they have the property that whenever finitely many things have been labeled by natural numbers (say with $0, 1, 2, \ldots, n$) there is a unique smallest natural number that comes next (here $n + 1$).

Trying to find an analog of natural numbers inside $\mathcal{M}$ readily leads to the notion of $\mathcal{M}$-ordinals. The principle of induction even works for more general $\mathcal{M}$-classes, which are called well-founded.

Well-founded relations

Definition 2.3.1. *A binary $\mathcal{M}$-class relation $\prec$ on C is called* well-founded on C *if every non-empty $\mathcal{M}$-class $C' \subseteq C$ has a* $\prec$-minimal *element, i.e. an element $Y \in C'$ such that there is no $X \in C'$ with $X \prec Y$.*

A binary $\mathcal{M}$-class relation $\prec$ on C is called set-like *if*

$$C_{\prec X} \;=\; \{U \in C : U \prec X\}$$

forms an $\mathcal{M}$-set for all $X \in C$.

Theorem 2.3.2 (Well-founded induction)**.** *Let $\prec$ be a well-founded $\mathcal{M}$-class relation on Ω and let $C \subseteq \Omega$ be an $\mathcal{M}$-class. Then $C = \Omega$ if for all $\beta \in \Omega$:*

$$\Omega_{\prec\beta} \subseteq C \implies \beta \in C$$

Proof. For every $\prec$-minimal element β in the $\mathcal{M}$-class $\Omega \setminus C$ we must have $\Omega_{\prec\beta} \subseteq C$, so $\beta \in C$ by assumption, which is absurd. $\square$

Ordinals

Definition 2.3.3. *An $\mathcal{M}$-class C is* transitive *if $\alpha \sqsubset \beta \Rightarrow \alpha \in C$ for every $\beta \in C$. An $\mathcal{M}$-set X is* transitive *if $\sqsubset^{-1}(X)$ is transitive.*

A binary $\mathcal{M}$-class relation is a well-order *if it is a well-founded strict total order.*

$\mathcal{M}$-ordinals *are transitive $\mathcal{M}$-sets γ well-ordered by $[\langle\alpha, \beta\rangle : \alpha \sqsubset \beta \sqsubset \gamma]$.*

We will denote by $\mathbb{O}$ the $\mathcal{M}$-class of all $\mathcal{M}$-ordinals equipped with the $\mathcal{M}$-class relation $<$ given by $\alpha < \beta \Leftrightarrow \alpha \sqsubset \beta$.

An $\mathcal{M}$-class Ω is called an initial segment of $\mathbb{O}$ *if either $\Omega = \mathbb{O}$ or $\Omega = \mathbb{O}_{<\alpha}$ for some $\mathcal{M}$-ordinal α.*

Remark 2.3.4. An $\mathcal{M}$-set X is transitive iff $\bigsqcup X \sqsubseteq X$ iff $Y \sqsubseteq X$ for all $Y \sqsubset X$.

If an $\mathcal{M}$-set X is transitive, so is $X \sqcup [X]$ and, if it exists, $\mathsf{P}(X)$.

If C is a non-empty $\mathcal{M}$-class of transitive $\mathcal{M}$-sets, then $\bigsqcup C$ and $\bigsqcap C$ are transitive.

Example 2.3.5. $\underline{0} = \varnothing$, $\underline{1} = \underline{0} \sqcup [\underline{0}]$, $\underline{2} = \underline{1} \sqcup [\underline{1}]$, $\ldots$ are $\mathcal{M}$-ordinals.

$\underline{2} \sqcup [[\underline{1}]]$ is transitive but not an $\mathcal{M}$-ordinal.

We will now show step by step that the $\mathcal{M}$-class $\mathbb{O}$ of all $\mathcal{M}$-ordinals is transitive and that $<$ is a set-like well-order on it.

Lemma 2.3.6. $\mathbb{O}$ *is transitive.*

Proof. Let $\alpha \mathrel{\varepsilon} \beta \in \mathbb{O}$. We need to show $\alpha \in \mathbb{O}$. Since ε well-orders β, it also well-orders α because of $\alpha \sqsubseteq \beta$ by transitivity of β. To verify that α is transitive take $\gamma \mathrel{\varepsilon} \delta \mathrel{\varepsilon} \alpha$. Firstly, the transitivity of β yields $\delta \mathrel{\varepsilon} \beta$ and then also $\gamma \mathrel{\varepsilon} \beta$ such that, secondly, the transitivity of ε on β implies $\gamma \mathrel{\varepsilon} \alpha$ as required. $\square$

Corollary 2.3.7. *The relation $<$ on $\mathbb{O}$ is set-like. More precisely, for all $\alpha \in \mathbb{O}$*

$$\alpha \; = \; [\gamma : \gamma < \alpha] \; = \; [\mathbb{O}_{<\alpha}] \; .$$

Proof. Use EXTENSIONALITY and the fact that $\mathcal{M}$-elements of $\mathcal{M}$-ordinals are again $\mathcal{M}$-ordinals according to Lemma 2.3.6. $\square$

Lemma 2.3.8. *For $\mathcal{M}$-ordinals α and β the following hold:*

(1) $\alpha < \beta$ *iff* $\alpha \sqsubsetneq \beta$ *and in this case* $\alpha = \min(\beta \smallsetminus \alpha)$*.*

(2) $\alpha \sqsubseteq \beta$ *or* $\beta \sqsubseteq \alpha$*.*

Proof. (1) $\Rightarrow$ holds by transitivity of β and since $<$ is trichotomous on β.

For $\Leftarrow$ assume $\alpha \sqsubsetneq \beta$ and let $\gamma = \min(\beta \smallsetminus \alpha)$. Clearly, $\gamma < \beta$ such that it suffices to show $\alpha = \gamma$. Since γ is also an $\mathcal{M}$-ordinal by Lemma 2.3.6 and β is transitive, this is equivalent to $\delta < \alpha \Leftrightarrow \delta < \gamma$ for all $\delta < \beta$. Now $\delta < \gamma$ implies $\delta < \alpha$ by the minimality of γ. Vice versa, $\delta < \alpha$ and $\delta \not< \gamma$ would lead to $\gamma \leq \delta < \alpha$ and then to the contradiction $\gamma < \alpha$, since $<$ is a strict total order on β.

(2) Clearly, $\gamma = \alpha \sqcap \beta$ is an $\mathcal{M}$-ordinal with $\gamma \sqsubseteq \alpha$ and $\gamma \sqsubseteq \beta$. If both of these inclusions were proper, we would have the contradiction $\gamma < \alpha \sqcap \beta = \gamma$ by (1). $\square$

Corollary 2.3.9. *The relation $<$ on $\mathbb{O}$ is a strict total order.*

Proof. Let $\alpha, \beta, \gamma \in \mathbb{O}$.

By Lemma 2.3.8 (1) we have $\alpha \not< \alpha$ and $\alpha < \beta < \gamma \Rightarrow \alpha < \gamma$.

By Lemma 2.3.8 (1) and (2) exactly one of $\alpha = \beta$, $\alpha < \beta$, $\beta < \alpha$ holds. $\square$

Lemma 2.3.10. *For non-empty $\mathcal{M}$-classes C of $\mathcal{M}$-ordinals their intersection $\bigsqcap C$ is an $\mathcal{M}$-ordinal, which is the minimum of C.*

Proof. It follows from $\bigsqcap C \sqsubseteq \alpha$ for every $\alpha \in C$ that $\bigsqcap C$ is an $\mathcal{M}$-ordinal and then by Lemma 2.3.8 that it is the infimum of C in $\mathbb{O}$. Hence, it is enough to check that C has a minimum. To do this, pick some $\gamma \in C$. If γ is not a minimum of C, then $[\alpha < \gamma : \alpha \in C] = \gamma \sqcap C$ is non-empty and the minimum δ of $\gamma \sqcap C$ is also the minimum of C: Indeed, if we had $\delta \not\leq \alpha$ for some $\alpha \in C$, by Corollary 2.3.9 we would have $\alpha < \delta$, then $\alpha < \gamma$ because of $\delta < \gamma$, and thus $\alpha \mathrel{\varepsilon} \gamma \sqcap C$. $\square$

Corollary 2.3.11. *The relation $<$ on $\mathbb{O}$ is a set-like well-order.*

Proof. Combine Corollaries 2.3.7 and 2.3.9 and Lemma 2.3.10. $\square$

Remark 2.3.12. $\mathbb{O}$ does not form an $\mathcal{M}$-set α, since otherwise we would have $\alpha \in \mathbb{O}$ by Lemma 2.3.6 and Corollary 2.3.11, i.e. the contradiction $\alpha < \alpha$.

Hence, M does not form an $\mathcal{M}$-set either (otherwise so would $\mathbb{O} = M \sqcap \mathbb{O}$).

Lemma 2.3.13. *For $\mathcal{M}$-sets X of $\mathcal{M}$-ordinals their union $\bigsqcup X$ is an $\mathcal{M}$-ordinal, which is the supremum of X in $\mathbb{O}$.*

Proof. Clearly, $\bigsqcup X$ is transitive and by Corollary 2.3.11 well-ordered, so $\bigsqcup X \in \mathbb{O}$. Lemma 2.3.8 implies that $\bigsqcup X$ is the supremum of X. $\square$

Lemma 2.3.14. *Every transitive $\mathcal{M}$-subclass C of $\mathbb{O}$ is an initial segment.*

Proof. If $C = \mathbb{O}$, there is nothing to show. Otherwise $C \subseteq \mathbb{O}_{<\alpha}$ for $\alpha = \min \mathbb{O} \setminus C$. Indeed, $\gamma \in C$ and $\alpha \leq \gamma$ would imply $\alpha < \gamma$ and then $\alpha \in C$ by transitivity of C. It follows that C forms an $\mathcal{M}$-set $\alpha \sqcap C$, which is transitive and well-ordered by $<$, i.e. it is an $\mathcal{M}$-ordinal. Now use Corollary 2.3.7. $\square$

Successor and limit ordinals

There are two fundamentally different types of ordinals: successor and limit ordinals. The existence of limit ordinals will only be guaranteed by Infinity.

Lemma 2.3.15. *For every $\mathcal{M}$-ordinal α also $\alpha + \underline{1} = \alpha \sqcup [\alpha]$ is an $\mathcal{M}$-ordinal and $\alpha + \underline{1} = \bigsqcap\{\gamma \in \mathbb{O} : \gamma > \alpha\}$, i.e. $\alpha + \underline{1}$ is the least $\mathcal{M}$-ordinal greater than α.*

Proof. Since α is transitive, so is $\alpha + \underline{1}$. It is also straightforward to see that $\alpha + \underline{1}$ is well-ordered by $\sqsubset$ with maximum α, since α is well-ordered by $\sqsubset$.

The second part follows from Lemma 2.3.10 since, as a consequence of Lemma 2.3.8 and Corollary 2.3.9, $\alpha + \underline{1}$ is the infimum of $\{\gamma \in \mathbb{O} : \gamma > \alpha\}$ in $\mathbb{O}$. $\square$

Definition 2.3.16. *An $\mathcal{M}$-ordinal β is a* successor ordinal in $\mathcal{M}$ *with* predecessor $\beta - \underline{1} = \alpha$ *if $\beta = \alpha + \underline{1}$ for some $\mathcal{M}$-ordinal α. Otherwise β is a* limit ordinal in $\mathcal{M}$.

Write $\mathbb{O}_{+\underline{1}}$ for the $\mathcal{M}$-class of all successor ordinals in $\mathcal{M}$.

Write $\mathbb{O}_{\lim}$ for the $\mathcal{M}$-class of all limit ordinals in $\mathcal{M}$.

Example 2.3.17. $\underline{0}$ is a limit ordinal.

$\underline{1} = \underline{0} + \underline{1}$, $\underline{2} = \underline{1} + \underline{1}$, ... are successor ordinals.

Lemma 2.3.18. *An $\mathcal{M}$-ordinal β is a successor ordinal iff $\beta = (\bigsqcup \beta) + \underline{1}$.*

An $\mathcal{M}$-ordinal β is a limit ordinal iff $\beta = \bigsqcup \beta$ iff $\alpha + \underline{1} < \beta$ for all $\alpha < \beta$.

Proof. Clearly, $\bigsqcup \beta \leq \beta \leq (\bigsqcup \beta) + \underline{1}$ by Corollary 2.3.7 and Lemmas 2.3.8 and 2.3.13. If $\bigsqcup \beta < \beta$, then $(\bigsqcup \beta) + \underline{1} = \beta$ by Lemma 2.3.15. If $\beta = \alpha + \underline{1}$, then $\bigsqcup \beta = \alpha < \beta$.

For all $\alpha < \beta$ we have $\alpha + \underline{1} \leq \beta$ by Lemma 2.3.15 where the inequality always is strict for limit ordinals β and is an equality for successor ordinals β with $\alpha = \beta - \underline{1}$. □

Theorem 2.3.19 (Transfinite induction). *Let Ω be an initial segment of $\mathbb{O}$ and let $C \subseteq \Omega$ be an $\mathcal{M}$-class. Then $C = \Omega$ if the following hold:*

(i) $\forall\, \beta \in \Omega \cap \mathbb{O}_{+\underline{1}} : (\beta - \underline{1} \in C \implies \beta \in C)$

(ii) $\forall\, \beta \in \Omega \cap \mathbb{O}_{\lim} : ((\forall\, \alpha < \beta : \alpha \in C) \implies \beta \in C)$

Proof. If $\Omega \setminus C$ were non-empty, it would have a minimum β. So if β were a successor, then $\beta - \underline{1} \in C$ contradicting (i). By (ii) β cannot be a limit ordinal either. □

Normal functions

Many important $\mathcal{M}$-class functions defined on the $\mathcal{M}$-ordinals are order-preserving and continuous (w.r.t. order topology). Such functions are called *normal.*

Convention 2.3.20. For this subsection fix an initial segment Ω in $\mathbb{O}$.

Definition 2.3.21. *An $\mathcal{M}$-subclass C of Ω is said to be* closed in Ω *if $\bigsqcup X \in C$ for every non-empty $\mathcal{M}$-set $X \sqsubseteq C$ with $\bigsqcup X \in \Omega$.*

As usual, if C forms an $\mathcal{M}$-set, then this $\mathcal{M}$-set is called closed in Ω, *if so is C.*

An $\mathcal{M}$-class function $f : \Omega \to \mathbb{O}$ is normal *if it preserves $<$ and* commutes with $\bigsqcup$, *i.e. $f(\bigsqcup X) = \bigsqcup f[X]$ for all non-empty $\mathcal{M}$-sets $X \sqsubseteq \Omega$ with $\bigsqcup X \in \Omega$.*

Whether a given function is normal can be checked with the following criterion:

Lemma 2.3.22. *An $\mathcal{M}$-class function $f : \Omega \to \mathbb{O}$ is normal iff $f(\beta - \underline{1}) < f(\beta)$ for successor ordinals $\beta \in \Omega$ and $f(\beta) = \bigsqcup f[\beta]$ for limit ordinals $\beta \in \Omega \setminus \{\underline{0}\}$.*

Proof. $\Rightarrow$ We get $f(\beta - \underline{1}) < f(\beta)$ for successor ordinals β, since f is order-preserving, and $f(\beta) = \bigsqcup f[\beta]$ for limit ordinals $\beta \in \Omega \setminus \{\underline{0}\}$, since then $\beta = \bigsqcup \beta$.

$\Leftarrow$ To verify that f preserves $<$ we apply transfinite induction to the $\mathcal{M}$-class

$$C \;=\; \{\beta \in \Omega : f(\alpha) < f(\beta) \text{ for all } \alpha < \beta\}\,.$$

Let $\alpha < \beta \in \Omega$. We must show $f(\alpha) < f(\beta)$. If β is a successor ordinal, then this holds by assumption, if $\alpha = \beta - \underline{1}$, and by induction, if $\alpha < \beta - \underline{1}$. If β is a limit ordinal, then $\alpha < \alpha + \underline{1} < \beta$ such that by assumption $f(\alpha) < f(\alpha + \underline{1}) \leq \bigsqcup f[\beta] = f(\beta)$.

It remains to verify $f(\bigsqcup X) = \bigsqcup f[X]$ for non-empty $\mathcal{M}$-sets X of $\mathcal{M}$-ordinals in Ω for which the supremum $\beta = \bigsqcup X$ of X belongs to Ω. If $\beta \sqin X$, then $f(\beta) = \bigsqcup f[X]$ because f is order-preserving. If $\beta \not\sqin X$, then β is a limit ordinal different from $\underline{0}$ and X is unbounded in β such that $f(\beta) = \bigsqcup f[\beta] = \bigsqcup f[X]$, where the first equality holds by assumption and the second one uses again that f is order-preserving. □

Example 2.3.23. The successor function $s\colon \mathbb{O} \to \mathbb{O}$, $\alpha \mapsto \alpha + \underline{1}$, preserves $<$.

However, it is does not commute with $\bigsqcup$, if there exists a non-zero limit ordinal β in $\mathcal{M}$, since then $s(\beta) = \beta + \underline{1} \neq \beta = \bigsqcup s[\beta]$.

Lemma 2.3.24. *If $f\colon \Omega \to \Omega$ is an $\mathcal{M}$-class function that preserves $<$, then f is an isomorphism $(\Omega, <) \to (C, <)$ onto its image C and $\alpha \leq f(\alpha)$ for all $\alpha \in \Omega$.*

Proof. If $\beta \leq \alpha \in \Omega$ with $f(\alpha) < f(\beta)$, then $f(\beta) \leq f(\alpha) < f(\beta)$, which is absurd.

Moreover, if $f(\alpha) < \alpha$, then $f(f(\alpha)) < f(\alpha)$. Hence, $\{\alpha \in \Omega : f(\alpha) < \alpha\}$ cannot have a minimum and must therefore be empty. $\square$

Lemma 2.3.25. *The image of a normal $\mathcal{M}$-class function $f\colon \Omega \to \Omega$ is closed in Ω and either is unbounded in Ω, if Ω has no maximum, or contains the maximum of Ω.*

Proof. Let $C = \operatorname{img}(f)$ and take $X \sqsubseteq C$ non-empty. By Lemma 2.3.24 the identity is dominated by f, hence the second statement, and $f^{-1}\colon C \to \Omega$ also preserves $<$.

Now, on the one hand, if $\bigsqcup f^{-1}[X] \in \Omega$, then $\bigsqcup X = f(\bigsqcup f^{-1}[X]) \in C$ because f commutes with $\bigsqcup$.

On the other hand, if $\bigsqcup f^{-1}[X] \notin \Omega$, then $f^{-1}[X]$ is unbounded in Ω and, consequently, $X = f[f^{-1}[X]]$ is unbounded in C such that by the already proved second statement X is unbounded in Ω, i.e. $\bigsqcup X \notin \Omega$. $\square$

We collect some more properties of normal $\mathcal{M}$-class functions:

Lemma 2.3.26. *For normal $\mathcal{M}$-class functions $f\colon \Omega \to \Omega$ the following hold:*

(a) $f(\beta) \in \mathbb{O}_{\text{lim}} \setminus \{\underline{0}\}$ *for all* $\beta \in \Omega \cap \mathbb{O}_{\text{lim}} \setminus \{\underline{0}\}$.

(b) $f(\beta) = \bigsqcup_{\alpha<\beta} f(\alpha + \underline{1})$ *for all* $\beta \in \Omega \setminus \{\underline{0}\}$.

(c) $\Omega^f = \{\alpha \in \Omega : f(\alpha) = \alpha\}$ *is closed in* Ω.

(d) If $g\colon \Omega \to \mathbb{O}$ *is a normal* $\mathcal{M}$*-class function, so is* $g \circ f\colon \Omega \to \mathbb{O}$.

(e) For each $\delta \in \Omega$ *with* $f(\underline{0}) \leq \delta$ *the* $\mathcal{M}$*-set* $(\delta + \underline{1}) \sqcap \operatorname{img}(f)$ *has a maximum, i.e. there is a greatest element in the image of* f *less than or equal to* δ.

Proof. (a) Because of $f(\beta) = \bigsqcup f[\beta]$ by Lemma 2.3.22 $f[\beta]$ is unbounded in $f(\beta)$. Hence, $f(\beta)$ must be a non-zero limit ordinal.

(b) If β is a limit, then $[f(\alpha + \underline{1}) : \alpha < \beta] = f[\beta]$ and the formula holds because f commutes with $\bigsqcup$.

If β is a successor, $[f(\alpha + \underline{1}) : \alpha < \beta]$ has maximum $f(\beta)$ because f preserves $<$.

(c) $f(\bigsqcup X) = \bigsqcup f[X] = \bigsqcup X$ for non-empty $X \sqsubseteq \Omega^f$ with $\bigsqcup X \in \Omega$, so $\bigsqcup X \in \Omega^f$.

(d) Preserving $<$ and commuting with $\bigsqcup$ clearly is preserved by composition.

(e) The $\mathcal{M}$-set $X = f^{-1}[\delta + \underline{1}]$ is non-empty with $\bigsqcup X \in \Omega$. The claim therefore follows from $f(\bigsqcup X) = \bigsqcup f[X] = \bigsqcup (\delta + \underline{1}) \sqcap \operatorname{img}(f)$. $\square$

Exercise 2.3.27. Consider Ω equipped with the order topology. Show:

(i) The limit points of Ω in Ω are the non-zero limit ordinals in Ω.

(ii) The isolated points of Ω consist of $\underline{0}$ and the successor ordinals in Ω.

(iii) $\mathcal{M}$-subclasses of Ω are closed in Ω if they are closed in the topological sense.

(iv) $\mathcal{M}$-class functions $f\colon \Omega \to \mathbb{O}$ that preserve $\leq$ commute with $\bigsqcup$ iff they are continuous.

Natural numbers

Definition 2.3.28. Natural numbers in $\mathcal{M}$ *are $\mathcal{M}$-ordinals α such that there is no limit ordinal γ in $\mathcal{M}$ with $\underline{0} < \gamma \leq \alpha$.*

Write $\mathbb{N}^{\mathcal{M}}$ for the $\mathcal{M}$-class of natural numbers in $\mathcal{M}$.

An $\mathcal{M}$-set X is finite *if there is a bijective $\mathcal{M}$-function $X \twoheadrightarrow \alpha$ for some natural number α in $\mathcal{M}$. Otherwise it is* infinite.

Theorem 2.3.29 (Induction for natural numbers)**.** *Let $C \subseteq \mathbb{N}^{\mathcal{M}}$ be an $\mathcal{M}$-class. Then $C = \mathbb{N}^{\mathcal{M}}$ if $\underline{0} \in C$ and $\alpha + \underline{1} \in C$ for all $\alpha \in C$.*

Proof. If we had $C \neq \mathbb{N}^{\mathcal{M}}$, then the assumptions would imply that $\beta = \min(\mathbb{N}^{\mathcal{M}} \setminus C)$ is a limit ordinal different from $\underline{0}$, which contradicts the definition of $\mathbb{N}^{\mathcal{M}}$. □

For instance, induction can be used to prove that the $\mathcal{M}$-subsets of a natural number always form an $\mathcal{M}$-set, even when POWER SET in general might not hold:

Lemma 2.3.30. *For all $\alpha \in \mathbb{N}^{\mathcal{M}}$ the $\mathcal{M}$-class $\mathsf{P}(\sqsubseteq^{-1}(\alpha))$ forms an $\mathcal{M}$-set $\mathsf{P}(\alpha)$.*

Proof. Let C be the $\mathcal{M}$-class of all natural numbers α in $\mathcal{M}$ for which $\mathsf{P}(\sqsubseteq^{-1}(\alpha))$ forms an $\mathcal{M}$-set $\mathsf{P}(\alpha)$. Clearly, $\underline{0} \in C$ because we can take $\mathsf{P}(\underline{0}) = [\varnothing]$ and if $\alpha \in C$ then we can take $\mathsf{P}(\alpha + \underline{1}) = \mathsf{P}(\alpha) \sqcup [X \sqcup [\alpha] : X \sqsubseteq \mathsf{P}(\alpha)]$ to conclude $\alpha + \underline{1} \in C$. □

Lemma 2.3.31. *Assuming $\mathcal{M}$ satisfies* INFINITY*, $\mathbb{N}^{\mathcal{M}}$ forms a set ω, which is the intersection of all inductive $\mathcal{M}$-sets.*

Proof. The set D of all inductive $\mathcal{M}$-sets is an $\mathcal{M}$-class, which is non-empty thanks to INFINITY. Clearly, $\omega = \bigsqcap D \sqcap \mathbb{N}^{\mathcal{M}}$ again is inductive such that $\sqsubseteq^{-1}(\omega) = \mathbb{N}^{\mathcal{M}}$ according to Theorem 2.3.29. Consequently, $\omega = \bigsqcap D = \left[\mathbb{N}^{\mathcal{M}}\right]$. □

Corollary 2.3.32. *Assuming $\mathcal{M}$ satisfies* INFINITY*, ω is the least limit ordinal in $\mathcal{M}$ different from $\underline{0}$.*

Proof. Evidently, $\mathbb{N}^{\mathcal{M}}$ is transitive and inherits the well-order of $\mathbb{O}$, so $\omega \in \mathbb{O}$. By definition $\mathcal{M}$-ordinals different from $\underline{0}$ and less than ω are successor ordinals. Also ω cannot itself be a successor ordinal, since then it would be a natural number, leading to the contradiction $\omega < \omega$. □

Remark 2.3.33. With the argument given in the proof of Corollary 2.3.32 and, given that limit ordinals are inductive, it is easy to see that $\mathcal{M}$ satisfies INFINITY iff $\mathbb{N}^{\mathcal{M}} \neq \mathbb{O}$ iff there exists at least one limit ordinal in $\mathcal{M}$ iff $\mathbb{N}^{\mathcal{M}}$ forms an $\mathcal{M}$-set.

Remarkably, there may be more natural numbers in $\mathcal{M}$ than those given by the "external" natural numbers:

Remark 2.3.34. The function $\mathbb{N} \to \mathbb{N}^{\mathcal{M}}$, $n \mapsto \underline{n}$, with $\underline{n+1} = \underline{n} + \underline{1}$, is injective because of $\left|\sqsubseteq^{-1}(\underline{n})\right| = n$. An argument similar to the one given in Example 1.6.16 shows that it is not necessarily surjective for every choice of $\mathcal{M}$ (if there is any choice at all), i.e. we might have

$$\underline{\mathbb{N}} \;=\; \{\underline{n} : n \in \mathbb{N}\} \subsetneq \mathbb{N}^{\mathcal{M}}$$

Then $\underline{\mathbb{N}}$ cannot be an $\mathcal{M}$-class because otherwise $\mathbb{N}^{\mathcal{M}} \setminus \underline{\mathbb{N}}$ would have a minimum β, but then $\beta - \underline{1} = \underline{n}$ for some $n \in \mathbb{N}$ leading to the contradiction $\beta = \underline{n} + \underline{1} = \underline{n+1} \in \underline{\mathbb{N}}$.

So if we want the natural numbers that are used to construct formulas to be the same as the natural numbers in $\mathcal{M}$, we must postulate this as an additional axiom.

2.4 Recursive definitions

Well-founded recursion

Definition 2.4.1. *Let $\prec$ be an $\mathcal{M}$-class relation on C.*

If $X \in C$ lies in an $\mathcal{M}$-set $Y \sqsubseteq C$ that is closed under $\prec$-predecessors*, i.e. with the property $U \prec V \Rightarrow U \mathrel{\varepsilon} Y$ for all $U \in C$ and $V \mathrel{\varepsilon} Y$, we define*

$$C_{\preceq^\infty X} \;=\; \bigsqcap\{Y \in M : Y \text{ is closed under } \prec\text{-predecessors and } X \mathrel{\varepsilon} Y \sqsubseteq C\}$$

and call it the $\prec$-predecessor closure *of X in C.*

We say that the relation $\prec$ admits predecessor closures in C *if every element of C has a $\prec$-predecessor closure in C.*

Example 2.4.2. Clearly, $\mathrel{\varepsilon}$ admits predecessor closures in any transitive $\mathcal{M}$-class C consisting of transitive $\mathcal{M}$-sets. In this case, $C_{\preceq^\infty X} = X \sqcup [X]$ for all $X \in C$.

In particular, the $<$-predecessor closure of $\alpha \in \mathbb{O}$ is $\alpha + \underline{1}$.

Lemma 2.4.3. *Set-like $\mathcal{M}$-class relations $\prec$ that admit predecessor closures in C are well-founded on C, if each non-empty $\mathcal{M}$-subset of C has a $\prec$-minimal element.*

Proof. We must prove that every non-empty $\mathcal{M}$-class $D \subseteq C$ has a minimal element.

To do this, pick $Y \in D$. Then there is a minimal element X in the $\mathcal{M}$-set $C_{\preceq^\infty Y} \sqcap D$. If X were not minimal in D, then there would be an $z \mathrel{\varepsilon} C_{\prec X} \sqcap D$, leading to the contradiction $z \mathrel{\varepsilon} C_{\preceq^\infty Y} \sqcap D$. □

Theorem 2.4.4 (Well-founded recursion)**.** *Let $\prec$ be a set-like well-founded $\mathcal{M}$-class relation on C that admits predecessor closures in C. Then for every $\mathcal{M}$-class function $g\colon M \to M$ there exists a unique $\mathcal{M}$-class function $f\colon C \to M$ such that*

$$f(X) \;=\; g\big(\langle X, [f|_{C_{\prec X}}]\rangle\big) \quad \textit{for all } X \in C.$$

Proof. Let $f = \bigcup_{h \in F} \hat{h}$ be the $\mathcal{M}$-class relation defined by

$$F \;=\; \big\{h : \text{there is } X \in C \text{ and } h\colon C_{\preceq^\infty X} \rightharpoonup M \\ \text{such that } h(Z) = g\big(\langle Z, h|_{C_{\prec Z}}\rangle\big) \text{ for all } Z \in C_{\preceq^\infty X}\big\}$$

To begin with, f is a function, since otherwise there would be $h\colon C_{\preceq^\infty X} \rightharpoonup M$ and $h'\colon C_{\preceq^\infty Y} \rightharpoonup M$ in F such that the $\mathcal{M}$-class $\{Z \sqin C_{\preceq^\infty X} \sqcap C_{\preceq^\infty Y} : h(Z) \neq h'(Z)\}$ would be non-empty, hence have a minimal element Z, leading to the contradiction

$$h(Z) \;=\; g\big(\langle Z, h|_{C_{\prec Z}}\rangle\big) \;=\; g\big(\langle Z, h'|_{C_{\prec Z}}\rangle\big) \;=\; h'(Z)\,.$$

We now prove $\mathrm{dom}(f) = C$ with well-founded induction (Theorem 2.3.2). For this, take $X \in C$ with $C_{\prec X} \subseteq \mathrm{dom}(f)$. It is easily checked that

$$C_{\preceq^\infty X} \;=\; [X] \sqcup \bigsqcup \big[C_{\preceq^\infty W} : W \sqin C_{\prec X}\big]\,.$$

From this and the definition of f we can therefore conclude $C_{\preceq^\infty X} \smallsetminus [X] \sqsubseteq \mathrm{dom}(f)$. Hence, $f|_{C_{\preceq^\infty X} \smallsetminus [X]} \sqcup \big[\big\langle X, g\big(\langle X, f|_{C_{\prec X}}\rangle\big)\big\rangle\big]$ lies in F, so $X \in \mathrm{dom}(f)$.

To verify uniqueness let $f'\colon C \to M$ be another $\mathcal{M}$-class function with the property that $f'(X) = g\big(\langle X, [f'|_{C_{\prec X}}]\rangle\big)$ for all $X \in C$. Then $D = \{X \in C : f(X) = f'(X)\}$ is an $\mathcal{M}$-class and for every $X \in C$ with $C_{\prec X} \subseteq D$ we have

$$f(X) \;=\; g\big(\langle X, [f|_{C_{\prec X}}]\rangle\big) \;=\; g\big(\langle X, [f'|_{C_{\prec X}}]\rangle\big) \;=\; f'(X)\,.$$

This shows $X \in D$. Using Theorem 2.3.2 once again yields $D = C$, i.e. $f = f'$. $\square$

Recursion for ordinals

Definition 2.4.5. *Let Ω be an initial segment of $\mathbb{O}$. An* Ω-sequence *is an $\mathcal{M}$-class function $x\colon \Omega \to M$, usually written as $\langle x_\alpha\rangle_{\alpha\in\Omega}$.*

For $\beta \in \Omega$ we also write $\langle x_\alpha\rangle_{\alpha<\beta}$ for the family $\langle x_\alpha\rangle_{\alpha\sqin\beta}$ of $\mathcal{M}$-sets.

Corollary 2.4.6 (Recursive definition of sequences)**.** *Let Ω be an initial segment of $\mathbb{O}$. For every $\mathcal{M}$-class function $f\colon M \to M$ there is an Ω-sequence $\langle x_\alpha\rangle_{\alpha\in\Omega}$ with*

$$x_\beta \;=\; f(\langle \beta, \langle x_\alpha\rangle_{\alpha<\beta}\rangle) \qquad \textit{for all } \beta \in \Omega.$$

Proof. This is a direct consequence of Theorem 2.4.4. $\square$

Lemma 2.4.7. *Assuming $\mathcal{M}$ satisfies* Infinity, *every $\mathcal{M}$-class relation $\prec$ on C admits predecessor closures in C.*

Explicitly, $C_{\preceq^\infty X}$ is the ω-iterate of $[X]$ under $V \mapsto \bigsqcup[[C_{\prec U}] : U \sqin V]$.

Proof. Let Z be the ω-iterate of $[X]$ under $V \mapsto \bigsqcup[[C_{\prec U}] : U \sqin V]$.

By definition it is $Z = \bigsqcup_{\alpha<\omega} Z_\alpha$ where $Z_{\underline{0}} = [X]$ and $Z_{\alpha+\underline{1}} = \bigsqcup[[C_{\prec U}] : U \sqin Z_\alpha]$. So we have to show for $Z \sqsubseteq C_{\preceq^\infty X}$ that

$$\big\{\alpha \in \mathbb{N}^{\mathcal{M}} : Z_\alpha \sqsubseteq C_{\preceq^\infty X}\big\} \;=\; \mathbb{N}^{\mathcal{M}}\,,$$

which clearly is true by induction. The inclusion $C_{\preceq^\infty X} \sqsubseteq Z$ holds because $Z \sqsubseteq C$ is closed under $\prec$-predecessors and contains X. $\square$

Remark 2.4.8. Using Lemmas 2.4.3 and 2.4.7 and assuming $\mathcal{M}$ satisfies Infinity, the relation $\sqsubseteq$ is seen to be well-founded on M iff $\mathcal{M}$ satisfies Regularity.

Iterates

Definition 2.4.9. *The* β-iterate $f^{\beta}(X)$ *of an* $\mathcal{M}$*-set* X *under an* $\mathcal{M}$*-class function* $f\colon M \to M$ *is given by the* $\mathbb{O}$*-sequence* $\mathrm{It}_{f,X} = \langle f^{\beta}(X)\rangle_{\beta\in\mathbb{O}}$ *recursively defined by*

$$\begin{aligned} f^{\underline{0}}(X) &= X\,,\\ f^{\beta}(X) &= f(f^{\beta-\underline{1}}(X)) \quad \text{for } \beta\in\mathbb{O}_{+\underline{1}},\\ f^{\beta}(X) &= \textstyle\bigsqcup_{\alpha<\beta} f^{\alpha}(X) \quad \text{for } \beta\in\mathbb{O}_{\lim}\setminus\{\underline{0}\}. \end{aligned}$$

Lemma 2.4.10. *Let* Ω *be an initial segment of* $\mathbb{O}$ *and* $\gamma\in\Omega$ *and* $f,g\colon M\to\Omega$ $\mathcal{M}$*-class functions with* $\mathrm{It}_{f,\gamma}(\Omega)\subseteq\Omega$ *and* $\mathrm{It}_{g,\gamma}(\Omega)\subseteq\Omega$*. Then the following hold:*

(a) If $\alpha < f(\alpha)$ *for all* $\alpha\in\Omega$*, then* $\mathrm{It}_{f,\gamma}|_{\Omega}$ *is normal.*

(b) If $\alpha \le f(\alpha)$ *for all* $\alpha\in\Omega$*, then* $\mathrm{It}_{f,\gamma}|_{\Omega}$ *preserves* $\le$*.*

(c) If $f|_{\Omega}$ *preserves* $\le$ *and* $f(\alpha)\le g(\alpha)$ *for all* $\alpha\in\Omega$*, then we have* $f^{\beta}(\delta)\le g^{\beta}(\gamma)$ *for all* $\delta\le\gamma$ *and* $\beta\in\Omega$*.*

Proof. (a) Because of $f^{\beta-\underline{1}}(X) < f(f^{\beta-\underline{1}}(X)) = f^{\beta}(X)$ for successor ordinals $\beta\in\Omega$ and $f^{\beta}(X) = \bigsqcup_{\alpha<\beta} f^{\alpha}(X)$ for limit ordinals $\beta\in\Omega\setminus\{\underline{0}\}$ we can apply Lemma 2.3.22.

(b) Use a similar argument as for (a).

(c) By induction we may assume $f^{\alpha}(\delta)\le g^{\alpha}(\gamma)$ for all $\alpha<\beta$. The claim is then clear for non-zero limit ordinals β. For $\beta=\underline{0}$ it holds anyway. If β is a successor ordinal, then $f^{\beta}(\delta) = f(f^{\beta-\underline{1}}(\delta)) \le f(g^{\beta-\underline{1}}(\gamma)) \le g(g^{\beta-\underline{1}}(\gamma)) = g^{\beta}(\gamma)$. □

As an application we can prove that normal $\mathcal{M}$-class functions $\mathbb{O}\to\mathbb{O}$ have arbitrarily large fixed points:

Lemma 2.4.11. *Assuming* $\mathcal{M}$ *satisfies* Infinity*, for every normal* $\mathcal{M}$*-class function* $f\colon\mathbb{O}\to\mathbb{O}$ *its* $\mathcal{M}$*-class* $\{\gamma\in\mathbb{O} : f(\gamma)=\gamma\}$ *of fixed points is unbounded in* $\mathbb{O}$*.*

Proof. Fix any limit ordinal $\delta\neq\underline{0}$ in $\mathcal{M}$ (e.g. take $\delta=\omega$). For every $\mathcal{M}$-ordinal α the $\mathcal{M}$-ordinal $\gamma = \bigsqcup\big[f^{\beta}(\alpha) : \beta<\delta\big]$ satisfies

$$f(\gamma) \;=\; \bigsqcup\big[f(f^{\beta}(\alpha)) : \beta<\delta\big] \;=\; \bigsqcup\big[f^{\beta+\underline{1}}(\alpha) : \beta<\delta\big] \;=\; \gamma \;\ge\; f^{\underline{0}}(\alpha) = \alpha\,.$$

□

Ordinal arithmetic

Definition 2.4.12. *For two* $\mathcal{M}$*-ordinals* α,β *their* ordinal sum $\alpha+\beta$ *is the* β*-iterate of* α *under* $\gamma\mapsto\gamma+\underline{1}$*, their* ordinal product $\alpha\cdot\beta$ *is the* β*-iterate of* $\underline{0}$ *under* $\gamma\mapsto\gamma+\alpha$*, and their* ordinal power $\alpha^{(\beta)}$ *is the* β*-iterate of* $\underline{1}$ *under* $\gamma\mapsto\gamma\cdot\alpha$*.*

Remark 2.4.13. Assuming $\mathcal{M}$ satisfies Infinity, it now is easy to construct integers and rational numbers in $\mathcal{M}$. For instance take $(\omega*\omega)/_{\sim}$ as the $\mathcal{M}$-set of *integers in* $\mathcal{M}$ where $\sim$ is defined by $\langle\alpha,\beta\rangle\sim\langle\alpha',\beta'\rangle \Leftrightarrow \alpha+\beta' = \alpha'+\beta$.

2.5 Transitive collapse

Definition 2.5.1. *Let C be an $\mathcal{M}$-class equipped with a set-like well-founded $\mathcal{M}$-class relation $\prec$ that admits predecessor closures in C.*

The rank function on C w.r.t. $\prec$ *is the $\mathcal{M}$-class function* $\mathrm{rk}_{C,\prec}\colon C \to \mathbb{O}$ *with*

$$\mathrm{rk}_{C,\prec}(X) \;=\; \bigsqcup \left[\mathrm{rk}_{C,\prec}(W) + \underline{1} : W \in C_{\prec X}\right] \quad \textit{for all } X \in C.$$

The unique surjective $\mathcal{M}$-class function $t_{C,\prec}\colon C \to T_{C,\prec}$ *satisfying*

$$t_{C,\prec}(X) \;=\; t_{C,\prec}[[C_{\prec X}]] \quad \textit{for all } X \in C$$

is called the transitive collapse *of* $(C, \prec)$.

Example 2.5.2. The transitive collapse of $(\{\underline{0}, \underline{2}, \underline{4}\}, <)$ is given by $\underline{n} \mapsto \frac{n}{\underline{2}}$.

For $\mathcal{M}$-ordinals β we have $\mathrm{rk}_{\mathbb{O},<}(\beta) = \beta = t_{\mathbb{O},<}(\beta)$.

Remark 2.5.3. For $\mathcal{M}$-subclasses $D \subseteq C$ we do not necessarily have $T_{D,\prec} \subseteq T_{C,\prec}$, e.g. for $D = \{\underline{1}, [\underline{1}], \underline{2} \sqcup [\underline{1}]\}$ and $C = \{\underline{0}\} \cup D$ it is $T_{D,\prec} = \{\underline{0}, \underline{1}, \underline{2}\} \not\subseteq C = T_{C,\prec}$.

Lemma 2.5.4. *Let $(C, \prec)$ be as in Definition 2.5.1. For every $\mathcal{M}$-subclass $B \subseteq C$ and all $X \in C$ we have* $\mathrm{rk}_{B,\prec}(X) \leq \mathrm{rk}_{C,\prec}(X)$ *with equality if* $C_{\preceq^\infty X} \sqsubseteq B$.

Proof. Assuming by induction $\mathrm{rk}_{B,\prec}(W) \leq \mathrm{rk}_{C,\prec}(W)$ for all $W \in B_{\prec X}$, we get

$$\begin{aligned}\mathrm{rk}_{B,\prec}(X) \;&=\; \bigsqcup[\mathrm{rk}_{B,\prec}(W) + \underline{1} : W \in B_{\prec X}] \\ &\leq\; \bigsqcup[\mathrm{rk}_{C,\prec}(W) + \underline{1} : W \in C_{\prec X}] \;=\; \mathrm{rk}_{C,\prec}(X)\,.\end{aligned}$$

and, assuming $\mathrm{rk}_{B,\prec}(W) = \mathrm{rk}_{C,\prec}(W)$ for $W \in B_{\prec X}$ with $C_{\preceq^\infty W} \sqsubseteq B$, even equality, if $C_{\preceq^\infty X} \sqsubseteq B$, since then $B_{\prec X} = C_{\prec X}$ and $C_{\preceq^\infty W} \sqsubseteq C_{\preceq^\infty X} \sqsubseteq B$ for $W \in C_{\prec X}$. □

Lemma 2.5.5. *$T_{C,\prec}$ is transitive and $t_{C,\prec}$ is a homomorphism* $(C, \prec) \to (T_{C,\prec}, \in)$.

Proof. Abbreviate $t = t_{C,\prec}$ and $T = T_{C,\prec}$.

Firstly, for all $U \in V \in T$ there is $X \in C$ with $V = t(X) = t[[C_{\prec X}]]$, so $U \in T$.

Secondly, $X \prec Y \Rightarrow t(X) \in t[[C_{\prec Y}]] = t(Y)$ for all $X, Y \in C$. □

Definition 2.5.6. *An $\mathcal{M}$-class relation $\prec$ is* extensional *on an $\mathcal{M}$-class C if the implication $C_{\prec X} = C_{\prec Y} \Rightarrow X = Y$ holds for all $X, Y \in C$.*

Example 2.5.7. The relation $\in$ is extensional on M (and thus also on any transitive $\mathcal{M}$-class) by Extensionality.

Lemma 2.5.8. *$\mathcal{M}$-class relations $\prec$ that well-order C are extensional on C.*

Proof. Just observe that X is the minimum of $C \setminus C_{\prec X}$. □

Mostowski collapse

Theorem 2.5.9 (Mostowski's Isomorphism)**.** *Let $(C, \prec)$ be as in Definition 2.5.1. If the relation $\prec$ is extensional on C, then the transitive collapse $t_{C,\prec}$ is the one and only isomorphism $(C, \prec) \to (D, \in)$ where D is a transitive $\mathcal{M}$-class.*

Proof. Abbreviate $t = t_{C,\prec}$ and $T = T_{C,\prec}$.

If $t\colon C \twoheadrightarrow T$ were not injective, then $\{Y \in C : t(X) = t(Y) \text{ for some } X \in C \setminus \{Y\}\}$ would have a minimal element Y. So we would have $t(X) = t(Y)$ for some $X \neq Y$, i.e. $t[C_{\prec X}] = t[C_{\prec Y}]$ and therefore $C_{\prec X} = C_{\prec Y}$ by the minimal choice of Y, contradicting extensionality.

To verify that t^{-1} is a morphism $(T, \in) \to (C, \prec)$ we show $t(X) \in t(Y) \Rightarrow X \prec Y$ for all $X, Y \in C$. So let $t(X) \in t(Y) = t[[C_{\prec Y}]]$. Then there exists $Z \in C_{\prec Y}$ with $t(X) = t(Z)$. The injectivity of t now implies $X = Z \prec Y$.

To prove uniqueness, let $t'\colon (C, \prec) \to (D, \in)$ be another isomorphism such that D is a transitive $\mathcal{M}$-class. Assuming $t \neq t'$, the $\mathcal{M}$-class $\{Y \in C : t(Y) \neq t'(Y)\}$ would have a minimal element Y. Then $t(Y) = t[[C_{\prec Y}]] = t'[[C_{\prec Y}]] \sqsubseteq t'(Y)$, where the inclusion is due to the fact that t' is a morphism. By transitivity of D every $U \in t'(Y) \setminus t(Y)$ belongs to D, so there would be $X \in C$ with $t'(X) = U \in t'(Y)$. Hence, $X \prec Y$ because t'^{-1} is a morphism, leading due to the minimal choice of Y to the contradiction $U = t'(X) = t(X) \in t(Y)$. □

As a corollary, $\mathcal{M}$-ordinals are up to isomorphism the only well-ordered $\mathcal{M}$-sets:

Corollary 2.5.10. *If an $\mathcal{M}$-set X is well-ordered by an $\mathcal{M}$-relation $\prec$, there is a unique isomorphism $t_{X,\prec}\colon \langle X, \prec\rangle \twoheadrightarrow \langle \alpha, <\rangle$ where $\operatorname{type}_{\mathcal{M}}(X, \prec) = \alpha$ is an $\mathcal{M}$-ordinal.*

Proof. Take $\operatorname{type}_{\mathcal{M}}(X, \prec) = [T_{X,\prec}]$ and use Lemma 2.5.8 and Theorem 2.5.9. □

Corollary 2.5.11. *The identity $\mathrm{id}_{\mathbb{O}}$ is the only automorphism of $(\mathbb{O}, <)$ and, if $\mathcal{M}$ satisfies* Infinity *and* Regularity*, id_M is the only automorphism of $(M, \in)$.*

Proof. Combine Lemma 2.3.6, Corollary 2.3.11, and Theorem 2.5.9 for the first and Remark 2.6.3, Example 2.5.7, and Theorem 2.5.9 for the second statement. □

In contrast to Remark 2.5.3 we have the following positive result:

Lemma 2.5.12. *Let $X \sqsubseteq \alpha$ where α is an $\mathcal{M}$-ordinal. Then $\operatorname{type}_{\mathcal{M}}(X, <) \leq \alpha$.*

Proof. Let $t = t_{X,<}$. It clearly is sufficient to prove $t(\beta) \leq \beta$ for all $\beta \in X$ because this implies $\operatorname{type}_{\mathcal{M}}(X, <) = t[X] \sqsubseteq \alpha$. But indeed we have $t(\beta) = t[X_{<\beta}] \sqsubseteq \beta$ if we assume inductively $t(\alpha) \leq \alpha$ for all $\alpha \in X_{<\beta}$. □

Lemma 2.5.13. *An $\mathcal{M}$-class C is closed and unbounded in $\mathbb{O}$ iff it is the image of a normal $\mathcal{M}$-class function $f\colon \mathbb{O} \to \mathbb{O}$.*

In this case, f is determined by C and is called the enumerator En_C *of C.*

Proof. The "if" part follows from Lemma 2.3.25.

For the "only if" part, consider the inverse $f\colon \Omega \to C$ of the transitive collapse $t_{C,<}$. Now Ω is a transitive $\mathcal{M}$-subclass of $\mathbb{O}$, so it is an initial segment by Lemma 2.3.14. Thus $\Omega = \mathbb{O}$, since $[\Omega] = \alpha$ with $\alpha \in \mathbb{O}$ would imply that C is bounded by $\bigsqcup \alpha$. We already know from Theorem 2.5.9 that f preserves $<$ such that it only remains to show that f commutes with $\bigsqcup$. So let X be any non-empty $\mathcal{M}$-set of $\mathcal{M}$-ordinals. Then $f[X]$ is bounded in C because X is bounded in $\mathbb{O}$ and f is an isomorphism. Since C is closed in $\mathbb{O}$, the supremum $\bigsqcup f[X]$ of $f[X]$ in $\mathbb{O}$ lies in C, so is the supremum in C. Using that f is an isomorphism, we can conclude $f(\bigsqcup X) = \bigsqcup f[X]$.

Due to Lemma 2.3.24 and Theorem 2.5.9 f^{-1} and hence f is determined by C. $\square$

Recall that, when $\mathcal{M}$ satisfies Infinity, the $\mathcal{M}$-class $\mathbb{O}^f$ of fixed points of every normal $\mathcal{M}$-class function $f\colon \mathbb{O} \to \mathbb{O}$ is closed and unbounded in $\mathbb{O}$.

Definition 2.5.14. *Assuming $\mathcal{M}$ satisfies* Infinity*, the* derivative *f' of a normal $\mathcal{M}$-class function $f\colon \mathbb{O} \to \mathbb{O}$ is the enumerator of $\mathbb{O}^f$.*

2.6 Well-founded sets

Convention 2.6.1. In this section, we assume that $\mathcal{M}$ satisfies Infinity.

All $\mathcal{M}$-sets that can be constructed by iterated "elementary operations" – such as taking unions, intersections, pairs, power sets etc. – from the empty $\mathcal{M}$-set $\varnothing$ are well-founded in the sense described below.

Definition 2.6.2. *The* transitive closure *of an $\mathcal{M}$-set X is the $\sqsubset$-predecessor closure of X in M. We denote it by X^∞.*

An $\mathcal{M}$-set X is said to be well-founded *if $\sqsubset$ is well-founded on X^∞.*

We denote by $\mathbb{W}$ the $\mathcal{M}$-class of all well-founded $\mathcal{M}$-sets and for $\alpha \in \mathbb{O}$ set

$$\mathbb{W}_\alpha = \{X \in \mathbb{W} : \mathrm{rk}(X) < \alpha\} \quad \textit{where} \quad \mathrm{rk}(X) = \mathrm{rk}_{X^\infty,\sqsubset}(X)\,.$$

Remark 2.6.3. $\mathcal{M}$ satisfies Regularity iff $\mathbb{W} = M$.

Lemma 2.6.4. *The following hold:*

(a) $\mathbb{W}$ *is transitive.*

(b) $\sqsubset$ *is well-founded on* $\mathbb{W}$.

(c) $\mathrm{rk}(X) = \mathrm{rk}_{\mathbb{W},\sqsubset}(X)$ *for all* $X \in \mathbb{W}$.

(d) $\mathrm{rk}(X) = \bigsqcup[\mathrm{rk}(W) + \underline{1} : W \sqsubset X]$ *for all* $X \in \mathbb{W}$.

(e) $\mathrm{rk}(X) < \mathrm{rk}(Y)$ *for all* $X \sqsubset Y \in \mathbb{W}$.

(f) $\mathbb{O} \subseteq \mathbb{W}$ *and* $\mathrm{rk}(\alpha) = \alpha$ *for all* $\alpha \in \mathbb{O}$.

(g) $X \in \mathbb{W}$ *for all $\mathcal{M}$-sets X with* $X \sqsubseteq \mathbb{W}$.

Proof. (a) Let $X \sqsubset Y \in \mathbb{W}$. If $\sqsubset$ is well-founded on Y^∞, then so it is on $X^\infty \sqsubseteq Y^\infty$.

(b) Using (g) apply Lemma 2.4.3.

(c) Using (g) apply Lemma 2.5.4 with $[B] = X^\infty \sqsubseteq \mathbb{W} = C$.

(d) Use (a) and (c).

(e) Use (a) and (d).

(f) $\mathbb{O} \subseteq \mathbb{W}$ by definition of $\mathcal{M}$-ordinal and Example 2.4.2. Apply Lemma 2.5.4 with $B = \mathbb{O} \subseteq \mathbb{W} = C$ and use (c) and (d) to get $\mathrm{rk}(\alpha) = \mathrm{rk}_{\mathbb{W},\sqsubset}(\alpha) = \mathrm{rk}_{\mathbb{O},\sqsubset}(\alpha) = \alpha$.

(g) Let $P = \bigsqcup_{W \sqsubset X} W^\infty$.

We must show that every non-empty $Y \sqsubseteq X^\infty = [X] \sqcup P$ has a minimal element. Now either $Y \setminus [X]$ is empty, in which case X is a minimal element of $Y = [X]$, since $X \sqsubset X \sqsubseteq \mathbb{W}$ is impossible. Or $Y \setminus [X]$ has an element Z minimal in $P \in \mathbb{W}$. But then also $X \not\sqsubset Z$, since otherwise we would have $X \sqsubset P$ such that by Lemma 2.4.7 there would be $\beta \in \mathbb{N}^{\mathcal{M}}$ and a sequence $\langle X_\alpha \rangle_{\alpha \le \beta}$ with $X_\alpha \sqsubset X_{\alpha-\underline{1}}$ and $X_{\underline{0}} = X_\beta = X$, so $[X_\alpha : \alpha \le \beta] \sqsubseteq X_{\underline{1}}^\infty$ would not have a minimal element, but $X_{\underline{1}} \sqsubset X \sqsubseteq \mathbb{W}$. ☐

Exercise 2.6.5. $\mathbb{W}$ is characterized as the smallest $\mathcal{M}$-class C with the property $X \sqsubseteq C \Leftrightarrow X \in C$ for all $\mathcal{M}$-sets X, i.e. any such C necessarily contains $\mathbb{W}$.

Theorem 2.6.6. *$\mathbb{W}_\beta = \mathsf{P}(\mathbb{W}_{\beta-\underline{1}})$ for successor ordinals β and $\mathbb{W}_\beta = \bigcup_{\alpha<\beta} \mathbb{W}_\alpha$ for limit ordinals β in $\mathcal{M}$.*

Proof. We need $X \sqsubseteq \mathbb{W}_{\beta-\underline{1}} \Leftrightarrow X \in \mathbb{W}_\beta$ for all $\mathcal{M}$-sets X and successor ordinals β. Here, $\Rightarrow$ follows from Lemma 2.6.4 (d) and $\Leftarrow$ from Lemma 2.6.4 (e).

The identity $\mathbb{W}_\beta = \bigcup_{\alpha<\beta} \mathbb{W}_\alpha$ is obvious for limit ordinals β. ☐

Corollary 2.6.7. *Each $\mathbb{W}_\beta$ forms an $\mathcal{M}$-set W_β for $\beta \le \omega$. It is finite for $\beta < \omega$.*

Assuming $\mathcal{M}$ satisfies Power set*, $\mathbb{W}_\beta$ forms an $\mathcal{M}$-set W_β even for all $\beta \in \mathbb{O}$. In this case, W_β is the β-iterate of $\varnothing$ under $X \mapsto \mathsf{P}(X)$.*

Proof. This is a direct consequence of Theorem 2.6.6 and Lemma 2.3.30. For the claim about finiteness prove by induction $X \simeq \alpha \Rightarrow \mathsf{P}(X) \simeq \underline{2}^{(\alpha)}$ for all $\alpha < \omega$. ☐

Remark 2.6.8. Assuming $\mathcal{M}$ satisfies (Infinity and) Regularity, $\mathcal{M}$ satisfies Power set iff $\mathbb{W}_\beta$ forms an $\mathcal{M}$-set for every $\mathcal{M}$-ordinal β.

2.7 Subuniverses

Using the results from the last section, we discuss how to obtain from a T-universe a $(T \cup \mathtt{REG})$-universe where T is either of the theories $\mathrm{ZF}^\circ \cup \mathtt{INF}$, ZF^-, or ZFC^-, which will demonstrate the relative consistency w.r.t. T of the axiom of regularity.

Convention 2.7.1. For this section, fix a ZF°-universe $\mathcal{M}$.

Definition 2.7.2. *For $T \subseteq M$ denote by $\mathcal{M}|_T$ the S^{Set}-structure with*

$$\underline{\mathcal{M}|_T} = T \quad \text{and} \quad \sharp^{\mathcal{M}|_T} = \sharp^{\mathcal{M}}|_{T\times T} \ \text{for } \sharp \in \{\equiv, \epsilon\}\,.$$

Lemma 2.7.3. *Let T be a transitive $\mathcal{M}$-class such that every $\mathcal{M}|_T$-class that forms an $\mathcal{M}$-set forms an $\mathcal{M}|_T$-set. Then $\mathcal{M}|_T$ is a* ZF$^\circ$*-universe where for all $X, Y \in T$:*

(1) $\left(\sqsubseteq^{\mathcal{M}|_T}\right)^{-1}(X) = \left(\sqsubseteq^{\mathcal{M}}\right)^{-1}(X)$.

(2) $\varnothing^{\mathcal{M}|_T} = \varnothing^{\mathcal{M}}$.

(3) $[X, Y]^{\mathcal{M}|_T} = [X, Y]^{\mathcal{M}}$.

(4) $\bigsqcup^{\mathcal{M}|_T} X = \bigsqcup^{\mathcal{M}} X$.

(5) $f[X]^{\mathcal{M}|_T} = f[X]^{\mathcal{M}}$ *for every partial function* $f\colon \sqsubseteq^{-1}(X) \nrightarrow T$ *definable in* $\mathcal{M}|_T$.

(6) If $\mathcal{M} \vDash \mathsf{POW}$*, then* $\mathcal{M}|_T \vDash \mathsf{POW}$ *and* $\mathsf{P}^{\mathcal{M}|_T}(X) = \mathsf{P}^{\mathcal{M}}(X) \sqcap T$.

(7,9) If $\mathcal{M} \vDash \mathsf{INF}$ *and* $T \subseteq \mathbb{W}^{\mathcal{M}}$*, then* $\mathcal{M}|_T \vDash \mathsf{INF} \cup \mathsf{REG}$ *and* $\mathbb{O}^{\mathcal{M}|_T} = \mathbb{O}^{\mathcal{M}} \cap T$.

Moreover, $\bigsqcap^{\mathcal{M}|_T} C = \bigsqcap^{\mathcal{M}} C \sqcap T$ *for every* $\mathcal{M}|_T$*-class* C

Proof. (1) This holds since T is transitive.

(2) $\emptyset$ is an $\mathcal{M}|_T$-class forming the $\mathcal{M}$-set $\varnothing^{\mathcal{M}}$.

(3) $\{X, Y\}$ is an $\mathcal{M}|_T$-class forming the $\mathcal{M}$-set $[X, Y]^{\mathcal{M}}$.

(4) $\bigsqcup^{\mathcal{M}} \sqsubseteq^{-1}(X) = \{V \in T : V \sqsubseteq W \text{ for some } W \in T \text{ with } W \sqsubseteq X\} = \bigsqcup^{\mathcal{M}|_T} \sqsubseteq^{-1}(X)$ by transitivity of T, showing it is an $\mathcal{M}|_T$-class, which forms an $\mathcal{M}$-set $\bigsqcup^{\mathcal{M}} X$.

(5) Let π be an S^{Set}-formula, $\mathrm{fvar}(\pi) \subseteq \{z, w_1, \ldots, w_m\}$, $W_1, \ldots, W_m \in M$ with

$$T \;=\; \left\{Z \in M : \mathcal{M} \vDash \pi[(z, w_1, \ldots, w_m)/(Z, W_1, \ldots, W_m)]\right\}.$$

and σ an S^{Set}-formula, $\mathrm{fvar}(\sigma) \subseteq \{u, v, w_{m+1}, \ldots, w_n\}$, $W_{m+1}, \ldots, W_n \in T$ with

$$f \;=\; \left\{(U, V) \in T^2 : \mathcal{M}|_T \vDash \sigma[(u, v, w_1, \ldots, w_n)/(V, W_1, \ldots, W_n)]\right\}.$$

Then f is also definable in $\mathcal{M}$ as

$$f \;=\; \left\{(U, V) \in M^2 : \mathcal{M} \vDash \sigma^\pi[(u, v, w_1, \ldots, w_n)/(V, W_1, \ldots, W_n)]\right\},$$

where σ^π is recursively given as

$$\sigma^\pi \;=\; \begin{cases} \bot & \text{for } \sigma = \bot, \\ \big((\pi(z/x) \wedge \pi(z/y)) \wedge x \,\sharp\, y\big) & \text{for } \sigma = x \,\sharp\, y \text{ with } \sharp \in \{\equiv, \epsilon\}, \\ \big(\varphi^\pi \to \psi^\pi\big) & \text{for } \sigma = (\varphi \to \psi), \\ \bigwedge_x \big(\pi(z/x) \to \varphi^\pi\big) & \text{for } \sigma = \bigwedge_x \varphi. \end{cases}$$

All in all, $f(\sqsubseteq^{-1}(X))$ is an $\mathcal{M}|_T$-class, which forms an $\mathcal{M}$-set $f[X]^{\mathcal{M}}$.

(6) $\mathsf{P}^{\mathcal{M}}(\sqsubseteq^{-1}(X)) \sqcap T = \{V \in T : U \sqsubseteq V \Rightarrow U \sqsubseteq X \text{ for all } U \in T\} = \mathsf{P}^{\mathcal{M}|_T}(\sqsubseteq^{-1}(X))$ by transitivity of T, showing it is an $\mathcal{M}|_T$-class, which forms an $\mathcal{M}$-set $\mathsf{P}^{\mathcal{M}}(X) \sqcap T$.

(7,9) $T \subseteq \mathbb{W}^{\mathcal{M}}$ means that $\sqsubseteq$ is well-founded on $T_{\sqsubseteq^\infty X} = M_{\sqsubseteq^\infty X}$ for every $X \in T$, hence $\mathcal{M} \vDash \mathsf{REG}$ by Remark 2.6.3.

Using $T \subseteq \mathbb{W}^{\mathcal{M}}$, Exercise 1 of Problem Set 4, and the transitivity of T we get

$$\mathbb{O}^{\mathcal{M}} \cap T \;=\; \{\alpha \in T : \alpha \text{ and all } \gamma < \alpha \text{ are transitive}\} \;=\; \mathbb{O}^{\mathcal{M}|_T} ,$$

which then also implies $\mathbb{N}^{\mathcal{M}|_T} = \mathbb{N}^{\mathcal{M}} \cap T$.

It remains to check $\mathcal{M}|_T \vDash \mathtt{INF}$. Now $\underline{0} \in \mathbb{N}^{\mathcal{M}|_T}$ by (2) and $\alpha + \underline{1} = \alpha \sqcup [\alpha] \in \mathbb{N}^{\mathcal{M}|_T}$ for all $\alpha \in \mathbb{N}^{\mathcal{M}|_T}$ by (3) and (4). Therefore we have $\mathbb{N}^{\mathcal{M}|_T} = \mathbb{N}^{\mathcal{M}}$ by induction in $\mathcal{M}$, showing that the $\mathcal{M}|_T$-class $\mathbb{N}^{\mathcal{M}|_T}$ forms an $\mathcal{M}$-set, which is inductive. $\square$

Assuming $\mathcal{M}$ satisfies Infinity, the $\mathcal{M}$-class $T = \mathbb{W}$ meets all the hypotheses of Lemma 2.7.3 by Lemma 2.6.4. This gives the following relative consistency result:

Theorem 2.7.4. *If $\mathcal{M} \vDash \mathrm{ZFC}^-$, then $\mathcal{M}|_{\mathbb{W}} \vDash \mathrm{ZFC}$.*

Hence, the consistency of ZFC$^-$ *implies the consistency of* ZFC.

Proof. Let us first argue that we can apply Lemma 2.7.3 to $T = \mathbb{W}$. Indeed, the $\mathcal{M}$-class $\mathbb{W}$ is transitive by Lemma 2.6.4 (a) and for every $\mathcal{M}|_{\mathbb{W}}$-class that forms an $\mathcal{M}$-set X we have $X \sqsubseteq \mathbb{W}$ such that $X \in \mathbb{W}$ by Lemma 2.6.4 (g).

Assume now that $\mathcal{M}$ satisfies Choice and let $X \in \mathbb{W}$ be such that all $V \sqsubset X$ are non-empty and pairwise disjoint. Choose an $\mathcal{M}$-set Y with $|{\sqsubset}^{-1}(V \sqcap Y)| = 1$ for every $V \sqsubset X$. By Lemma 2.6.4 (a) and (g) $Y' = Y \sqcap \bigsqcup X$ is an $\mathcal{M}|_{\mathbb{W}}$-set, which still has the property $|{\sqsubset}^{-1}(V \sqcap Y')| = 1$ for every $V \sqsubset X$.

This shows $\mathcal{M} \vDash \mathrm{ZFC}^- \Rightarrow \mathcal{M}|_{\mathbb{W}} \vDash \mathrm{ZFC}$. Now use the Completeness Theorem. $\square$

2.8 Cardinal numbers

Mostly we will require $\mathcal{M}$ to satisfy Choice in this section. But to a certain degree, comparing the sizes of $\mathcal{M}$-sets is possible without assuming Choice. For instance, the relation $\preceq$ on M, where $X \preceq Y$ holds iff there is an injective $\mathcal{M}$-function $X \twoheadrightarrow Y$, evidently is reflexive and transitive. The next result concerns the symmetry of $\prec$.

The equivalence theorem

Theorem 2.8.1 (Cantor–Bernstein–Schröder)**.** *Assume $\mathcal{M}$ satisfies* Power set *or* Infinity. *If there are injective $X \twoheadrightarrow Y$ and $Y \twoheadrightarrow X$, there is a bijective $X \twoheadrightarrow Y$.*

Proof assuming Power set. Let $f\colon X \twoheadrightarrow Y$ and $g\colon Y \twoheadrightarrow X$ be injective. Because $h\colon \mathsf{P}(X) \twoheadrightarrow \mathsf{P}(X)$, $W \mapsto X \smallsetminus g[Y \smallsetminus f[W]]$, preserves $\sqsubseteq$ it has a fixed point Z by Example 2.2.21 and Lemma 2.2.22. Then $f|_Z \sqcup g^{-1}|_{X \smallsetminus Z}\colon X \twoheadrightarrow Y$ is bijective $\square$

Proof assuming Infinity. Let $f\colon X \twoheadrightarrow Y$ and $g\colon Y \twoheadrightarrow X$ be injective. Recursively let $X_{\underline{0}} = X$, $X'_{\underline{0}} = g[Y]$ and $X_{\alpha+\underline{1}} = (g \circ f)[X_\alpha]$, $X'_{\alpha+\underline{1}} = (g \circ f)[X'_\alpha]$ for $\alpha \in \mathbb{N}^{\mathcal{M}}$. Then $X_\alpha \sqsupseteq X'_\alpha \sqsupseteq X_{\alpha+\underline{1}} \sqsupseteq X'_{\alpha+\underline{1}}$ and with $Z_\alpha = X_\alpha \smallsetminus X'_\alpha$

$$(g \circ f)|_{Z_\alpha}\colon Z_\alpha \twoheadrightarrow Z_{\alpha+\underline{1}} ,$$

are bijective. Combining them yields a bijective $h = (g \circ f)|_Z \sqcup \mathrm{id}_{X \smallsetminus Z}\colon X \twoheadrightarrow g[Y]$, with $Z = \bigsqcup \left[Z_\alpha : \alpha \in \mathbb{N}^{\mathcal{M}}\right]$. All in all, $g^{-1} \circ h\colon X \twoheadrightarrow Y$ is a bijective $\mathcal{M}$-function. $\square$

Remark 2.8.2. Assuming $\mathcal{M}$ satisfies Power set and Choice, Theorem 2.8.1 will become a direct consequence of Corollary 2.8.10 and Theorem 2.9.2 below.

Cardinals

Convention 2.8.3. From now on, let $\mathcal{M}$ be a $\mathrm{ZF}^\circ \cup \mathtt{POW}$-universe.

Definition 2.8.4. *An $\mathcal{M}$-*cardinal *is an $\mathcal{M}$-ordinal κ such that for all $\alpha < \kappa$ there is no bijective $\mathcal{M}$-function $\alpha \twoheadrightarrow \kappa$.*

We write $\mathbb{K}$ for the set of all $\mathcal{M}$-cardinals.

Remark 2.8.5. With Lemma 2.2.18 it is clear that $\mathbb{K}$ is an $\mathcal{M}$-class.

Observe that for every $\mathcal{M}$-set X by Power set we have an $\mathcal{M}$-set

$$\mathrm{WO}(X) \;=\; \left[\prec \sqsubseteq \mathsf{P}(X * X) : \prec \text{ is a well-order}\right].$$

It is then not hard to see that the following set is an $\mathcal{M}$-class

$$\mathbb{O}_{\simeq X} \;=\; \left\{\alpha \in \mathbb{O} : \mathrm{type}_{\mathcal{M}}(X, \prec) = \alpha \text{ for some } \prec \sqsubseteq \mathrm{WO}(X)\right\}.$$

Definition 2.8.6. *The* cardinality *of an $\mathcal{M}$-set X that is* well-orderable, *i.e. for which $\mathbb{O}_{\simeq X}$ is non-empty, is defined as $|X| = \bigsqcap \mathbb{O}_{\simeq X}$.*

Lemma 2.8.7. *The cardinality $|X|$ of a well-orderable $\mathcal{M}$-set X is the least $\kappa \in \mathbb{O}$ such that there exists a bijective $\mathcal{M}$-function $X \twoheadrightarrow \kappa$. In particular, $|X| \in \mathbb{K}$.*

Proof. Clearly, $|X| \in \mathbb{O}$ by Lemma 2.3.10 and there is a bijective $X \twoheadrightarrow |X|$.

Conversely, every bijective $f\colon X \twoheadrightarrow \kappa$ with $\kappa \in \mathbb{O}$ induces a well-order $\prec$ on X with $x \prec y \Leftrightarrow f(x) < f(y)$ such that $\mathrm{type}_{\mathcal{M}}(X, \prec) = \kappa$. Hence, $|X| \leq \kappa$. ☐

Corollary 2.8.8. *Let X, Y be well-orderable $\mathcal{M}$-sets. Then $|X| = |Y|$ is equivalent to the existence of a bijective $\mathcal{M}$-function $X \twoheadrightarrow Y$.*

Proof. Use Lemma 2.8.7 and that inverses of bijective $\mathcal{M}$-functions are bijective. ☐

Theorem 2.8.9. *For all well-orderable $\mathcal{M}$-sets X, Y the following are equivalent:*

(a) $|X| \leq |Y|$

(b) There is an injective $X \twoheadrightarrow Y$.

(c) There is a surjective $Y \twoheadrightarrow X$ or $X = \varnothing$.

Proof. (a) $\Rightarrow$ (b) This is clear because of $|X| \leq |Y| \Leftrightarrow |X| \sqsubseteq |Y|$.

(b) $\Rightarrow$ (c) If $X \neq \varnothing$, then for each $f\colon X \twoheadrightarrow Y$ there exists some $g\colon Y \setminus f[X] \twoheadrightarrow X$. If f is injective, then $f^{-1} \sqcup g$ is a surjective $\mathcal{M}$-function $Y \twoheadrightarrow X$.

(c) $\Rightarrow$ (b) If $g\colon Y \twoheadrightarrow X$ is surjective, fix some well-order $\prec$ on Y. Then $f\colon X \twoheadrightarrow Y$ given by $U \mapsto \min_\prec g^{-1}[[U]]$ is injective because of $g \circ f = \mathrm{id}_X$.

(b) $\Rightarrow$ (a) If there is an injective $X \twoheadrightarrow Y$, there also is an injective $f\colon X \twoheadrightarrow |Y|$. Then $|X| \leq \mathrm{type}_{\mathcal{M}}(f[X], <) \leq |Y|$ in view of Lemmas 2.5.12 and 2.8.7. ☐

Corollary 2.8.10. *For all well-orderable $\mathcal{M}$-sets X, Y:*

$$X \preceq Y \text{ and } Y \preceq X \quad \Rightarrow \quad |X| = |Y|$$

Proof. This is immediate by Theorem 2.8.9. □

Lemma 2.8.11. *The natural numbers in $\mathcal{M}$ are the finite $\mathcal{M}$-cardinals.*

Proof. It is enough to show that $C = \{\alpha \in \mathbb{N}^{\mathcal{M}} : \gamma \simeq \alpha \text{ for some } \gamma < \alpha\}$ is empty. If not, let $\alpha = \min C$ and let $f\colon \gamma \twoheadrightarrow \alpha$ be a bijective $\mathcal{M}$-function with $\gamma < \alpha$. Since $\alpha \neq \underline{0}$, also $\gamma \neq \underline{0}$ and we obtain a bijective $\mathcal{M}$-function $(g \circ f)|_{\gamma-\underline{1}}\colon \gamma - \underline{1} \twoheadrightarrow \alpha - \underline{1}$ where $g\colon \alpha \twoheadrightarrow \alpha$ is the transposition of $\alpha - \underline{1}$ and $f(\gamma - \underline{1})$. Hence, $\alpha - \underline{1} \in C$, in contradiction to $\alpha = \min C$. □

Lemma 2.8.12. *All infinite $\mathcal{M}$-cardinals are limit ordinals in $\mathcal{M}$.*

Proof. By Remark 2.3.33 there is nothing to show if $\mathcal{M}$ does not satisfy Infinity, so let us suppose it does. It is enough to check $|\alpha + \underline{1}| = |\alpha|$ for all infinite $\alpha \in \mathbb{O}$. But this is indeed the case, since e.g. the $\mathcal{M}$-function $f\colon \alpha + \underline{1} \twoheadrightarrow \alpha$ with $f(\alpha) = \underline{0}$ and $f(\gamma) = \gamma + \underline{1}$ for all $\gamma < \omega$ and $f(\gamma) = \gamma$ for all $\omega \leq \gamma < \alpha$ is bijective. □

Lemma 2.8.13. *For every $\mathcal{M}$-set X of $\mathcal{M}$-cardinals $\bigsqcup X$ is an $\mathcal{M}$-cardinal.*

Proof. If we had $|\alpha| < \alpha$ for $\alpha = \bigsqcup X$, then by Lemma 2.3.13 there would be $\kappa \sqsubset X$ with $|\alpha| < \kappa \leq \alpha$, contradicting Corollary 2.8.10 and Lemma 2.8.7. □

Successor cardinals

Definition 2.8.14. *The* Hartogs number *of an $\mathcal{M}$-set X is defined as*

$$|X|^{+} \;=\; \bigsqcup[\alpha + \underline{1} : \alpha \in \mathbb{O} \textit{ with } \alpha \preceq X]\,.$$

Remark 2.8.15. The $\mathcal{M}$-class $\{\alpha + \underline{1} : \alpha \in \mathbb{O} \text{ with } \alpha \preceq X\}$ forms indeed an $\mathcal{M}$-set, as is easily implied by the observation that $\alpha \preceq X$ if and only if there exists an $\mathcal{M}$-subset Y of X and a well-order $<$ on Y such that $\operatorname{type}_{\mathcal{M}}(Y, <) = \alpha$.

Lemma 2.8.16. $|X|^{+} = \bigsqcap[\alpha \in \mathbb{O} : \alpha \npreceq X] \in \mathbb{K}$ *for all $\mathcal{M}$-sets X.*

Proof. Let $\kappa = \bigsqcap[\alpha \in \mathbb{O} : \alpha \npreceq X]$. Then $\kappa \leq |X|^{+}$ since $|X|^{+} \npreceq X$ by definition. If $\alpha < |X|^{+}$, then there exists $\beta \preceq X$ with $\alpha \leq \beta$, so $\alpha \preceq X$. This implies $|X|^{+} = \kappa$. Finally, $\kappa \in \mathbb{K}$, since otherwise $|\kappa| < \kappa$ and then $\kappa \simeq |\kappa| \preceq X$, which is absurd. □

Remark 2.8.17. If X is well-orderable, then $\kappa \npreceq X \Leftrightarrow \kappa > |X|$ by Theorem 2.8.9 for all $\mathcal{M}$-cardinals κ. So in this case, $|X|^{+}$ is the least $\mathcal{M}$-cardinal greater than $|X|$.

Corollary 2.8.18. $\mathbb{K}$ *is closed and unbounded in* $\mathbb{O}$.

Proof. Lemma 2.8.13 states that $\mathbb{K}$ is closed in $\mathbb{O}$. For each $\alpha \in \mathbb{O}$ we have with Lemma 2.8.16 that $\alpha < |\alpha|^{+} \in \mathbb{K}$. □

If $\mathcal{M}$ satisfies INFINITY, the $\mathcal{M}$-class $\mathbb{K}_\infty = \mathbb{K} \setminus \mathbb{N}^{\mathcal{M}}$ of infinite $\mathcal{M}$-cardinals is still closed and unbounded in $\mathbb{O}$. Its enumerator is the so-called "aleph function":

Definition 2.8.19. *If $\mathcal{M}$ satisfies* INFINITY, *write $\aleph$ for the enumerator of $\mathbb{K}_\infty$.*

Lemma 2.8.20. *If $\mathcal{M}$ satisfies* INFINITY, *$\aleph_\beta$ is the β-iterate of ω under $\kappa \mapsto \kappa^+$.*

Proof. Let $f = \mathrm{It}_{\kappa\mapsto\kappa^+,\omega}$. Then f is normal by Lemma 2.4.10 (a).

By Lemma 2.5.13 it is enough to check that the image of f is $\mathbb{K}_\infty$.

For the inclusion $\mathrm{img}(f) \subseteq \mathbb{K}_\infty$ by induction and Lemmas 2.8.13 and 2.8.16 it is enough to show that ω is an $\mathcal{M}$-cardinal. Assuming the contrary, $|\omega| < \omega$ would be a natural number in $\mathcal{M}$, so $|\omega|^+ = |\omega| + \underline{1} < \omega$ by Lemma 2.8.11, which is impossible.

If there were $\kappa \in \mathbb{K}_\infty \setminus \mathrm{img}(f)$, we will show by induction $\kappa > f(\beta)$ for all $\beta \in \mathbb{O}$, which is a contradiction since the image of f is unbounded in $\mathbb{O}$ by Lemma 2.3.25.

Indeed, $\kappa > \omega = f(\underline{0})$ by Corollary 2.3.32 and Lemma 2.8.12.

For $\beta \in \mathbb{O}_{+\underline{1}}$ with $\kappa > f(\beta - \underline{1})$ we have $\kappa \geq f(\beta - \underline{1})^+ = f(\beta) \neq \kappa$.

For $\beta \in \mathbb{O}_{\lim}$ with $\kappa > f(\alpha)$ for all $\alpha < \beta$ we have $\kappa \geq \bigsqcup f[\beta] = f(\beta) \neq \kappa$. □

Definition 2.8.21. *$\mathcal{M}$-cardinals of the form κ^+ with $\kappa \in \mathbb{K}$ are* successor cardinals *and infinite $\mathcal{M}$-cardinals that are not successor cardinals are* limit cardinals in $\mathcal{M}$.

Write $\mathbb{K}_{+\underline{1}}$ for the $\mathcal{M}$-class of all successor cardinals in $\mathcal{M}$.

Write $\mathbb{K}_{\lim}$ for the $\mathcal{M}$-class of all limit cardinals in $\mathcal{M}$.

Remark 2.8.22. All non-zero natural numbers in $\mathcal{M}$ are successor cardinals.

The infinite successor cardinals are the $\aleph_\beta$ where β is a successor ordinal.

The limit cardinals are the $\aleph_\beta$ where β is a limit ordinal.

2.9 Well-orderable sets

Theorem 2.9.1 (Well-ordering theorem)**.** *An $\mathcal{M}$-set X is well-orderable iff there is an $\mathcal{M}$-function $f\colon \mathsf{P}(X) \smallsetminus [\varnothing] \twoheadrightarrow X$ with $f(Y) \in Y$ for all non-empty $Y \sqsubseteq X$.*

Proof. ⇒ Choose an $\mathcal{M}$-relation $<$ that well-orders X. Then the rule $V \mapsto \min_< V$ yields an $\mathcal{M}$-function $\mathsf{P}(X) \smallsetminus [\varnothing] \twoheadrightarrow X$.

⇐ It suffices to find an $\mathcal{M}$-bijection $\gamma \twoheadrightarrow X$ for some $\mathcal{M}$-ordinal γ. Taking some $\mathcal{M}$-set Z with $Z \notin X$,Theorem 2.4.4 yields an $\mathcal{M}$-class function $h\colon \mathbb{O} \twoheadrightarrow M$ with

$$h(\alpha) \;=\; \begin{cases} f(X \smallsetminus h[\alpha]) & \text{if } X \not\sqsubseteq h[\alpha], \\ Z & \text{otherwise.} \end{cases}$$

For all $\alpha < \beta$ in $C = [\delta \in \mathbb{O} : h(\delta) \in X]$ we have $h(\alpha) \in h[\beta]$ and $h(\beta) \in X \smallsetminus h[\beta]$, hence $h(\alpha) \neq h(\beta)$. This proves that $h|_C\colon C \to \in^{-1}(X)$ is injective and applying Lemma 2.2.11 to $(h|_C)^{-1}$ shows that $h|_C$ forms an $\mathcal{M}$-set $\tilde{h}\colon \gamma \twoheadrightarrow X$. It is not hard to see that $\gamma = \min\{\delta \in \mathbb{O} : h(\delta) \notin X\}$. In particular, $h(\gamma) \notin X$ such that $X \sqsubseteq h[\gamma]$. Hence, $\tilde{h}\colon \gamma \twoheadrightarrow X$ is bijective, which finishes the proof. □

There are many equivalent formulations of CHOICE. To collect just a few of them:

Theorem 2.9.2. *The following are equivalent:*

(i) *$\mathcal{M}$ satisfies* CHOICE.

(ii) $\ast_{i \sqsubset I} X_i \neq \varnothing$ *for all* $\langle X_i \rangle_{i \sqsubset I}$ *with* $X_i \neq \varnothing$.

(iii) *Every $\mathcal{M}$-set is well-orderable.*

(iv) ZORN'S LEMMA. *For each $\mathcal{M}$-set X each $\mathcal{M}$-relation $<$ on X has a $<$-maximal element, if $<$ partially orders X and every* $<$-chain in X *(i.e. $\mathcal{M}$-subset of X that is totally ordered by $<$) has an upper bound in X w.r.t. $<$.*

(v) *Every surjective* $p\colon X \twoheadrightarrow Y$ *has a* right inverse $s\colon Y \twoheadrightarrow X$, *i.e.* $p \circ s = \mathrm{id}_Y$.

Proof. (i) ⇒ (ii) Obviously, $X = [[i] * X_i : i \sqsubset I]$ is an $\mathcal{M}$-set of pairwise disjoint non-empty $\mathcal{M}$-sets. Hence, there is an $\mathcal{M}$-set Y such that $V \sqcap Y$ has exactly one $\mathcal{M}$-element for all $V \sqsubset X$. It follows that $Y \sqcap \bigsqcup X \sqsubset \ast_{i \sqsubset I} X_i$.

(ii) ⇒ (iii) Let Z be an arbitrary $\mathcal{M}$-set. Take for X the identity on $I = \mathsf{P}(Z) \setminus [\varnothing]$ to obtain $f\colon \mathsf{P}(Z) \setminus [\varnothing] \to M$ with $f(Y) \sqsubset Y$ for all $Y \sqsubseteq Z$. Now use Theorem 2.9.1 to conclude that Z is well-orderable.

(iii) ⇒ (iv) Let $\prec$ be a well-order on X. To arrive at a contradiction, we assume $<$ has no maximal elements. Let K be the $\mathcal{M}$-set of $<$-chains in X and let $s\colon K \twoheadrightarrow X$ associate with each $<$-chain its upper bound w.r.t. $<$ that is minimal w.r.t. $\prec$.

By Theorem 2.4.4 we get an $\mathcal{M}$-class function $f\colon \mathbb{O} \to M$ with $f(\alpha) = s(f[\alpha])$ for each $\alpha \in \mathbb{O}$ with $f[\alpha] \in K$. Indeed, we will then have $f[\alpha] \in K$ for all $\alpha \in \mathbb{O}$, assuming by induction for all $\beta, \gamma < \alpha$ that $\beta < \gamma \Rightarrow f(\beta) < f(\gamma)$. This assumption is justified, since for all $\beta < \alpha$ we have $f(\beta) < s(f[\alpha]) = f(\alpha)$ because $f(\beta) \sqsubset f[\alpha]$ and $<$ has no maximal elements.

All in all, this shows that f is a homomorphism $(\mathbb{O}, <) \to (\sqsubset^{-1}(X), <)$. Hence, it is injective. So Lemma 2.2.11 yields the contradiction that $\mathbb{O}$ forms an $\mathcal{M}$-set.

(iv) ⇒ (v) Let $p\colon X \twoheadrightarrow Y$ be surjective. Consider the $\mathcal{M}$-set

$$S \;=\; \left[Z \overset{s}{\twoheadrightarrow} X : Z \sqsubseteq Y \text{ and } p \circ s = \mathrm{id}_Z \right]$$

equipped with the $\mathcal{M}$-relation $\sqsubsetneq$, which clearly is a partial order and satisfies $\bigsqcup K \sqsubset S$ for every $\sqsubsetneq$-chain K in S. It follows that there is maximal element $s\colon Z \twoheadrightarrow X$ in S. We must have $Z = Y$, since otherwise there would be $y \sqsubset Y \setminus Z$ and then $x \sqsubset p^{-1}[[y]]$ due to the surjectivity of p such that $s \sqsubsetneq s \sqcup [\langle y, x \rangle] \sqsubset S$.

(v) ⇒ (i) Let X be an $\mathcal{M}$-set such that the $\mathcal{M}$-elements V of X are non-empty and pairwise disjoint. Define $Z = \bigsqcup X$ and let $p\colon Z \twoheadrightarrow X$ be the surjective $\mathcal{M}$-function sending each $U \sqsubset Z$ to the uniquely determined $V \sqsubset X$ with $U \sqsubset V$. Choose a right inverse $s\colon X \twoheadrightarrow Z$ of p. Then for $Y = \mathrm{img}(s)$ it is $V \sqcap Y = [s[V]]$ for all $V \sqsubset X$. □

Corollary 2.9.3. *Assuming $\mathcal{M}$ satisfies* CHOICE, *every $\mathcal{M}$-equivalence relation $\sim$ on X has a* system of representatives $\langle s_U \rangle_{U \sqsubset X/_\sim}$, *i.e.* $s_U \sqsubset U$ *for all* $U \in X/_\sim$.

Proof. Due to Theorem 2.9.2 (v) the canonical surjective $\mathcal{M}$-function $f_\sim : X \twoheadrightarrow X/_\sim$ (see Lemma 2.2.15) has a right inverse $\langle s_U \rangle_{U \in X/_\sim}$. □

2.10 Cardinal arithmetic

Convention 2.10.1. From now on, we assume $\mathcal{M}$ to be a ZFC$^-$-universe. Recall that this means that $\mathcal{M}$ satisfies all ZFC axioms except possibly Regularity.

Definition 2.10.2. Sums*,* products*, and* powers *of $\mathcal{M}$-cardinals κ and λ are*

$$\kappa \oplus \lambda \;=\; \big|\kappa \mathbin{\dot\sqcup} \lambda\big|, \qquad \kappa \otimes \lambda \;=\; \big|\kappa * \lambda\big|, \qquad \kappa^\lambda \;=\; \big|[\lambda \twoheadrightarrow \kappa]\big|.$$

More generally, define for families $\langle \kappa_i \rangle_{i \in I}$ of $\mathcal{M}$-cardinals:

$$\bigoplus_{i \in I} \kappa_i \;=\; \Big|\dot{\bigsqcup_{i \in I}} \kappa_i\Big|, \qquad \bigotimes_{i \in I} \kappa_i \;=\; \Big|\mathop{\ast}_{i \in I} \kappa_i\Big|$$

Remark 2.10.3. For all $\mathcal{M}$-sets X and Y it is easy to check the formulas

$$|X| \oplus |Y| \;=\; \big|X \mathbin{\dot\sqcup} Y\big|, \qquad |X| \otimes |Y| \;=\; \big|X * Y\big|, \qquad |X|^{|Y|} \;=\; \big|[Y \twoheadrightarrow X]\big|.$$

Moreover, $|\mathsf{P}(X)| = \underline{2}^{|X|}$, since the $\mathcal{M}$-function $\mathsf{P}(X) \twoheadrightarrow [X \twoheadrightarrow \underline{2}]$ mapping $Y \sqsubseteq X$ to the characteristic $\mathcal{M}$-function of Y is bijective. Finally, with the help of Choice, it is straightforward to verify for all families $\langle X_i \rangle_{i \in I}$ of $\mathcal{M}$-sets the identities:

$$\bigoplus_{i \in I} |X_i| \;=\; \Big|\dot{\bigsqcup_{i \in I}} X_i\Big|, \qquad \bigotimes_{i \in I} |X_i| \;=\; \Big|\mathop{\ast}_{i \in I} X_i\Big|$$

Exercise 2.10.4. Show that $\alpha \oplus \beta = \alpha + \beta$ and $\alpha \otimes \beta = \alpha \cdot \beta$ and $\alpha^\beta = \alpha^{(\beta)}$ for all natural numbers α and β in $\mathcal{M}$.

Exercise 2.10.5. For all $\kappa, \lambda, \mu, \nu \in \mathbb{K}$, $\langle \kappa_i \rangle_{i \in I}, \langle \lambda_i \rangle_{i \in I} \in \{I \twoheadrightarrow \mathbb{K}\}$, $\bigcirc \in \{\oplus, \otimes\}$:

(i) $\underline{0} \oplus \kappa = \kappa$

(ii) $\underline{0} \otimes \kappa = \underline{0}$

(iii) $\underline{1} \otimes \kappa = \kappa$

(iv) $\kappa^{\underline{0}} = \underline{1} = \underline{1}^\kappa$

(v) $\underline{0} < \kappa \implies \underline{0}^\kappa = \underline{0}$

(vi) $\kappa \bigcirc \lambda = \lambda \bigcirc \kappa$

(vii) $(\kappa \bigcirc \lambda) \bigcirc \mu = \kappa \bigcirc (\lambda \bigcirc \mu)$

(viii) $\kappa \otimes (\lambda \oplus \mu) = (\kappa \otimes \lambda) \oplus (\kappa \otimes \mu)$

(ix) $(\kappa \otimes \lambda)^\mu = \kappa^\mu \otimes \lambda^\mu$

(x) $\kappa^{\lambda \oplus \mu} = \kappa^\lambda \otimes \kappa^\mu$

(xi) $(\kappa^\lambda)^\mu = \kappa^{\lambda \otimes \mu}$

(xii) $\kappa \le \lambda \;\&\; \mu \le \nu \implies \kappa \bigcirc \mu \le \lambda \bigcirc \nu$

(xiii) $\kappa \le \lambda \;\&\; \underline{0} < \mu \le \nu \implies \kappa^\mu \le \lambda^\nu$

(xiv) $\bigoplus_{\alpha < \lambda} \kappa = \lambda \otimes \kappa$

(xv) $\bigotimes_{\alpha < \lambda} \kappa = \kappa^\lambda$

(xvi) $(\bigotimes_{i \in I} \kappa_i)^\mu = \bigotimes_{i \in I} \kappa_i^\mu$

(xvii) $\kappa^{\bigoplus_{i \in I} \lambda_i} = \bigotimes_{i \in I} \kappa^{\lambda_i}$

(xviii) $\kappa_i \le \lambda_i \,\forall\, i \in I \Rightarrow \bigoplus_{i \in I} \kappa_i \le \bigoplus_{i \in I} \lambda_i$

(xix) $\kappa_i \le \lambda_i \,\forall\, i \in I \Rightarrow \bigotimes_{i \in I} \kappa_i \le \bigotimes_{i \in I} \lambda_i$

The anti-lexicographic order $<$ on $\mathbb{O}*\mathbb{O}$, which restricts to the "natural" well-orders on the cartesian products $\alpha*\beta$ for $\alpha,\beta\in\mathbb{O}$ (see Problem Set 4, Exercise 4) is not set-like. The following alternative order $\prec$ however is:

Lemma 2.10.6. *Let $\prec$ be the $\mathcal{M}$-class relation on $D=\mathbb{O}*\mathbb{O}$ given by*

$$\begin{aligned}\langle\alpha,\beta\rangle\prec\langle\alpha',\beta'\rangle \quad\Leftrightarrow\quad &\big(\max\left[\alpha,\beta\right]<\max\left[\alpha',\beta'\right] \;\; \textit{or}\\ &\big(\max\left[\alpha,\beta\right]=\max\left[\alpha',\beta'\right] \;\; \textit{and} \;\; \langle\alpha,\beta\rangle<\langle\alpha',\beta'\rangle\big)\big)\,.\end{aligned}$$

Then the following hold:

(i) $\prec$ is a set-like well-order with transitive collapse $t\colon D\to\mathbb{O}$.

(ii) $D_{\prec\langle\alpha,\underline{0}\rangle}$ forms the $\mathcal{M}$-set $\alpha*\alpha$ for each $\alpha\in\mathbb{O}$.

(iii) $t[\lambda*\lambda]=\lambda$ for each infinite $\lambda\in\mathbb{K}$.

Proof. (i) It is not hard to check that $\prec$ is a well-order on D, which is set-like because of $D_{\prec X}\subseteq\,\Subset^{-1}(\gamma*\gamma)$ for all $X=\langle\alpha,\beta\rangle\in D$ with $\gamma=\max\left[\alpha,\beta\right]$.

According to Corollary 2.5.10 we have $t(X)=\mathrm{type}_{\mathcal{M}}(X,\prec)\in\mathbb{O}$ for all $X\in D$. It only remains to show that every $\alpha\in\mathbb{O}$ belongs to the image of t. Since the $\mathcal{M}$-class function $g\colon\mathbb{O}\to D$ given by $\alpha\mapsto\langle\alpha,\underline{0}\rangle$ is a morphism $(\mathbb{O},<)\to(D,\prec)$, $f=t\circ g$ is an endomorphism of $(\mathbb{O},<)$. Hence, $\alpha\leq f(\alpha)\in\mathrm{img}(t)$ by Lemma 2.3.24. Given that $\mathrm{img}(t)$ is transitive, this proves $\alpha\in\mathrm{img}(t)$.

(ii) This is obvious.

(iii) Since $t(\langle\alpha,\underline{0}\rangle)=t\big[\big[D_{\prec\langle\alpha,\underline{0}\rangle}\big]\big]$ by (ii) we get $\alpha\leq f(\alpha)=t[\alpha*\alpha]$ for all $\alpha\in\mathbb{O}$. Actually, f is normal because for limit ordinals β in $\mathcal{M}$ it is

$$f(\beta)\;=\;t[\beta*\beta]\;=\;\bigsqcup\nolimits_{\alpha<\beta}t[\alpha*\alpha]\;=\;\bigsqcup\nolimits_{\alpha<\beta}f(\alpha)\,.$$

We need to show that the $\mathcal{M}$-class $C=\big\{\lambda\in\mathbb{K}_\infty : f(\lambda)\neq\lambda\big\}$ is empty. Assume not and let $\lambda=\min C$. Then $\omega\leq\lambda<f(\lambda)$.

On the one hand, we cannot have $\lambda=\omega$, since for $\alpha<\omega$ we have the inequality $|f(\alpha)|=|t[\alpha*\alpha]|=|\alpha*\alpha|=|\alpha|\cdot|\alpha|<\omega$, so $f(\alpha)<\omega$, hence $f(\omega)\leq\omega$.

On the other hand, we also cannot have $\omega<\lambda$, since then $\lambda<f(\lambda)=\bigsqcup_{\omega\leq\alpha<\lambda}f(\alpha)$ by Lemma 2.8.12, i.e. there would be $\omega\leq\alpha<\lambda$ with $\lambda<f(\alpha)$. But this would give the absurd $\lambda\leq|f(\alpha)|=|\alpha|\leq\alpha$ where the equality is due to the choice of λ. $\square$

Theorem 2.10.7. *For all $\kappa,\lambda\in\mathbb{K}$ with infinite λ we have*

$$\begin{aligned}\kappa\oplus\lambda\;&=\;\max\left[\kappa,\lambda\right],\\ \kappa\otimes\lambda\;&=\;\max\left[\kappa,\lambda\right] && \textit{if } \kappa\neq 0,\\ \kappa^\lambda\;&=\;\underline{2}^\lambda && \textit{if } \underline{2}\leq\kappa\leq\underline{2}^\lambda.\end{aligned}$$

Proof. We may assume $\underline{0}<\kappa\leq\lambda$. Then Exercise 2.10.5 yields the inequalities

$$\begin{gathered}\lambda\;=\;\underline{0}\oplus\lambda\;\leq\;\kappa\oplus\lambda\;\leq\;\lambda\oplus\lambda\;=\;\underline{2}\otimes\lambda\;\leq\;\lambda\otimes\lambda\,,\\ \lambda\;=\;\underline{1}\otimes\lambda\;\leq\;\kappa\otimes\lambda\;\leq\;\lambda\otimes\lambda\,,\\ \underline{2}^\lambda\;\leq\;\kappa^\lambda\;\leq\;(\underline{2}^\lambda)^\lambda\;=\;\underline{2}^{\lambda\otimes\lambda}\,.\end{gathered}$$

So we just need $\lambda\otimes\lambda=\lambda$, which follows from Lemma 2.10.6 (iii) since t is bijective. $\square$

Lemma 2.10.8 (König's Theorem). *Let $\langle \kappa_i \rangle_{i \in I}$ and $\langle \lambda_i \rangle_{i \in I}$ be two families of $\mathcal{M}$-cardinals with $\kappa_i < \lambda_i$ for all $i \in I$. Then the following strict inequality holds:*

$$\bigoplus_{i \in I} \kappa_i < \bigotimes_{i \in I} \lambda_i .$$

Proof. According to Theorem 2.8.9 it suffices to show that no $\mathcal{M}$-function $f \colon X \twoheadrightarrow Y$ with $X = \dot{\bigsqcup}_{i \in I} \kappa_i$ and $Y = \mathop{\huge\ast}_{i \in I} \lambda_i$ can be surjective. Denote by p_i the canonical projection $Y \twoheadrightarrow \lambda_i$ and define $Y_i = (p_i \circ f)\big[[i]*\kappa_i\big]$. We then clearly have $|Y_i| \leq \kappa_i < \lambda_i$. In particular, $\lambda_i \setminus Y_i \neq \varnothing$ and according to Theorem 2.9.2 then also $\mathop{\huge\ast}_{i \in I} \lambda_i \setminus Y_i \neq \varnothing$. Because of $\mathop{\huge\ast}_{i \in I} \lambda_i \setminus Y_i \sqsubseteq Y \setminus \operatorname{img}(f)$ it follows that f is not surjective. $\square$

Lemma 2.10.9 (Cantor's Theorem). *$\kappa < \underline{2}^{\kappa}$ for all $\mathcal{M}$-cardinals κ.*

Consequently, $\kappa^{\kappa} = \underline{2}^{\kappa}$ for all infinite $\mathcal{M}$-cardinals κ.

Proof. Exercise 2.10.5 and Lemma 2.10.8 directly give $\kappa = \bigoplus_{\alpha < \kappa} \underline{1} < \bigotimes_{\alpha < \kappa} \underline{2} = \underline{2}^{\kappa}$. The last part then follows from the first with Theorem 2.10.7. $\square$

Proof without Choice. Each $f \colon \kappa \twoheadrightarrow \mathsf{P}(\kappa)$ is not surjective, since $[x \in \kappa : x \notin f(x)]$ does not belong to the image of f. Now use Remark 2.10.3 and Theorem 2.8.9. $\square$

Lemma 2.10.10. *For every family $\langle \kappa_i \rangle_{i \in \lambda}$ of $\mathcal{M}$-cardinals with $\kappa_i > \underline{1}$ for all $i \in \lambda$ the following inequality holds:*

$$\bigoplus_{i \in \lambda} \kappa_i \leq \bigotimes_{i \in \lambda} \kappa_i .$$

Proof. Abbreviate $\sigma = \bigoplus_{i \in \lambda} \kappa_i$ and $\rho = \bigotimes_{i \in \lambda} \kappa_i$.

Let us first treat the case where λ is infinite. We have $\lambda < \underline{2}^{\lambda} = \bigotimes_{i \in \lambda} \underline{2} \leq \rho$ in view of Lemma 2.10.9 and Exercise 2.10.5. Theorem 2.10.7 therefore implies $\lambda \otimes \rho = \rho$. By Theorem 2.9.2 (ii) $\lambda * \mathop{\huge\ast}_{i \in \lambda} \kappa_i \twoheadrightarrow \dot{\bigsqcup}_{i \in \lambda} \kappa_i$ given by $\langle i, f \rangle \mapsto \langle i, f_i \rangle$ is surjective. Thus we get $\sigma \leq \lambda \otimes \rho = \rho$ as desired.

Let us now deal with finite λ. For $\lambda \leq \underline{1}$ all is clear. For $\lambda > \underline{1}$ let $\sigma' = \bigoplus_{i \in \lambda - \underline{1}} \kappa_i$ and $\rho' = \bigotimes_{i \in \lambda - \underline{1}} \kappa_i$. We assume $\sigma' \leq \rho'$ and $\rho' > \underline{1}$ by induction. Then

$$\rho = \rho' \otimes \kappa_{\lambda - \underline{1}} > \underline{1} \quad \text{and} \quad \sigma = \sigma' \oplus \kappa_{\lambda - \underline{1}} \leq \rho' \oplus \kappa_{\lambda - \underline{1}} \leq \rho' \otimes \kappa_{\lambda - \underline{1}} = \rho .$$

Here, the last inequality holds by Theorem 2.10.7 if ρ' or $\kappa_{\lambda - \underline{1}}$ is infinite. If both are finite, calculate with the natural numbers $m = \rho' - \underline{2}$ and $n = \kappa_{\lambda - \underline{1}} - \underline{2}$ in $\mathcal{M}$. $\square$

Lemma 2.10.11. *For all families $\langle \kappa_i \rangle_{i \in \lambda}$ of $\mathcal{M}$-cardinals we have*

$$\bigoplus_{i \in \lambda} \kappa_i \leq \lambda \otimes \bigsqcup_{i \in \lambda} \kappa_i .$$

We even have equality, if λ is infinite and $\kappa_i > \underline{0}$ for all $i \in \lambda$.

Proof. Abbreviate $\kappa = \bigsqcup_{i \sqsubset \lambda} \kappa_i$ and $\sigma = \bigoplus_{i \sqsubset \lambda} \kappa_i$. The observation $\dot{\bigsqcup}_{i \sqsubset \lambda} \kappa_i \sqsubseteq \lambda * \kappa$ gives the inequality. Since $\kappa_i \leq \sigma$ for all i, it is $\kappa \leq \sigma$. If moreover $\kappa_i > \underline{0}$ for all i, then $\lambda = \bigoplus_{i \sqsubset \lambda} \underline{1} \leq \sigma$, so $\lambda \otimes \kappa = \max[\lambda, \kappa] \leq \sigma$ for infinite λ by Theorem 2.10.7. □

The slogan "Countable unions of countable $\mathcal{M}$-sets are countable." can be seen as a special case of Lemma 2.10.11. More generally, we obtain the following result:

Corollary 2.10.12. *For infinite $\mathcal{M}$-cardinals κ and all families $\langle X_i \rangle_{i \sqsubset I}$ we have:*

$$|I| \leq \kappa \text{ and } |X_i| \leq \kappa \text{ for all } i \sqsubset I \implies \left| \bigsqcup_{i \sqsubset I} X_i \right| \leq \kappa\,.$$

Proof. Let $\lambda = |I|$ and $\kappa_i = |X_i|$. Then with Lemma 2.10.11 and Theorem 2.10.7

$$|\textstyle\bigsqcup_{i \sqsubset I} X_i| \;\leq\; |\bigoplus_{i \sqsubset \lambda} \kappa_i| \;\leq\; \lambda \otimes \bigsqcup_{i \sqsubset \lambda} \kappa_i \;\leq\; \kappa \otimes \kappa \;=\; \kappa\,.$$

□

Lemma 2.10.13. *For all families $\langle \kappa_i \rangle_{i \sqsubset \lambda}$ of $\mathcal{M}$-cardinals we have*

$$\bigotimes_{i \sqsubset \lambda} \kappa_i \;\leq\; \Big(\bigsqcup_{i \sqsubset \lambda} \kappa_i \Big)^{\lambda}.$$

We even have equality, if λ is infinite and $\kappa_i > \underline{0}$ for all $i \sqsubset \lambda$ and $\langle \kappa_i \rangle_{i \sqsubset \lambda}$ is increasing.

Proof. Abbreviate $\kappa = \bigsqcup_{i \sqsubset \lambda} \kappa_i$. The inequality holds because of $\mathop{\Large *}_{i \sqsubset \lambda} \kappa_i \sqsubseteq [\lambda \rightharpoonup \kappa]$. If λ is infinite, take an $\mathcal{M}$-bijection $f\colon \lambda * \lambda \rightharpoonup \lambda$, which is possible by Theorem 2.10.7. Now λ is the disjoint union of the pairwise disjoint $\mathcal{M}$-sets $X_j = f\big[[j] * \lambda\big]$ with $j \sqsubset \lambda$. Because of $|X_j| = \lambda$, it follows that X_j must be unbounded in λ, hence $\bigsqcup_{i \sqsubset X_j} \kappa_i = \kappa$ since $\langle \kappa_i \rangle_{i \sqsubset \lambda}$ is increasing. So $\kappa^\lambda = \bigotimes_{j \sqsubset \lambda} \bigsqcup_{i \sqsubset X_j} \kappa_i \leq \bigotimes_{j \sqsubset \lambda} \bigotimes_{i \sqsubset X_j} \kappa_i = \bigotimes_{i \sqsubset \lambda} \kappa_i$, the inequality holding since $\kappa_i \leq \bigotimes_{i \sqsubset X_j} \kappa_i$ for all $i \sqsubset X_j$ given that $\kappa_i > \underline{0}$ for all i. □

Cofinality

When trying to evaluate cardinal powers κ^λ for infinite $\mathcal{M}$-cardinals κ and λ, one naturally comes across the notion of cofinality. Understanding cofinal $\mathcal{M}$-subsets only is relevant for non-zero limit ordinals, in which case cofinal simply means unbounded.

Definition 2.10.14. *If $X \sqsubseteq Y$ are $\mathcal{M}$-subsets of $\mathbb{O}$, we say that X is* cofinal in Y *if for every $\gamma \sqsubset Y$ there exists some $\delta \sqsubset X$ with $\gamma \leq \delta$.*

The cofinality *of $\alpha \in \mathbb{O}$ is the least cardinality of a cofinal $\mathcal{M}$-set in α, formally*

$$\operatorname{cof}(\alpha) \;=\; \bigsqcap \big[\kappa \in \mathbb{K} : \kappa = |X| \text{ for some } \mathcal{M}\text{-set } X \sqsubseteq \alpha \text{ with } X \text{ cofinal in } \alpha\big]\,.$$

An infinite $\mathcal{M}$-cardinal κ is regular *if $\operatorname{cof}(\kappa) = \kappa$. Otherwise it is* singular.

Remark 2.10.15. Let X be an $\mathcal{M}$-subset of an $\mathcal{M}$-ordinal α.

If α is a limit ordinal, then X is cofinal in α iff X is unbounded in α.

If α is a successor, then X is cofinal in α iff X contains the maximum $\alpha - \underline{1}$ of α.

In particular, $\operatorname{cof}(\alpha) = \underline{1}$ iff α is a successor ordinal.

Example 2.10.16. ω is regular since every unbounded subset of ω is infinite.

Lemma 2.10.17. *Let $\alpha \in \mathbb{O}$ and $X \sqsubseteq \alpha$ with $\bigsqcup X = \alpha$. Then* $\mathrm{cof}(\alpha) \leq |X|$.

Proof. Just observe that $\bigsqcup X = \alpha$ means that X is unbounded in α. $\square$

Lemma 2.10.18. *Let X and Y be $\mathcal{M}$-subsets of an $\mathcal{M}$-ordinal α both cofinal in α. Then there is $f\colon X \twoheadrightarrow Y$ that preserves $\leq$ whose image is cofinal in Y.*

Proof. Take for $f\colon X \twoheadrightarrow Y$ the $\mathcal{M}$-function given by $x \mapsto \min\,[y \sqin Y : x \leq y]$. $\square$

Corollary 2.10.19. *Let α and β be $\mathcal{M}$-ordinals such that there exists an $\mathcal{M}$-function $f\colon \alpha \twoheadrightarrow \beta$ preserving $<$ whose image is cofinal in β. Then* $\mathrm{cof}(\alpha) = \mathrm{cof}(\beta)$.

Proof. For every cofinal X in α the image $f[X]$ is cofinal in β, hence $\mathrm{cof}(\beta) \leq \mathrm{cof}(\alpha)$.

Vice versa, for every cofinal Y in β by Lemma 2.10.18 there is $g\colon Y \twoheadrightarrow f[\alpha]$ with cofinal image. Then $(f^{-1} \circ g)[Y]$ is cofinal in α because $f^{-1}\colon f[\alpha] \twoheadrightarrow \alpha$ preserves $<$. This shows that we also have $\mathrm{cof}(\alpha) \leq \mathrm{cof}(\beta)$. $\square$

Remark 2.10.20. Let β be a non-zero limit ordinal in $\mathcal{M}$. Then $\mathrm{cof}(\aleph_\beta) = \mathrm{cof}(\beta)$. Since $\mathrm{cof}(\beta) \leq \beta \leq \aleph_\beta$, it follows that, if $\aleph_\beta$ is regular, then necessarily $\aleph_\beta = \beta$.

In particular, $\aleph_\omega$ is singular.

Lemma 2.10.21. *For all $\alpha \in \mathbb{O}$ the cofinality* $\mathrm{cof}(\alpha)$ *is the least $\lambda \in \mathbb{K}$ such that there is a normal $f\colon \lambda \twoheadrightarrow \alpha$ whose image is cofinal in α.*

Proof. Clearly, $\mathrm{cof}(\alpha) \leq \lambda$ if there exists $f\colon \lambda \twoheadrightarrow \alpha$ whose image is cofinal in α.

Now let $\lambda = \mathrm{cof}(\alpha)$. Note that $\lambda = \underline{1}$, if α is a successor, and in any case there exists $g\colon \lambda \twoheadrightarrow \alpha$ with cofinal image. Take for f the $\mathcal{M}$-function recursively given by

$$f(\beta) \;=\; \begin{cases} \max\,[g(\beta - \underline{1}), f(\beta - \underline{1}) + \underline{1}] & \text{if } \beta \text{ is a successor ordinal,} \\ \bigsqcup f[\beta] & \text{if } \beta \text{ is a limit ordinal.} \end{cases}$$

It follows by transfinite induction that f has all the properties claimed in the lemma. In particular, observe that $\bigsqcup f[\beta] < \alpha$ for every non-zero limit ordinal $\beta < \lambda$, given that $f[\beta] \sqsubseteq \alpha$ cannot be cofinal in α because of $|f[\beta]| \leq |\beta| \leq \beta < \lambda = \mathrm{cof}(\alpha)$. $\square$

Lemma 2.10.22. *For all $\alpha \in \mathbb{O}$ it is* $\mathrm{cof}(\mathrm{cof}(\alpha)) = \mathrm{cof}(\alpha) \leq |\alpha| \leq \alpha$.

Proof. The inequalities $\mathrm{cof}(\mathrm{cof}(\alpha)) \leq \mathrm{cof}(\alpha) \leq |\alpha| \leq \alpha$ hold since α is cofinal in α.

The inequality $\mathrm{cof}(\mathrm{cof}(\alpha)) \geq \mathrm{cof}(\alpha)$ follows from Lemma 2.10.18 because the composition of increasing $\mathcal{M}$-functions with cofinal image again has cofinal image. $\square$

Corollary 2.10.23. $\mathrm{cof}(\alpha)$ *is regular for every non-zero limit ordinal α in $\mathcal{M}$.*

Proof. We have $\mathrm{cof}(\mathrm{cof}(\alpha)) = \mathrm{cof}(\alpha)$ according to Lemma 2.10.22. Evidently, $\mathrm{cof}(\alpha)$ is not finite, since then by Remark 2.10.15 we would have $\mathrm{cof}(\alpha) = \mathrm{cof}(\mathrm{cof}(\alpha)) = \underline{1}$ or $\mathrm{cof}(\alpha) = \underline{0}$, so α would itself be a successor ordinal or $\underline{0}$. $\square$

Theorem 2.10.24. *For every $\kappa \in \mathbb{K}_\infty$ the cofinality* $\operatorname{cof}(\kappa)$ *is the least* $\lambda \in \mathbb{K}$ *such that there exists* $\langle \mu_\gamma \rangle_{\gamma<\lambda}$ *with* $\mathcal{M}$*-cardinals* $\mu_\gamma < \kappa$ *for all* $\gamma < \lambda$ *and* $\kappa = \bigoplus_{\gamma<\lambda} \mu_\gamma$.

Proof. Firstly, let $\lambda = \operatorname{cof}(\kappa)$. By Lemma 2.10.18 there is $f\colon \lambda \twoheadrightarrow \kappa$ with cofinal image. Take $\mu_\gamma = |f(\gamma)|$. Then clearly $\mu_\gamma < \kappa$ and in view of Lemma 2.8.12 $\kappa = \bigsqcup_{\gamma<\lambda} f(\gamma)$. Lemma 2.10.11 and Theorem 2.10.7 yield $\kappa \le \bigoplus_{\gamma<\lambda} \mu_\gamma \le \lambda \otimes \bigsqcup_{\gamma<\lambda} \mu_\gamma \le \lambda \otimes \kappa \le \kappa$.

Secondly, let $\lambda \in \mathbb{K}$ with $\lambda < \operatorname{cof}(\kappa)$. If $\kappa = \bigoplus_{\gamma<\lambda} \mu_\gamma$ with $\mathcal{M}$-cardinals $\mu_\gamma < \kappa$, Lemma 2.10.11 and Theorem 2.10.7 would yield $\kappa \le \lambda \otimes \bigsqcup_{\gamma<\lambda} \mu_\lambda \le \kappa$, so $\kappa = \bigsqcup_{\gamma<\lambda} \mu_\lambda$ because of $\lambda < \kappa$. Lemma 2.10.17 would then give the contradiction $\operatorname{cof}(\kappa) \le \lambda$. □

Corollary 2.10.25. *Let $\kappa \in \mathbb{K}_\infty$. Then κ is singular iff there is an $\mathcal{M}$-cardinal $\lambda < \kappa$ and a family $\langle \mu_\gamma \rangle_{\gamma<\lambda}$ of $\mathcal{M}$-cardinals with $\mu_\gamma < \kappa$ for all $\gamma < \lambda$ and $\kappa = \bigoplus_{\gamma<\lambda} \mu_\gamma$.*

Proof. $\Rightarrow$ Take $\lambda = \operatorname{cof}(\kappa) < \kappa$ and apply Theorem 2.10.24.

$\Leftarrow$ Theorem 2.10.24 implies $\operatorname{cof}(\kappa) \le \lambda < \kappa$. □

It follows that the $\mathcal{M}$-cardinals $\aleph_{\beta+\underline{1}}$ are regular:

Lemma 2.10.26. *Every infinite successor cardinal in $\mathcal{M}$ is regular.*

Proof. Let $\sigma \in \mathbb{K}$ and $\kappa = \sigma^+$ and $\lambda = \operatorname{cof}(\kappa)$. Theorem 2.10.24 allows us to write $\kappa = \bigoplus_{\gamma<\lambda} \mu_\gamma$ with $\mathcal{M}$-cardinals $\mu_\gamma < \kappa$, i.e. $\mu_\gamma \le \sigma$. If κ were singular, then $\lambda \le \sigma$, leading with Theorem 2.10.7 to the contradiction $\kappa \le \lambda \otimes \bigsqcup_{\gamma<\lambda} \mu_\gamma \le \sigma$. □

Let us write $\mathbb{K}_{\mathrm{reg}}$ for the $\mathcal{M}$-class consisting of all regular cardinals in $\mathcal{M}$.

Corollary 2.10.27. $\mathbb{K}_{\mathrm{reg}}$ *is unbounded in* $\mathbb{O}$.

Proof. This follows from Lemma 2.10.26 because $\mathbb{K}_{+\underline{1}}$ is unbounded in $\mathbb{O}$. □

Corollary 2.10.28. *The regular cardinals in $\mathcal{M}$ are precisely ω, all infinite successor cardinals in $\mathcal{M}$, and all regular $\mathcal{M}$-cardinals that are fixed points of $\aleph$.*

Proof. Use Example 2.10.16, Remarks 2.8.22 and 2.10.20, and Lemma 2.10.26 and observe that every $\mathcal{M}$-cardinal κ with $\aleph_\kappa = \kappa$ is a limit cardinal. □

Remark 2.10.29. Regular limit cardinals in $\mathcal{M}$ different from ω are referred to as *weakly inaccessible*. They are the regular fixed points of $\aleph$ in $\mathbb{K}$.

There are choices of $\mathcal{M}$ (if there are any at all) such that $\mathcal{M}$ does not have weakly inaccessible cardinals. Namely, if κ is the smallest weakly inaccessible $\mathcal{M}$-cardinal, then $\mathcal{M}|_{\mathbb{L}_\kappa}$ is a ZFC-universe without weakly inaccessible cardinals where $\mathbb{L}_\kappa$ is the $\mathcal{M}$-class of constructible sets of rank κ (to be constructed in Definition 2.11.5).

Cardinal exponentiation

The point of this subsection is to derive that in order to compute the powers κ^λ for infinite $\mathcal{M}$-cardinals κ and λ it is enough to know $\underline{2}^\nu$ and $\nu^{\mathrm{cof}(\nu)}$ for all $\nu \in \mathbb{K}_\infty$.

Lemma 2.10.30. $\kappa < \kappa^{\mathrm{cof}(\kappa)}$ *for all* $\kappa \in \mathbb{K}_\infty$.

Proof. For $\lambda = \mathrm{cof}(\kappa)$ there are according to Theorem 2.10.24 $\mathcal{M}$-cardinals $\mu_\gamma < \kappa$ with $\kappa = \bigoplus_{\gamma<\lambda} \mu_\gamma$ such that Lemma 2.10.8 yields $\kappa < \bigotimes_{\gamma<\lambda} \kappa = \kappa^\lambda$. □

The following result strengthens Lemma 2.10.9 for infinite $\mathcal{M}$-cardinals:

Lemma 2.10.31. $\kappa < \mathrm{cof}(\underline{2}^\kappa)$ *for all* $\kappa \in \mathbb{K}_\infty$.

Proof. For every family $\langle \mu_\gamma \rangle_{\gamma<\kappa}$ of $\mathcal{M}$-cardinals with $\mu_\gamma < \underline{2}^\kappa$ Lemma 2.10.8 and Theorem 2.10.7 show $\bigoplus_{\gamma<\kappa} \mu_\gamma < \bigotimes_{\gamma<\kappa} \underline{2}^\kappa = (\underline{2}^\kappa)^\kappa = \underline{2}^\kappa$. Now apply Theorem 2.10.24. □

Example 2.10.32. $\underline{2}^\omega \neq \aleph_\omega$ because of $\mathrm{cof}(\aleph_\omega) = \omega$.

So there are restrictions for what the cardinality $\underline{2}^\omega$ of the "continuum in $\mathcal{M}$" can be apart from the obvious limitation $\omega < \underline{2}^\omega$ due to Cantor's Theorem.

Lemma 2.10.33. *For all* $\kappa, \lambda \in \mathbb{K}_\infty$ *with* $\lambda < \mathrm{cof}(\kappa)$ *it is*

$$\kappa^\lambda \;=\; \max\left[\kappa,\, \bigsqcup\nolimits_{|\mu|=\mu<\kappa} \mu^\lambda\right].$$

Proof. Because of $\lambda < \mathrm{cof}(\kappa)$ we have the identity $[\lambda \rightharpoonup \kappa] = \bigsqcup_{\gamma<\kappa} [\lambda \rightharpoonup \gamma]$. So

$$\kappa^\lambda \;\leq\; \bigoplus\nolimits_{\gamma<\kappa} |\gamma|^\lambda \;\leq\; \kappa \otimes \bigsqcup\nolimits_{\gamma<\kappa} |\gamma|^\lambda \;\leq\; \kappa \otimes \kappa^\lambda \;=\; \kappa^\lambda$$

and thus $\kappa^\lambda = \max\left[\kappa,\, \bigsqcup_{|\mu|=\mu<\kappa} \mu^\lambda\right]$ by Lemma 2.10.11 and Theorem 2.10.7. □

Lemma 2.10.34 (Hausdorff's formula)**.** *For all* $\mu, \lambda \in \mathbb{K}_\infty$ *we have*

$$(\mu^+)^\lambda \;=\; \mu^+ \otimes \mu^\lambda\,.$$

Proof. Let $\kappa = \mu^+$.

By Lemma 2.10.26 we have $\mathrm{cof}(\kappa) = \kappa$ such that in case $\lambda < \kappa = \mathrm{cof}(\kappa)$ we get with Lemma 2.10.33 and Theorem 2.10.7 that $\kappa^\lambda = \max\left[\kappa,\, \bigsqcup_{|\sigma|=\sigma<\kappa} \sigma^\lambda\right] \leq \kappa \otimes \mu^\lambda \leq \kappa^\lambda$.

In case $\kappa \leq \lambda$ Lemma 2.10.9 and Theorem 2.10.7 show that $\mu < \kappa \leq \lambda < \underline{2}^\lambda = \mu^\lambda$. We then conclude once again with Theorem 2.10.7 that $\kappa \otimes \mu^\lambda = \mu^\lambda = \underline{2}^\lambda = \kappa^\lambda$. □

Lemma 2.10.35. *For all* $\kappa \in \mathbb{K}_{\mathrm{lim}}$ *and* $\lambda \in \mathbb{K}$ *with* $\mathrm{cof}(\kappa) \leq \lambda$ *we have*

$$\kappa^\lambda \;=\; \rho^{\mathrm{cof}(\kappa)} \quad \textit{with} \quad \rho \;=\; \bigsqcup\nolimits_{|\mu|=\mu<\kappa} \mu^\lambda\,.$$

Proof. Let $\sigma = \mathrm{cof}(\kappa)$. By Theorem 2.10.24 there is a family $\langle \mu_\gamma \rangle_{\gamma<\sigma}$ of $\mathcal{M}$-cardinals with $\underline{1} < \mu_\gamma < \kappa$ and $\kappa = \bigoplus_{\gamma<\sigma} \mu_\gamma$. Using Lemma 2.10.10 and Theorem 2.10.7 we then get $\kappa^\lambda \leq \left(\bigotimes_{\gamma<\sigma} \mu_\gamma\right)^\lambda = \bigotimes_{\gamma<\sigma} \mu_\gamma^\lambda \leq \bigotimes_{\gamma<\sigma} \rho = \rho^\sigma \leq (\kappa^\lambda)^\sigma = \kappa^\lambda$. □

All in all, we have the following rules for computing cardinal powers:

Theorem 2.10.36. *For all $\kappa, \lambda \in \mathbb{K}_\infty$ the following holds:*

$$\kappa^\lambda = \begin{cases} \mu^\lambda & \text{if } \lambda < \operatorname{cof}(\kappa) \text{ and if } \mu^\lambda \geq \kappa \text{ for some } |\mu| = \mu < \kappa, \\ \kappa & \text{if } \lambda < \operatorname{cof}(\kappa) \text{ and if } \mu^\lambda < \kappa \text{ for every } |\mu| = \mu < \kappa, \\ \kappa^{\operatorname{cof}(\kappa)} & \text{if } \operatorname{cof}(\kappa) \leq \lambda < \kappa \text{ and if } \mu^\lambda < \kappa \text{ for every } |\mu| = \mu < \kappa, \\ \underline{2}^\lambda & \text{if } \kappa \leq \lambda \quad . \end{cases}$$

Proof. The first case follows from $\mu^\lambda \leq \kappa^\lambda \leq (\mu^\lambda)^\lambda = \mu^\lambda$ with Theorem 2.10.7.

The second case follows for $\kappa = \mu^+$ from Lemma 2.10.34 and Theorem 2.10.7.

If κ is a limit cardinal, then in the second and third case $\kappa = \bigsqcup_{|\mu|=\mu<\kappa} \mu^\lambda$ such that the second case is due to Lemma 2.10.33 and the third case due to Lemma 2.10.35.

By Lemma 2.10.26 the third case cannot occur, if κ is a successor cardinal.

The fourth case follows with Theorem 2.10.7 and Lemma 2.10.9. □

Corollary 2.10.37. *For all $\kappa, \lambda \in \mathbb{K}_\infty$ the power κ^λ has the form κ or $\underline{2}^\lambda$ or $\nu^{\operatorname{cof}(\nu)}$ for some $\nu \in \mathbb{K}$ with $\operatorname{cof}(\nu) \leq \lambda < \nu \leq \kappa$.*

Proof. Let us assume $\kappa^\lambda \neq \kappa$ and $\kappa^\lambda \neq \underline{2}^\lambda$. We may choose $\nu \in \mathbb{K}_\infty$ minimal with the property $\nu^\lambda = \kappa^\lambda$. By Theorem 2.10.36 we then have $\nu^\lambda = \nu^{\operatorname{cof}(\nu)}$ with $\operatorname{cof}(\nu) \leq \lambda < \nu$, since neither $\nu \leq \lambda$, as this implies $\underline{2}^\lambda = \nu^\lambda = \kappa^\lambda$, nor $\mu < \nu \leq \mu^\lambda$ for some $\mu = |\mu|$, as this implies $\mu^\lambda = \nu^\lambda = \kappa^\lambda$, nor $\nu^\lambda = \nu$, as this implies $\kappa^\lambda = \nu^\lambda = \nu \leq \kappa < \kappa^\lambda$. □

Continuum hypothesis

For every $\mathcal{M}$-cardinal κ Cantor's Theorem tells us that $\underline{2}^\kappa$ is an $\mathcal{M}$-cardinal greater than κ. The *General Continuum Hypothesis (GCH)* postulates $\underline{2}^\kappa = \kappa^+$ for all infinite $\mathcal{M}$-cardinals κ, i.e. it insists that there is no $\mathcal{M}$-cardinal between κ and $\underline{2}^\kappa$.

Lemma 2.10.38 (Exponentiation under GCH)**.** *Assuming $\mu^+ = \underline{2}^\mu$ for all $\mu \in \mathbb{K}_\infty$, the powers of all $\kappa, \lambda \in \mathbb{K}_\infty$ can be computed as follows:*

$$\kappa^\lambda = \begin{cases} \kappa & \text{if } \lambda < \operatorname{cof}(\kappa)\,, \\ \kappa^+ & \text{if } \operatorname{cof}(\kappa) \leq \lambda < \kappa \quad , \\ \lambda^+ & \text{if } \kappa \leq \lambda \quad . \end{cases}$$

Proof. In case $\lambda < \operatorname{cof}(\kappa)$ we have $\kappa \leq \kappa^\lambda = \max\left[\kappa, \bigsqcup_{|\mu|=\mu<\kappa} \mu^\lambda\right] \leq \kappa$ by Lemma 2.10.33 and since $\mu^\lambda \leq \underline{2}^{\mu\otimes\lambda} = (\mu \otimes \lambda)^+ \leq \kappa$ for $|\mu| = \mu < \kappa$ by Theorem 2.10.7.

In case $\operatorname{cof}(\kappa) \leq \lambda < \kappa$ use $\kappa < \kappa^\lambda \leq \kappa^\kappa = \underline{2}^\kappa = \kappa^+$ by Lemmas 2.10.9 and 2.10.30.

Finally, in case $\kappa \leq \lambda$ we have $\kappa^\lambda = \underline{2}^\lambda = \kappa^+$ according to Theorem 2.10.7. □

Remark 2.10.39. Let $f^{\mathcal{M}} \colon \mathbb{K}_\infty \to \mathbb{K}_\infty$ be the $\mathcal{M}$-class function given by $\kappa \mapsto \underline{2}^\kappa$. Then $f^{\mathcal{M}}$ preserves $\leq$ and $\kappa < \operatorname{cof}(f^{\mathcal{M}}(\kappa))$ for all $\kappa \in \mathbb{K}_\infty$ by Lemma 2.10.31.

Remarkably, this turns out to be the only provable restriction for the value of $\underline{2}^\kappa$ for regular $\mathcal{M}$-cardinals κ, a result known as Easton's Theorem:

Given an S^{Set}-formula with two free variables defining in each ZFC-universe $\mathcal{M}$ an $\mathcal{M}$-class function $g^{\mathcal{M}} \colon \mathbb{K}_{\mathrm{reg}} \to \mathbb{K}_\infty$ that preserves $\leq$ with the property $\kappa < \operatorname{cof}(g^{\mathcal{M}}(\kappa))$ for all $\kappa \in \mathbb{K}_{\mathrm{reg}}$. Then there exists a ZFC-universe $\mathcal{M}$ with $f^{\mathcal{M}}|_{\mathbb{K}_{\mathrm{reg}}} = g^{\mathcal{M}}$.

2.11 Constructible sets

The next goal is to show that CHOICE and GCH are satisfied in the ZF-subuniverse $\mathcal{M}|_{\mathbb{L}}$ of any ZF-universe $\mathcal{M}$ where $\mathbb{L}$ is the smallest transitive $\mathcal{M}$-class that contains all $\mathcal{M}$-ordinals. We now turn to the somewhat technical construction of $\mathbb{L}$.

Convention 2.11.1. In this section, we assume that $\mathcal{M}$ is a $(\mathrm{ZF}^{\circ} \cup \mathtt{INF})$-universe.

Internal definability

Let $H = \bigcup_{n\in\mathbb{N}} H_n$ be recursively defined by $H_0 = \emptyset$ and $H_n = \mathcal{P}(H_{n-1})$ for all $n > 0$. Consider the canonical map $H \to \mathbb{W}_\omega = \mathbb{HF}$ which is given by the recursive rule

$$x \mapsto \underline{x} = \left[\underline{u} : u \in x\right].$$

This map is injective and – assuming now for simplicity $0 = \emptyset$ and $n + 1 = n \cup \{n\}$ for all $n \in \mathbb{N}$ – it extends the map $\mathbb{N} \to \mathbb{N}^{\mathcal{M}}$ described in Remark 2.3.34.

Example 2.11.2. For instance, we have $\underline{\{(0, 1), 2\}} = \left[\langle\underline{0}, \underline{1}\rangle, \underline{2}\right]$.

Without loss of generality we will assume that the vocabulary $S = S^{\mathrm{Set}}$ was chosen such that $S_V = \mathbb{N}$ and $S_L \cup S_R = \{\bot, \to, \bigwedge, \equiv, \epsilon\} \subseteq H$. This makes it possible to mimic inside $\mathcal{M}$ the definition of S to obtain an analogous vocabulary $S^{\mathcal{M}}$ in $\mathcal{M}$. Repeating the constructions of Definition 1.2.1 inside $\mathcal{M}$, we define $S^{\mathcal{M}}$-terms and $S^{\mathcal{M}}$-formulas in the obvious way. In particular, for every S-formula φ we have the corresponding $S^{\mathcal{M}}$-formula $\underline{\varphi}$. Note however that not every $S^{\mathcal{M}}$-formula needs to arise in this way. Namely, there are "more" $S^{\mathcal{M}}$-formulas than S-formulas whenever the canonical map $\mathbb{N} \to \mathbb{N}^{\mathcal{M}}$ is not bijective. After all, we can mimic the definition of $\models$ as presented in Definition 1.3.1 inside $\mathcal{M}$. Here, the following special case will be sufficient for what we need:

Definition 2.11.3. *For every $\mathcal{M}$-set X consider the $\mathcal{M}$-set*

$$F_X = \left[\langle\pi, f\rangle : \pi \textit{ is an } S^{\mathcal{M}}\textit{-formula and } f\colon \mathrm{fvar}(\pi) \twoheadrightarrow X\right].$$

Define an $\mathcal{M}$-function $F_X \twoheadrightarrow \underline{2}$, $\langle\pi, f\rangle \mapsto \pi^{X,f}$, such that $\pi^{X,f} = \underline{1}$ if and only if

- $\pi = x \underline{\equiv} y$ *and* $f(x) = f(y)$ *or,*
- $\pi = x \underline{\epsilon} y$ *and* $f(x) \sqsubset f(y)$ *or,*
- $\pi = (\varphi \underline{\to} \psi)$ *and* $\left(\varphi^{X,f|_{\mathrm{fvar}(\varphi)}} = \underline{0} \textit{ or } \psi^{X,f|_{\mathrm{fvar}(\psi)}} = \underline{1}\right)$ *or,*
- $\pi = \underline{\bigwedge}_x \varphi$ *and* $\underline{0} \not\sqsubset \left[\varphi^{X,f^x_u} : u \sqsubset X\right]$ *with* $f^x_u(x) = u$ *and* $f^x_u(y) = f(y)$ *for* $y \neq x$.

For $\gamma < \omega$ let $T_{X,\gamma}$ be the $\mathcal{M}$-set consisting of all $\langle\pi, x, a\rangle$ where π is an $S^{\mathcal{M}}$-formula, $x = \langle x_\alpha\rangle_{\alpha<\beta}$ is a family of pairwise distinct variables of $S^{\mathcal{M}}$ and $a = \langle a_\alpha\rangle_{\alpha<\beta-\gamma}$ is a family of $\mathcal{M}$-elements of X such that $\gamma \leq \beta < \omega$ and $\mathrm{fvar}(\pi) \sqsubseteq [x_\alpha : \alpha < \beta]$.

We will write $X \vDash \pi[x/a]$ *synonymously for* $\pi^{X, f^x_a|_{\mathrm{fvar}(\pi)}} = \underline{1}$ *where* $\langle \pi, x, a\rangle \mathrel{\varepsilon} T_{X,\underline{0}}$ *and* $f\colon [x_\alpha : \alpha < \beta] \twoheadrightarrow X$ *given as* $f^x_a(x_\alpha) = a_\alpha$ *for* $\alpha < \beta$.

An $\mathcal{M}$*-subset* Y *of* X *is said to be* definable in X *if there exists* $\langle \pi, x, a\rangle \mathrel{\varepsilon} T_{X,\underline{1}}$ *such that* $Y = X_{\pi,x,a}$ *where with* $\beta = \mathrm{dom}(x)$

$$X_{\pi,x,a} \;=\; \Big[b_{\beta-\underline{1}} \mathrel{\varepsilon} X : X \vDash \pi[x/b] \text{ for } b\colon \beta \twoheadrightarrow X \text{ with } b|_{\beta-\underline{1}} = a \Big] .$$

Actually, to justify this definition we need to prove that indeed there exists $\mathcal{M}$-sets F_X and $T_{X,\gamma}$ and $X_{\pi,x,a}$ as described above. This is clear by the following lemma, which holds even without knowing that $\mathcal{M}$ satisfies Power set:

Lemma 2.11.4 (Function set with finite domain)**.** *For all* $\mathcal{M}$*-sets* X *and* Y *with finite* X *there is an* $\mathcal{M}$*-set*

$$[X \twoheadrightarrow Y] \;=\; [f : f \text{ is an } \mathcal{M}\text{-function } X \twoheadrightarrow Y] .$$

Proof. For every finite $\mathcal{M}$-subset Y' of Y also $X * Y'$ is finite, so $\mathsf{P}(X * Y')$ exists due to Lemma 2.3.30 and then so does $[X \twoheadrightarrow Y']$ as an $\mathcal{M}$-subset. Since the image of every $X \twoheadrightarrow Y$ is finite, we can take $[X \twoheadrightarrow Y] = \bigsqcup_{Y' \sqsubseteq Y \text{ finite}} [X \twoheadrightarrow Y']$. □

Constructible sets

Definition 2.11.5. *Denote by* $\mathsf{D}\colon M \to M$ *the* $\mathcal{M}$*-class function given by*

$$\mathsf{D}(X) \;=\; \big[Y : Y \text{ is a definable } \mathcal{M}\text{-subset in } X \big] ,$$

i.e. $\mathsf{D}(X)$ *is the image of the* $\mathcal{M}$*-function* $p_X\colon T_{X,\underline{1}} \twoheadrightarrow M$ *given by* $\langle \pi, x, a\rangle \mapsto X_{\pi,x,a}$.

Let L_β *be the* β*-iterate of* $\varnothing$ *under* D *and define* $\mathbb{L} = \bigcup_{\beta \in \mathbb{O}} \mathbb{L}_\beta$ *where* $\mathbb{L}_\beta = \mathrel{\varepsilon}^{-1}(L_\beta)$.

The $\mathcal{M}$*-sets in* $\mathbb{L}$ *are said to be* constructible*.*

We call $\mathcal{M}$ constructible *if every* $\mathcal{M}$*-set is constructible.*

Remark 2.11.6. $\mathsf{D}(X)$ forms a boolean subalgebra of $\mathsf{P}(\mathrel{\varepsilon}^{-1}(X))$ in $\mathcal{M}$.

Lemma 2.11.7. *For all* $\mathcal{M}$*-sets* X *and* $\mathcal{M}|_{\varepsilon^{-1}(X)}$*-classes* C *we have* $[C] \mathrel{\varepsilon} \mathsf{D}(X)$.

More generally, for every S*-formula* π *and every tuple* $(x_m)_{m<n}$ *of distinct variable symbols in* S_V *that contains all free variables of* π*, and* $(a_m)_{m<n}$ *with* $a_m \mathrel{\varepsilon} X$ *it is*

$$\mathcal{M}|_{\varepsilon^{-1}(X)} \vDash \pi[(x_m)_{m<n}/(a_m)_{m<n}] \;\Leftrightarrow\; X \vDash \underline{\pi}\big[\langle x_\alpha\rangle_{\alpha<\underline{n}}/\langle a_\alpha\rangle_{\alpha<\underline{n}}\big]$$

where $\langle x_\alpha\rangle_{\alpha<\underline{n}}$ *and* $\langle a_\alpha\rangle_{\alpha<\underline{n}}$ *are the families of* $\mathcal{M}$*-sets with* $x_{\underline{m}} = \underline{x_m}$ *and* $a_{\underline{m}} = a_m$.

Proof. By structural induction on π. □

Lemma 2.11.8. $X \sqsubseteq \mathsf{D}(X)$ *and* $\mathsf{D}(X)$ *is transitive for all transitive* $\mathcal{M}$*-sets* X.

Consequently, $\langle L_\alpha\rangle_{\alpha \in \mathbb{O}}$ *is a cumulative* $\mathcal{M}$*-hierarchy (see Problem Set 5).*

Proof. For every $V \sqsubseteq X$ by transitivity $V = [U \sqsubseteq X : U \sqsubseteq V] \sqsubseteq \mathsf{D}(X)$. In particular, for all $V \sqsubseteq W \sqsubseteq \mathsf{D}(X)$ we have $V \sqsubseteq \mathsf{D}(X)$ because of $V \sqsubseteq W \sqsubseteq X$. □

Lemma 2.11.9. *For every finite $\mathcal{M}$-subset Y of an $\mathcal{M}$-set X we have $Y \sqsubseteq \mathsf{D}(X)$.*

Proof. Let $f\colon \alpha \twoheadrightarrow Y$ be bijective with $\alpha \in \mathbb{N}^{\mathcal{M}}$. Clearly, $\varnothing \sqsubseteq \mathsf{D}(X)$, so assume $Y \neq \varnothing$ and inductively $Y' = Y \setminus [f(\alpha - \underline{1})] \sqsubseteq \mathsf{D}(X)$. But then $Y = Y' \sqcup [f(\alpha - \underline{1})] \sqsubseteq \mathsf{D}(X)$. □

Lemma 2.11.10. *For convenience define $\mathbb{L}_\infty = \mathbb{L}$. The following hold:*

(a) $\mathbb{L}_\beta$ is transitive for every $\beta \in \mathbb{O} \cup \{\infty\}$.

(b) $\mathbb{L}_\beta = \mathbb{W}_\beta$ for every $\beta \leq \omega$ and $\mathbb{L}_\beta \subseteq \mathbb{W}_\beta$ for every $\beta \in \mathbb{O} \cup \{\infty\}$.

(c) $\mathbb{L}_\beta \cap \mathbb{O} = \mathbb{O}_{<\beta}$ for all $\beta \in \mathbb{O} \cup \{\infty\}$, so $\beta \in \mathbb{L}_{\beta+\underline{1}} \setminus \mathbb{L}_\beta$ for all $\beta \in \mathbb{O}$.

(d) $X \in \mathbb{L}$ for each $\mathcal{M}|_{\mathbb{L}}$-class that forms an $\mathcal{M}$-set X.

Proof. (a) This is a direct consequence of Lemma 2.11.8 and Remark 2.3.4.

(b) Assuming by induction $L_{\beta-\underline{1}} = W_{\beta-\underline{1}}$ we get $L_\beta = \mathsf{D}(L_{\beta-\underline{1}}) = \mathsf{P}(W_{\beta-\underline{1}}) = W_\beta$ for all $\underline{0} < \beta < \omega$ by Corollary 2.6.7 and Lemma 2.11.9, which implies the first claim. The second claim follows by induction from Theorem 2.6.6.

(c) In view of (a) and (b) we have $\mathbb{L}_\beta \cap \mathbb{O} = \mathbb{O}^{\mathcal{M}|_{\mathbb{L}_\beta}}$ for all $\beta \in \mathbb{O} \cup \{\infty\}$ just as in the proof of Lemma 2.7.3 (7,9) where $\mathbb{O}^{\mathcal{M}|_{\mathbb{L}_\beta}}$ is the $\mathcal{M}|_{\mathbb{L}_\beta}$-class consisting of all $\alpha \in \mathbb{L}_\beta$ such that α and all $\gamma \sqsubseteq \alpha$ are transitive. We need to verify $\beta = L_\beta \sqcap \mathbb{O}$. Assume by induction that $\alpha = L_\alpha \sqcap \mathbb{O}$ for every $\alpha < \beta$. If β is a limit ordinal, then $\beta = \bigsqcup_{\alpha<\beta} \alpha = \bigsqcup_{\alpha<\beta} L_\alpha \sqcap \mathbb{O} = L_\beta \sqcap \mathbb{O}$. If β is a successor ordinal, then

$$\beta - \underline{1} \;=\; L_{\beta-\underline{1}} \sqcap \mathbb{O} \;=\; \left[\mathbb{O}^{\mathcal{M}|_{\mathbb{L}_{\beta-\underline{1}}}}\right] \sqsubseteq \mathsf{D}(L_{\beta-\underline{1}}) \sqcap \mathbb{O} \;=\; L_\beta \sqcap \mathbb{O}\,,$$

where the penultimate step used Lemma 2.11.7. By transitivity we get $\beta \sqsubseteq L_\beta \sqcap \mathbb{O}$. If the inclusion were proper, there would be an $\mathcal{M}$-ordinal $\gamma \geq \beta$ with $\gamma \sqsubseteq L_\beta$, so again by transitivity $\beta \sqsubseteq L_\beta = \mathsf{D}(L_{\beta-\underline{1}})$ and thus $\beta - \underline{1} \sqsubseteq \beta \sqsubseteq L_{\beta-\underline{1}}$, a contradiction.

(d) Let π be an S^{Set}-formula, $\mathrm{fvar}(\pi) \subseteq \{v, w_1, \ldots, w_n\}$, $W_1, \ldots, W_n \in \mathbb{L}$ with

$$C \;=\; \left\{V \in \mathbb{L} \;\; : \mathcal{M}|_{\mathbb{L}} \;\; \vDash \pi[(v, w_1, \ldots, w_n)/(V, W_1, \ldots, W_n)]\right\},$$

forming an $\mathcal{M}$-set. Since $C \to \mathbb{O}$, $V \mapsto \min\{\alpha \in \mathbb{O} : V \in \mathbb{L}_{\alpha+\underline{1}}\}$, forms an $\mathcal{M}$-set, there exists $\alpha \in \mathbb{O}$ with $W_1, \ldots, W_n \in \mathbb{L}_\alpha$ and $C \subseteq \mathbb{L}_\alpha$. As a consequence, we have

$$C \;=\; \left\{V \in \mathbb{L}_\alpha : \mathcal{M}|_{\mathbb{L}} \;\; \vDash \pi[(v, w_1, \ldots, w_n)/(V, W_1, \ldots, W_n)]\right\}.$$

By the Reflection Principle (see the solution of Exercise 1 on Problem Set 5), replacing α if necessary by a larger $\mathcal{M}$-ordinal, we may assume that

$$C \;=\; \left\{V \in \mathbb{L}_\alpha : \mathcal{M}|_{\mathbb{L}_\alpha} \vDash \pi[(v, w_1, \ldots, w_n)/(V, W_1, \ldots, W_n)]\right\}.$$

But then $[C] \sqsubseteq \mathsf{D}(L_\alpha) = L_{\alpha+\underline{1}} \sqsubseteq \mathbb{L}$ in view of Lemma 2.11.7. □

Theorem 2.11.11. *$\mathcal{M}|_{\mathbb{L}}$ is a constructible* $(\mathrm{ZF}^\circ \cup \mathsf{INF} \cup \mathsf{REG})$*-universe.*

If $\mathcal{M}$ satisfies Power set*, then so does $\mathcal{M}|_{\mathbb{L}}$.*

Finally, if $\mathcal{M}|_C$ is a $(\mathrm{ZF}^\circ \cup \mathsf{INF} \cup \mathsf{REG})$*-universe for some transitive $\mathcal{M}$-class C, then necessarily $\mathbb{L}^{\mathcal{M}|_C} = \mathbb{L}_\beta$ for some $\beta \in \mathbb{O} \cup \{\infty\}$. In particular, $\mathbb{L}^{\mathcal{M}|_{\mathbb{L}}} = \mathbb{L}$.*

Proof. Using Lemmas 2.7.3 and 2.11.10 we can conclude $\mathcal{M}|_{\mathbb{L}} \vDash \mathrm{ZF}^\circ \cup \mathsf{INF} \cup \mathsf{REG}$ and $\mathcal{M} \vDash \mathsf{POW} \Rightarrow \mathcal{M}|_{\mathbb{L}} \vDash \mathsf{POW}$.

Now for every transitive $\mathcal{M}$-class C such that $\mathcal{M}|_C$ is a $(\mathrm{ZF}^\circ \cup \mathsf{INF} \cup \mathsf{REG})$-universe we have with the argument given in the proof of Lemma 2.7.3 (7,9) $\mathbb{O}^{\mathcal{M}|_C} = \mathbb{O} \cap C$ and $\mathbb{N}^{\mathcal{M}|_C} = \mathbb{N}^{\mathcal{M}}$. It then follows that $\mathbb{W}_\omega^{\mathcal{M}|_C} = \mathbb{W}_\omega$ and the $\mathcal{M}|_C$-set of $S^{\mathcal{M}|_C}$-formulas agrees with the $\mathcal{M}$-set of $S^{\mathcal{M}}$-formulas such that analyzing the definition of D we may conclude $\mathsf{D}^{\mathcal{M}|_C} = \mathsf{D}|_C$. Hence, $\mathbb{L}_\alpha^{\mathcal{M}|_C} = \mathbb{L}_\alpha$ for all $\alpha \in \mathbb{O} \cap C$ by recursion in $\mathcal{M}$.

By Lemma 2.3.14 $\mathbb{O} \cap C = \mathbb{O}_{<\beta}$ for some $\beta \in \mathbb{O} \cup \{\infty\}$. Moreover, since $\mathbb{O}^{\mathcal{M}|_C}$ is inductive, β is a limit ordinal or ∞. So $\mathbb{L}^{\mathcal{M}|_C} = \bigcup_{\alpha \in \mathbb{O}^{\mathcal{M}|_C}} \mathbb{L}_\alpha^{\mathcal{M}|_C} = \bigcup_{\alpha<\beta} \mathbb{L}_\alpha = \mathbb{L}_\beta$. $\square$

Remark 2.11.12. Let $\mathcal{M}$ be a ZFC-universe. It is not hard to see that $\mathbb{K} \subseteq \mathbb{K}^{\mathcal{M}|_{\mathbb{L}}}$. However, equality does not hold in general.

Exercise 2.11.13. The $\mathcal{M}$-class $\mathbb{L}$ in any ZFC-universe $\mathcal{M}$ is the smallest transitive $\mathcal{M}$-class C containing all $\mathcal{M}$-ordinals such that $\mathcal{M}|_C$ is again a ZFC-universe.

Relative consistency of the axiom of choice

This subsection will establish that $\mathcal{M}|_{\mathbb{L}}$ satisfies Choice. The simple reason for this is that it is not hard to explicitly define a well-order on the $\mathcal{M}$-class $\mathbb{L}$.

Lemma 2.11.14. *For every well-orderable $\mathcal{M}$-set X also $\mathsf{D}(X)$ is well-orderable.*

More precisely, denoting by WO *the $\mathcal{M}$-class of all $\mathcal{M}$-relations that are well-orders, there exists an $\mathcal{M}$-class function* $\mathrm{WO} \to \mathrm{WO}$, $\prec \mapsto \prec^{\mathsf{D}}$*, such that for every well-order $\prec$ of X the $\mathcal{M}$-set $\prec^{\mathsf{D}}$ is a well-order of $\mathsf{D}(X)$.*

Proof. It is not very hard to see that the $\mathcal{M}$-set of all $S^{\mathcal{M}}$-formulas is well-orderable. So we can fix an $\mathcal{M}$-relation $<^{\mathcal{F}}$ that well-orders it.

For every well-order $\prec$ on an $\mathcal{M}$-set X denote by $\prec^T$ the well-order on $T_{X,\underline{1}}$ that compares elements $\langle \pi, x, a\rangle$ in $T_{X,\underline{1}}$ lexicographically looking firstly at π with $<^{\mathcal{F}}$, secondly at $\mathrm{dom}(x) = \mathrm{dom}(a) + \underline{1}$ with $<$, thirdly at x with the lexicographic order on the product $\mathop{\Large *}_{\alpha<\mathrm{dom}(x)} \omega$ induced by $<$, and finally at a with the lexicographic order on the product $\mathop{\Large *}_{\alpha<\mathrm{dom}(a)} X$ induced by $\prec$.

Finally, we use the canonical surjective $\mathcal{M}$-function $p_X \colon T_{X,\underline{1}} \twoheadrightarrow \mathsf{D}(X)$ to obtain a well-order $\prec^{\mathsf{D}}$ given by $U \prec^{\mathsf{D}} V \Leftrightarrow \min_{\prec^T} p_X^{-1}[[U]] \prec^T \min_{\prec^T} p_X^{-1}[[V]]$. $\square$

Theorem 2.11.15. *There is an $\mathcal{M}$-class relation $\prec$ that well-orders $\mathbb{L}$.*

Proof. Let $\gamma \in \mathbb{O}$ and recursively assume that well-orders $\prec_\alpha$ are given on all L_α with $\alpha < \gamma$ such that $\prec_\alpha$ is the restriction of $\prec_\beta$ for all $\alpha \leq \beta < \gamma$.

If γ is a limit ordinal, define $\prec_\gamma = \bigsqcup_{\alpha<\gamma} \prec_\alpha$.

If γ is a successor ordinal, define $\prec_\gamma$ for $X, Y \in L_\gamma$ via

$$X \prec_\gamma Y \;\Leftrightarrow\; \begin{cases} X \in L_{\gamma-\underline{1}} \text{ and } Y \in L_{\gamma-\underline{1}} \text{ and } X \prec_{\gamma-\underline{1}} Y \text{ or,} \\ X \notin L_{\gamma-\underline{1}} \text{ and } Y \notin L_{\gamma-\underline{1}} \text{ and } X \prec^{\mathsf{D}}_{\gamma-\underline{1}} Y \text{ or,} \\ X \in L_{\gamma-\underline{1}} \text{ and } Y \notin L_{\gamma-\underline{1}}. \end{cases}$$

Clearly, in either case $\prec_\gamma$ will be a well-order on L_γ with the property that $\prec_\alpha$ is the restriction of $\prec_\beta$ for all $\alpha \le \beta \le \gamma$. By recursion we now obtain an $\mathcal{M}$-class function $\alpha \mapsto \prec_\alpha$ and we can take for $\prec$ the $\mathcal{M}$-class relation $\bigcup_{\alpha\in\mathbb{O}} \hat{\prec}_\alpha$.

With Lemma 2.4.3 it follows that $\prec$ well-orders $\mathbb{L}$. □

Corollary 2.11.16. *If $\mathcal{M} \vDash \mathrm{ZF}^-$, then $\mathcal{M}|_{\mathbb{L}} \vDash \mathrm{ZFC}$.*

Hence, the consistency of ZF^- *implies the consistency of* ZFC.

Proof. Assume that $\mathcal{M} \vDash \mathrm{ZF}^-$. Theorem 2.11.11 already states that $\mathcal{M}|_{\mathbb{L}} \vDash \mathrm{ZF}$ holds. By Theorem 2.9.2 it thus suffices to show that every $\mathcal{M}|_{\mathbb{L}}$-set X is well-orderable. But according to Theorems 2.11.11 and 2.11.15 there exists an $\mathcal{M}|_{\mathbb{L}}$-class relation $\prec$ on $\mathbb{L}$ that well-orders $\mathbb{L}$. Since $\mathbb{L}$ is transitive, $\prec$ induces a well-order on $X \sqsubseteq \mathbb{L}$. □

Relative consistency of GCH

Before proving that $\mathcal{M}|_{\mathbb{L}}$ satisfies GCH, we need some preparations. Most importantly, the Löwenheim–Skolem Theorem and an internal version of the Reflection Principle.

Convention 2.11.17. In this subsection, we assume that $\mathcal{M}$ is a ZFC-universe.

With Choice at hand we can calculate the cardinalities of the $\mathcal{M}$-sets L_β:

Lemma 2.11.18. *For every infinite $\mathcal{M}$-set X we have $|\mathsf{D}(X)| = |X|$.*

Consequently, $|L_\beta| = |\beta|$ for all infinite $\mathcal{M}$-ordinals β.

Proof. On the one hand, we have $|X| \le |\mathsf{D}(X)|$ because of $X \sqsubseteq \mathsf{D}(X)$, and, on the other hand, $|\mathsf{D}(X)| \le |T_{X,\underline{1}}| = \max\big[|X|, \omega\big] = |X|$.

Now take $\omega \le \beta \in \mathbb{O}$. Clearly, $|\beta| \le |L_\beta|$ because of $\beta \sqsubseteq L_\beta$. Since L_α is finite for $\alpha < \omega$ and using induction we may assume $|L_\alpha| \le \max\big[|\alpha|, \omega\big] \le |\beta|$ for all $\alpha < \beta$.

If β is a limit ordinal, then $|L_\beta| = |\bigsqcup_{\alpha<\beta} L_\alpha| \le \bigoplus_{\alpha<\beta} |L_\alpha| \le |\beta|$.

If β is a successor ordinal, then $|L_\beta| = |\mathsf{D}(L_{\beta-\underline{1}})| = |L_{\beta-\underline{1}}| = |\beta - \underline{1}| = |\beta|$. □

Definition 2.11.19. *An $\mathcal{M}$-subset Y of an $\mathcal{M}$-set X is said to be* elementary *if the equivalence $Y \vDash \pi[x/a] \Leftrightarrow X \vDash \pi[x/a]$ holds for every $\langle \pi, x, a\rangle \in T_{Y,\underline{0}}$.*

The following criterion is useful for determining whether an $\mathcal{M}$-subset is elementary and saves us from having to do structural induction on $S^{\mathcal{M}}$-formulas each time:

Lemma 2.11.20. *An $\mathcal{M}$-subset Y of an $\mathcal{M}$-set X is elementary if and only if the implication $X \smallsetminus X_{\pi,x,a} \ne \varnothing \Rightarrow Y \smallsetminus X_{\pi,x,a} \ne \varnothing$ holds for every $\langle \pi, x, a\rangle \in T_{Y,\underline{1}}$.*

Proof. Let $\langle \pi, x, a\rangle \sqsubseteq T_{Y,\underline{0}}$. We have to show $Y \vDash \pi[x/a] \Leftrightarrow X \vDash \pi[x/a]$.

This is clear for atomic π.

By structural induction it holds in case $\pi = (\varphi \underline{\rightarrow} \psi)$.

Finally, consider $\pi = \underline{\bigwedge}_y \varphi$. We may then assume $y \neq x_\alpha$ for all $\alpha < \beta = |x|$ and let $\langle \varphi, z, a\rangle \sqsubseteq T_{Y,\underline{1}}$ with $z|_\beta = x$ and $z_\beta = y$. Then $\Leftarrow$ holds because $Y \sqsubseteq X$ and by structural induction. To see $\Rightarrow$ let $X \nvDash \pi[x/a]$, i.e. $X \smallsetminus X_{\varphi,z,a} \neq \varnothing$, so by assumption then $Y \smallsetminus X_{\varphi,z,a} \neq \varnothing$. Structural induction yields $Y \smallsetminus Y_{\varphi,z,a} \neq \varnothing$, i.e. $Y \nvDash \pi[x/a]$. $\square$

Lemma 2.11.21 (Downward Löwenheim–Skolem Theorem)**.** *For all $\mathcal{M}$-sets $Z \sqsubseteq X$ there exists an elementary $\mathcal{M}$-subset Y of X with $Z \sqsubseteq Y$ and $|Y| \leq \max\big[|Z|, \omega\big]$.*

Proof. Choose a well-order $\prec$ on X. According to Lemma 2.11.20 $Y = \bigsqcup_{\alpha<\omega} Y_\alpha$ with

$$Y_{\underline{0}} \;=\; Z \;\text{ and }\; Y_{\alpha+\underline{1}} \;=\; Y_\alpha \sqcup \big[\min\nolimits_\prec W : \langle \pi, x, a\rangle \sqsubseteq T_{Y_\alpha,\underline{1}} \text{ with } W = X \smallsetminus X_{\pi,x,a} \neq \varnothing\big]$$

is an elementary $\mathcal{M}$-subset of X. Since $|Y_{\underline{0}}| = |Z| \leq \max\big[|Z|, \omega\big]$ induction yields

$$|Y_{\alpha+\underline{1}}| \;\leq\; |Y_\alpha| \oplus |T_{Y_\alpha,\underline{1}}| \;\leq\; \max\big[|Z|, \omega\big]$$

because of $|T_{Y_\alpha,\underline{1}}| \leq \max\big[|Y_\alpha|, \omega\big]$. Consequently, $|Y| \leq \max\big[|Z|, \omega\big]$, too. $\square$

Lemma 2.11.22 (Internal Reflection Principle)**.** *For each regular $\mathcal{M}$-cardinal $\kappa > \omega$ and cumulative $\mathcal{M}$-hierarchy $\langle V_\alpha\rangle_{\alpha \in \mathbb{O}}$ with the property $|V_\alpha| < \kappa$ for all $\alpha < \kappa$ the $\mathcal{M}$-set $\big[\alpha < \kappa : V_\alpha$ is an elementary $\mathcal{M}$-subset of $V_\kappa\big]$ is closed and unbounded in κ.*

Proof. To prove closedness, let W be a non-empty $\mathcal{M}$-subset of κ such that V_γ is an elementary $\mathcal{M}$-subset of V_κ for all $\gamma \sqsubseteq W$. Let $\alpha = \bigsqcup W$. We have to show that $Y = V_\alpha$ is an elementary $\mathcal{M}$-subset of $X = V_\kappa$. We use Lemma 2.11.20.

So take $\langle \pi, x, a\rangle \sqsubseteq T_{Y,\underline{1}}$ with $X \smallsetminus X_{\pi,x,a} \neq \varnothing$. Because the $\mathcal{M}$-function a has finite image and the $\mathcal{M}$-sets V_γ with $\gamma \sqsubseteq W$ are pairwise comparable by $\sqsubseteq$, there is $\gamma \sqsubseteq W$ with $\langle \pi, x, a\rangle \sqsubseteq T_{V_\gamma,\underline{1}}$. Then $V_\gamma \smallsetminus X_{\pi,x,a} \neq \varnothing$, so $Y \smallsetminus X_{\pi,x,a} \neq \varnothing$ because of $\gamma \leq \alpha$.

To prove unboundedness, let $\varepsilon < \kappa$. Define $\alpha = \bigsqcup_{\sigma<\omega} \alpha_\sigma$ recursively by

$$\alpha_{\underline{0}} \;=\; \varepsilon \;\text{ and }\; \alpha_{\sigma+\underline{1}} \;=\; \bigsqcup \Big[\gamma_{\pi,x,a,\alpha_\sigma} : \langle \pi, x, a\rangle \sqsubseteq T_{V_{\alpha_\sigma},\underline{1}} \text{ with } X \smallsetminus X_{\pi,x,a} \neq \varnothing\Big]$$

where $\gamma_{\pi,x,a,\alpha_\sigma} = \min\big[\delta < \kappa : \alpha_\sigma \leq \delta \text{ with } V_\delta \smallsetminus X_{\pi,x,a} \neq \varnothing\big]$ and as before $X = V_\kappa$. Again we use the criterion Lemma 2.11.20 to verify that $Y = V_\alpha$ is an elementary $\mathcal{M}$-subset of X.

So take $\langle \pi, x, a\rangle \sqsubseteq T_{Y,\underline{1}}$ with $X \smallsetminus X_{\pi,x,a} \neq \varnothing$. Similarly as above when we proved closedness, we have $\langle \pi, x, a\rangle \sqsubseteq T_{V_{\alpha_\sigma},\underline{1}}$ for some $\sigma < \omega$. Hence, $V_{\gamma_{\pi,x,a,\alpha_\sigma}} \smallsetminus X_{\pi,x,a} \neq \varnothing$ and thus $Y \smallsetminus X_{\pi,x,a} \neq \varnothing$ because of $\gamma_{\pi,x,a,\alpha_\sigma} \leq \alpha$.

Clearly, $\varepsilon \leq \alpha$. So all that remains to be checked is $\alpha < \kappa$. Assuming $\alpha_\sigma < \kappa$ by induction on $\sigma < \omega$ we derive $\alpha_{\sigma+\underline{1}} < \kappa$ from $\big|T_{V_{\alpha_\sigma},\underline{1}}\big| \leq \max\big[|V_{\alpha_\sigma}|, \omega\big] < \kappa = \mathrm{cof}(\kappa)$. But then we immediately get $\alpha < \kappa$ because of $\omega < \kappa = \mathrm{cof}(\kappa)$. $\square$

Theorem 2.11.23. $\mathcal{M}|_{\mathbb{L}_\kappa}$ *is a constructible* $(\mathrm{ZF}^\circ \cup \mathsf{INF} \cup \mathsf{CHO} \cup \mathsf{REG})$*-universe for every regular* $\mathcal{M}$*-cardinal* $\kappa > \omega$*, i.e. it satisfies all* ZFC *axioms except* POWER SET.

Proof. Applying the lemma from the solution of Exercise 3 on Problem Set 5 we immediately see that $\mathcal{M}|_{\mathbb{L}_\kappa} \vDash \mathsf{EXT} \cup \mathsf{EMP} \cup \mathsf{PAI} \cup \mathsf{UNI} \cup \mathsf{INF} \cup \mathsf{CHO} \cup \mathsf{REG}$. Since $\mathbb{L}_\kappa$ is transitive, the constructibility of $\mathcal{M}|_{\mathbb{L}_\kappa}$ holds then by Theorem 2.11.11.

It remains to check $\mathcal{M}|_{\mathbb{L}_\kappa} \vDash \mathsf{REP}$. To do this, let $f\colon \mathrel{\varepsilon}^{-1}(X) \twoheadrightarrow \mathbb{L}_\kappa$ with $X \in \mathbb{L}_\kappa$ be definable in $\mathcal{M}|_{\mathbb{L}_\kappa}$. Then there exists an S-formula π with $\mathrm{fvar}(\pi) \subseteq \{v, w_1, \dots, w_n\}$ and with $W_1, \dots, W_n \in \mathbb{L}_\kappa$ such that as $\mathcal{M}$-sets

$$f[X] \;=\; \big[V \mathrel{\varepsilon} L_\alpha : \mathcal{M}|_{\mathbb{L}_\kappa} \vDash \pi[(v, w_1, \dots, w_n)/(V, W_1, \dots, W_n)]\big]\,.$$

where $\alpha = \bigsqcup_{V \mathrel{\varepsilon} f[X]} \min\big[\delta < \kappa : V \mathrel{\varepsilon} L_\delta\big]$. Picking some $\omega \le \gamma < \kappa$ such that $X \mathrel{\varepsilon} L_\gamma$, the transitivity of L_γ gives $X \sqsubseteq L_\gamma$, so using Lemma 2.11.18 $|X| \le |L_\gamma| = |\gamma| < \kappa$. Since κ is regular, we can conclude that $\alpha < \kappa$. Replacing α by a larger $\mathcal{M}$-ordinal less than κ, we may assume $W_1, \dots, W_n \mathrel{\varepsilon} L_\alpha$ and additionally in view of Lemma 2.11.22 that L_α is an elementary $\mathcal{M}$-subset of L_κ. We then get $f[X] \mathrel{\varepsilon} \mathsf{D}(L_\alpha) = L_{\alpha+\underline{1}} \sqsubseteq L_\kappa$ with the help of Lemma 2.11.7. This finishes the proof of $\mathcal{M}|_{\mathbb{L}_\kappa} \vDash \mathsf{REP}$. □

Since we have already proven that the consistency of ZF^- implies the consistency of ZFC (see Corollary 2.11.16), the next result establishes the consistency of ZFC together with GCH granted the consistency of ZF^-.

Theorem 2.11.24. *If* $\mathcal{M}$ *is constructible, then it satisfies GCH.*

Proof. Let λ be an infinite $\mathcal{M}$-cardinal. We have to prove $\underline{2}^\lambda = \lambda^+$. Given that we already now $\lambda^+ \le \underline{2}^\lambda$ from Lemma 2.10.9, it is sufficient to check $\mathsf{P}(\lambda) \sqsubseteq L_{\lambda^+}$ as this with Lemma 2.11.18 will yield the missing inequality $\underline{2}^\lambda = |\mathsf{P}(\lambda)| \le |L_{\lambda^+}| = \lambda^+$.

Take $W \mathrel{\varepsilon} \mathsf{P}(\lambda)$. Since $\mathcal{M}$ is constructible, there is some $\mathcal{M}$-cardinal $\kappa > \lambda$ such that $W \mathrel{\varepsilon} L_\kappa$. In view of Lemma 2.10.26 we may assume that κ is regular, which will ensure that $\mathcal{M}|_{\mathbb{L}_\kappa}$ is a constructible $(\mathrm{ZF}^\circ \cup \mathsf{INF} \cup \mathsf{REG})$-universe by Theorem 2.11.23.

The Löwenheim–Skolem Theorem (Lemma 2.11.21) lets us choose an elementary $\mathcal{M}$-subset Y of L_κ such that $Z = \lambda \sqcup [W] \sqsubseteq Y$ and $|Y| \le \max\big[|Z|, \omega\big] = \lambda$.

According to Mostowski's Isomorphism Theorem (Theorem 2.5.9) the transitive collapse is an isomorphism $t\colon (\mathrel{\varepsilon}^{-1}(Y), \mathrel{\varepsilon}) \to (T, \mathrel{\varepsilon})$ where T is a transitive $\mathcal{M}$-class.

Since Y is an elementary $\mathcal{M}$-subset of L_κ and t is an isomorphism, $\mathcal{M}|_T$ is a constructible $(\mathrm{ZF}^\circ \cup \mathsf{INF} \cup \mathsf{REG})$-universe, so $T = \mathbb{L}_\beta$ for some infinite $\mathcal{M}$-ordinal β by Theorem 2.11.11.

The transitivity of λ and the fact $W \sqsubseteq \lambda$ imply that Z is transitive. It is easy to see that the transitive collapse t is the identity when restricted to transitve $\mathcal{M}$-subsets of Y. Hence, $W = t(W) \in T = \mathbb{L}_\beta$. To obtain as required $W \mathrel{\varepsilon} L_{\lambda^+}$, it only remains to observe $L_\beta \sqsubseteq L_{\lambda^+}$ because $|\beta| = |L_\beta| = |Y| \le \lambda$ yields $\beta < \lambda^+$. □

2.12 Boolean-valued sets

In this section we will often cite [Jec03] and [Bel11] for technical proofs and definitions. The treatment in these sources is sometimes a little sketchy. The most recommendable, well-written, and self-contained source on the subject is in my opinion still [Ros69].

It is natural to identify $\mathcal{M}$-sets X with their characteristic $\mathcal{M}$-functions $X \multimap \underline{2}$. Of course, this approach is not completely consistent in the sense that the $\mathcal{M}$-elements of the domain of these $\mathcal{M}$-functions are not characteristic $\mathcal{M}$-functions themselves. This can be remedied by a hierarchical construction as given in the next definition, taking there for $\mathbb{A}$ the boolean ring $\mathbb{F}_2$. But why we are really interested in this construction is that clever choices of $\mathbb{A}$ will give rise to boolean-valued models of ZFC where certain S-formulas (e.g. describing GCH) are not satisfied.

Convention 2.12.1. In this section, we will assume that $\mathcal{M}$ is a ZFC-universe and fix a complete boolean ring $\langle A, +, \cdot \rangle$ in $\mathcal{M}$ and consider the $\mathcal{M}$-class $\mathbb{A} = \sqsubseteq^{-1}(A)$ as a boolean ring with addition given by $\hat{+}$ and multiplication given by $\hat{\cdot}$.

Remark 2.12.2. $\mathbb{A}$ is complete w.r.t. $\mathsf{P}(\mathbb{A})$ but in general not complete w.r.t. $\mathcal{P}(\mathbb{A})$, i.e. each $\mathcal{M}$-subset of $\mathbb{A}$ has a supremum but not necessarily each subset of $\mathbb{A}$.

It is easy to see that the following is well-defined using well-founded recursion:

Definition 2.12.3. *Let $\mathbb{B} = \bigcup_{\alpha \in \mathbb{O}} \mathbb{B}_\alpha$ with $[\mathbb{B}_\alpha] = B_\alpha$ and let $\approx, \in : \mathbb{B}^2 \to \mathbb{A}$ be the functions definable in $\mathcal{M}$ determined by the following properties for all $X, Y \in \mathbb{B}$:*

(a) $\langle B_\alpha \rangle_{\alpha \in \mathbb{O}}$ is the cumulative $\mathcal{M}$-hierarchy such that for all $\alpha \in \mathbb{O}$

$$B_{\alpha+\underline{1}} = \left[Z : Z \text{ is an } \mathcal{M}\text{-function } B \multimap A \text{ with } B \sqsubseteq B_\alpha \right].$$

(b) $X \in Y = \bigvee_{U \sqsubseteq \mathrm{dom}(Y)} (X \approx U \wedge Y(U))$.

(c) $X \approx Y = X \Subset Y \wedge Y \Subset X$ where $X \Subset Y = \bigwedge_{U \sqsubseteq \mathrm{dom}(X)} (X(U) \to U \in Y)$.

We will denote by $\mathcal{M}_\mathbb{A}$ the unassigned boolean-valued S-structure with $\underline{\mathcal{M}_\mathbb{A}} = \mathbb{B}$ and $R^{\mathcal{M}_\mathbb{A}} = \mathbb{A}$ and $\equiv^{\mathcal{M}_\mathbb{A}} = \approx$ and $\epsilon^{\mathcal{M}_\mathbb{A}} = \in$.

We refer to the elements of $\mathbb{B}$ as $\mathbb{A}$-valued $\mathcal{M}$-sets.

As could be hoped for, this construction gives rise to a boolean-valued model of ZFC. The verification is lengthy but straightforward. To give an idea of how it can be undertaken, we sketch a proof of the following result but we will later just refer to the literature for the computations necessary to check the remaining axioms of ZFC:

Lemma 2.12.4. *$\mathcal{M}_\mathbb{A}$ is a boolean-valued model of an S-theory with equality $\equiv$.*

Proof. We have to check that $\mathcal{M}_\mathbb{A}$ satisfies (R), (S), (T), (C) in Definition 1.4.1.

This is obvious for (S).

To check (R) take $X \in \mathbb{B}$. By induction assume $U \approx U = 1$ for all $U \sqsubseteq \operatorname{dom}(X)$. Then $U \boldsymbol{\in} X \geq U \approx U \wedge X(U) = X(U)$, i.e. $X(U) \to U \boldsymbol{\in} X = 1$, so $X \approx X = 1$.

To verify (T) and (C) it is enough to prove that $X \approx Y \leq Y \approx Z \to X \approx Z$ and $X \approx Y \leq X \boldsymbol{\in} Z \to Y \boldsymbol{\in} Z$ and $X \approx Y \leq Z \boldsymbol{\in} X \to Z \boldsymbol{\in} Y$ for all $X, Y, Z \in \mathbb{B}$. We do this simultaneously by induction following the proof of [Jec03, Lemma 14.16].

Firstly, by induction $Y \approx Z \leq U \boldsymbol{\in} Y \to U \boldsymbol{\in} Z$, i.e. $U \boldsymbol{\in} Y \leq Y \approx Z \to U \boldsymbol{\in} Z$ for all $U \sqsubseteq \operatorname{dom}(X)$, so $X(U) \to U \boldsymbol{\in} Y \leq Y \approx Z \to (X(U) \to U \boldsymbol{\in} Z)$. Taking the infimum over all such U yields $X \Subset Y \leq Y \approx Z \to X \Subset Z$. With a similar argument $Y \Subset X \leq Y \approx Z \to Z \Subset X$. Hence, $X \approx Y \leq Y \approx Z \to X \approx Z$.

Secondly, $X \approx Y \leq X \approx U \to Y \approx U \leq (X \approx U \wedge Z(U)) \to (Y \approx U \wedge Z(U))$ by induction for all $U \sqsubseteq \operatorname{dom}(Z)$. Hence, $X \approx Y \leq X \boldsymbol{\in} Z \to Y \boldsymbol{\in} Z$.

Thirdly, $X \approx Y \leq X \Subset Y \leq X(U) \to U \boldsymbol{\in} Y \leq (Z \approx U \wedge X(U)) \to Z \boldsymbol{\in} Y$ where the last inequality holds since $Z \approx U \leq U \boldsymbol{\in} Y \to Z \boldsymbol{\in} Y$ by induction for all $U \sqsubseteq \operatorname{dom}(X)$. Thus $X \approx Y \leq Z \boldsymbol{\in} X \to Z \boldsymbol{\in} Y$, finishing the proof. □

Remark 2.12.5. One might think that it is necessary in Definition 2.12.3 to replace $\mathbb{B}$ with $\mathbb{B}' = \{X \in \mathbb{B} : X(U) = U \boldsymbol{\in} X\}$ to make sure that the axiom of extensionality is satisfied in $\mathcal{M}_{\mathbb{A}}$. Even though it is indeed possible to make this replacement (as done in [Ros69]) it is not necessary (as follows from the presentation in [Jec03; Bel11]).

Actually, for every $\mathbb{A}$-valued $\mathcal{M}$-set X the $\mathbb{A}$-valued $\mathcal{M}$-set X' with domain $\operatorname{dom}(X)$ and $X'(U) = U \boldsymbol{\in} X$ for all $U \sqsubseteq \operatorname{dom}(X)$ belongs to $\mathbb{B}'$ and satisfies $X' \approx X = 1$.

We can map each $\mathcal{M}$-set in the obvious way to its corresponding "characteristic function" in the boolean-valued model $\mathcal{M}_{\mathbb{A}}$ as follows:

Definition 2.12.6. *For $\mathcal{M}$-sets X define $\mathbb{1}_X$ recursively as the $\mathbb{A}$-valued $\mathcal{M}$-set with domain $[\, \mathbb{1}_V : V \sqsubseteq X \,]$ taking constant value $1 \sqsubseteq A$.*

Denote by $\mathbb{1}$ the $\mathcal{M}$-class function $M \to \mathbb{B}$, $X \mapsto \mathbb{1}_X$.

Lemma 2.12.7. *For all $\mathbb{A}$-valued $\mathcal{M}$-sets U and $\mathcal{M}$-sets X we have*

$$U \boldsymbol{\in} \mathbb{1}_X \;=\; \bigvee_{V \sqsubseteq X} U \approx \mathbb{1}_V \,.$$

Proof. See [Bel11, Theorem 1.23]. □

An S-formula π is said to be *restricted* if every quantifier in it *appears restricted*, i.e. for every subformula of π of the form $\bigwedge_x \varphi$ the S-formula φ has the form $(x \epsilon y \to \psi)$ for some $y \in S_V$ and some S-formula ψ. We leave it to the reader to turn this into a precise definition. Logicians often call such formulas Δ_0 (see Lévy hierarchy).

The following observation is very helpful for verifying that $\mathcal{M}_{\mathbb{A}} \vDash \mathrm{ZFC}$:

Lemma 2.12.8. *For every restricted S-formula π and all tuples $\vec{x} = (x_1, \ldots, x_r)$ of pairwise distinct symbols in S_V and $\vec{a} = (a_1, \ldots, a_r) \in M^r$ and $\mathbb{1}_{\vec{a}} = \big(\mathbb{1}_{a_1}, \ldots, \mathbb{1}_{a_r}\big)$*

$$\mathcal{M} \vDash \pi[\vec{x}/\vec{a}] \;\Leftrightarrow\; \mathcal{M}_{\mathbb{A}} \vDash \pi[\vec{x}/\mathbb{1}_{\vec{a}}]\,.$$

In particular, we have for all $\mathcal{M}$-sets X, Y the following equivalences:

$$X = Y \;\Leftrightarrow\; \mathbb{1}_X \approx \mathbb{1}_Y = 1$$
$$X \sqsubset Y \;\Leftrightarrow\; \mathbb{1}_X \in \mathbb{1}_Y = 1$$

Proof. See [Bel11, Theorem 1.23] or [Jec03, Lemma 14.21]. □

Remark 2.12.9. In case $\mathbb{A} = \mathbb{F}_2$ Lemma 2.12.8 is true for general S-formulas π, i.e. also for unrestricted ones, and $\mathbb{1}$ is injective. Hence, $\mathcal{M}_{\mathbb{F}_2}$ is equivalent to $\mathcal{M}$ in the sense that $\mathbb{1}$ embeds $\mathcal{M}$ into $\mathcal{M}_{\mathbb{F}_2}$ as an elementary substructure.

To work with the boolean-valued structure $\mathcal{M}_{\mathbb{A}}$ it is convenient to know that it has witnesses. This fact can be deduced with the following useful construction:

Definition 2.12.10. *The* boolean mixture *of families $\langle A_i\rangle_{i\sqsubset I}$ and $\langle B_i\rangle_{i\sqsubset I}$ of $\mathcal{M}$-sets with $A_i \in \mathbb{A}$ and $B_i \in \mathbb{B}$ for all $i \sqsubset I$ is the $\mathbb{A}$-valued $\mathcal{M}$-set $X = \sum_{i\sqsubset I} A_i B_i$ with domain $\bigsqcup_{i\sqsubset I} \operatorname{dom}(B_i)$ such that for all $U \sqsubset \operatorname{dom}(X)$ we have*

$$\Big(\sum_{i\sqsubset I} A_i B_i\Big)(U) \;=\; \bigvee_{i\sqsubset I} A_i \wedge U \in B_i \,.$$

Lemma 2.12.11 (Mixing Lemma)**.** *Let $\langle A_i\rangle_{i\sqsubset I}$ and $\langle B_i\rangle_{i\sqsubset I}$ be families of $\mathcal{M}$-sets with $A_i \in \mathbb{A}$ and $B_i \in \mathbb{B}$ for all $i \sqsubset I$ such that $A_i \wedge A_j \leq B_i \approx B_j$ for all $i, j \sqsubset I$. Then for all $j \sqsubset I$ the following inequality holds:*

$$A_j \;\leq\; B_j \approx \sum_{i\sqsubset I} A_i B_i$$

Proof. See [Bel11, Lemma 1.25]. □

Corollary 2.12.12 (Maximum Principle)**.** *$\mathcal{M}_{\mathbb{A}}$ has witnesses.*

Proof. See [Bel11, Lemma 1.27]. □

Given an S-formula π whose distinct free variables are $x_1, \ldots, x_r$ and given furthermore $\mathbb{A}$-valued $\mathcal{M}$-sets $X_1, \ldots, X_{r-1}$, an $\mathbb{A}$-valued $\mathcal{M}$-set X_r is said to be an *existential witness for π via $x_1, \ldots, x_r \mapsto X_1, \ldots, X_r$* if $\mathcal{M}_{\mathbb{A}} \vDash \pi[(x_1, \ldots, x_r)/(X_1, \ldots, X_r)]$.

We are now in the position to state the central results about boolean-valued sets.

Theorem 2.12.13. *$\mathcal{M}_{\mathbb{A}}$ is a boolean-valued model of* ZFC*. Moreover:*

(EMP) $\mathbb{1}_{\varnothing} = \varnothing$ *is an existential witness for*

$$\bigwedge_v v \notin x \;\; \textit{via } x \mapsto \varnothing \,.$$

(PAI) *For $\mathbb{A}$-valued $\mathcal{M}$-sets X and Y the $\mathbb{A}$-valued $\mathcal{M}$-set $[X, Y]^{\mathbb{A}}$ taking constant value 1 with domain $[X, Y]$ is an existential witness for*

$$\bigwedge_v ((v \equiv x \vee v \equiv y) \leftrightarrow v \in z) \;\; \textit{via } x, y, z \mapsto X, Y, [X, Y]^{\mathbb{A}} \,.$$

(UNI) *For* $\mathbb{A}$*-valued* $\mathcal{M}$*-sets* X *the* $\mathbb{A}$*-valued* $\mathcal{M}$*-set* $\bigsqcup^{\mathbb{A}} X$ *taking constant value* 1 *with domain* $\bigsqcup[\mathrm{dom}(V) : V \in \mathrm{dom}(X)]$ *is an existential witness for*

$$\bigwedge_v (\bigvee_w (v \epsilon w \wedge w \epsilon x) \leftrightarrow v \epsilon y) \ \textit{via}\ x, y \mapsto X, \bigsqcup^{\mathbb{A}} X\,.$$

(POW) *For* $\mathbb{A}$*-valued* $\mathcal{M}$*-sets* X *the* $\mathbb{A}$*-valued* $\mathcal{M}$*-set* $\mathsf{P}^{\mathbb{A}}(X)$ *taking constant value* 1 *with domain* $[V \in [\mathrm{dom}(X) \twoheadrightarrow A] : V(U) \leq X(U)$ *for all* $U \in \mathrm{dom}(X)]$ *is an existential witness for*

$$\bigwedge_v (v \subseteq x \leftrightarrow v \epsilon y) \ \textit{via}\ x, y \mapsto X, \mathsf{P}^{\mathbb{A}}(X)\,.$$

(INF) $\mathbb{1}_\omega$ *is an existential witness for*

$$(\varnothing \epsilon x \wedge \bigwedge_v (v \epsilon x \to v \cup \{v\} \epsilon x)) \ \textit{via}\ x \mapsto \mathbb{1}_\omega\,.$$

Proof. See [Jec03, Theorem 14.24] and [Bel11, Theorem 1.33]. □

Let Ordinal(x) be an S-formula π such that for all ZFC-universes $\mathcal{N}$ the equivalence $\mathcal{N} \vDash \pi[x/\alpha] \Leftrightarrow \alpha \in \mathbb{O}^{\mathcal{N}}$ holds. Similarly, we fix three further S-formulas Cardinal(x), InjectsInto(x, y), Bijective(x, y) with the obvious properties indicated by the naming.

For each $\mathbb{A}$-valued $\mathcal{M}$-set X with Ordinal$(X) = 1$ let $\aleph_X^{\mathbb{A}}$ be an existential witness for an S-formula π via $x, y \mapsto X, \aleph_X^{\mathbb{A}}$ with the property that for all ZFC-universes $\mathcal{N}$ the equivalence $\mathcal{N} \vDash \pi[(x, y)/(\alpha, \kappa)] \Leftrightarrow (\alpha \in \mathbb{O}^{\mathcal{N}}$ and $\kappa = \aleph_\alpha^{\mathcal{N}})$ holds.

For $\mathbb{A}$-valued $\mathcal{M}$-sets X and Y we will write $X \preceq Y$ in case InjectsInto$(X, Y) = 1$ and $X \simeq Y$ in case Bijective$(X, Y) = 1$.

Lemma 2.12.14. *For all* $\mathbb{A}$*-valued* $\mathcal{M}$*-sets* X *and* Y *the following are true:*

(i) $X \preceq \mathbb{1}_{\mathrm{dom}(X)}$.

(ii) $|X| = |Y| \Rightarrow \mathbb{1}_X \simeq \mathbb{1}_Y$.

(iii) $\mathrm{Ordinal}(X) = \bigvee_{\alpha \in \mathbb{O}} X \approx \mathbb{1}_\alpha$, *in particular* $\mathrm{Ordinal}(\mathbb{1}_\alpha) = 1$ *for all* $\alpha \in \mathbb{O}$.

(iv) $\mathbb{1}_{\aleph_\alpha} \preceq \aleph_{\mathbb{1}_\alpha}^{\mathbb{A}}$ *for all* $\alpha \in \mathbb{O}$.

(v) $\mathrm{Cardinal}(\mathbb{1}_\alpha) = 1$ *for all* $\alpha \leq \omega$.

(vi) $\aleph_{\mathbb{1}_{\underline{0}}}^{\mathbb{A}} \approx \mathbb{1}_\omega = 1$.

Proof. See [Bel11, Theorems 1.44, 1.49, 1.50 and Lemma 1.52]. □

Remark 2.12.15. A perhaps less clumsy formulation of Lemma 2.12.14 could be given in terms of the ZFC-universes $\mathcal{M}_{\mathfrak{m}} = (\mathcal{M}_{\mathbb{A}}/\mathfrak{m})/_{\equiv}$ for maximal ideals $\mathfrak{m}$ of $\mathbb{A}$.

For instance, denoting by $X_{\mathfrak{m}}$ the $\mathcal{M}_{\mathfrak{m}}$-set that is the equivalence class of the $\mathcal{M}_{\mathbb{A}}$-set $\mathbb{1}_X$ in $\underline{\mathcal{M}_{\mathfrak{m}}}$, property (iv) translates into $(\aleph_\alpha)_{\mathfrak{m}} \leq \aleph_{\alpha_{\mathfrak{m}}}^{\mathcal{M}_{\mathfrak{m}}}$ for all $\mathfrak{m} \in \mathrm{Spec}(\mathbb{A})$.

Definition 2.12.16. *An* antichain *in a boolean ring $\langle A, +, \cdot\rangle$ in $\mathcal{M}$ is an $\mathcal{M}$-subset X of A such that $x \wedge y \neq 0$ for all $x, y \mathrel{\epsilon} X$.*

We say that a boolean ring $\langle A, +, \cdot\rangle$ in $\mathcal{M}$ is ccc *(or that it satisfies the* **c***ountable anti***c***hain* **c***ondition) if all of its antichains X are at most countable, i.e. $|X| \leq \omega$.*

Lemma 2.12.17. *If $\langle A, +, \cdot\rangle$ is ccc, the equivalences $X \in \mathbb{K} \Leftrightarrow \mathrm{Cardinal}(\mathbb{1}_X) = 1$ and $|X| = |Y| \Leftrightarrow \mathbb{1}_X \simeq \mathbb{1}_Y$ hold for all $\mathcal{M}$-sets X,Y and $\mathbb{1}_{\aleph_\alpha} \simeq \aleph^{\mathbb{A}}_{\mathbb{1}_\alpha}$ for all $\alpha \in \mathbb{O}$.*

Proof. See [Bel11, Theorem 1.51]. □

2.13 Forcing

We briefly discuss the relation between boolean-valued $\mathcal{M}$-sets and so-called forcing. We do this just out of curiosity. The only thing needed later from this section are Examples 2.13.5 and 2.13.8. As before, we will use the ideal language and relate it to the language of filters that is more commonly encountered in the literature.

By Theorem 2.12.13, Corollary 2.12.12, and Lemmas 1.3.9 and 1.4.2 we know that the quotients $\mathcal{M}_{\mathfrak{m}} = (\mathcal{M}_{\mathbb{A}}/\mathfrak{m})/_{\equiv}$ are ZFC-universes for all maximal ideals $\mathfrak{m}$ in $\mathbb{A}$. Furthermore, via $\mathbb{1}$ every $\mathcal{M}$-set X can be mapped to a corresponding $\mathcal{M}_{\mathfrak{m}}$-set $[\mathbb{1}_X]$. It is thus natural to wonder whether along these lines it is possible to regard $\mathcal{M}_{\mathfrak{m}}$ in a certain sense “as an extension of $\mathcal{M}$ by adjoining $\mathfrak{m}$”.

This works under (very strong) assumptions on $\mathcal{M}$ and $\mathfrak{m}$.

Namely, if $\mathcal{M}$ is *transitive*, i.e. $\mathrel{\epsilon}$ is given by $\in$, we can define another transitive $\equiv$-respecting S-structure $\mathcal{M}[\mathfrak{m}]$ with underlying set $M[\mathfrak{m}] = \{X^{\mathfrak{m}} : X \in \mathbb{B}\}$ where

$$X^{\mathfrak{m}} = \{U^{\mathfrak{m}} : U \in \mathrm{dom}(X) \text{ such that } X(U) \notin \mathfrak{m}\}.$$

If now $\mathfrak{m}$ is an *$\mathcal{M}$-prime ideal*, i.e. for families $\langle x_i\rangle_{i \mathrel{\epsilon} I}$ of $\mathcal{M}$-sets with $x_i \in \mathbb{A}$ we have $\bigwedge_{i \mathrel{\epsilon} I} x_i \in \mathfrak{m}$ only if $x_i \in \mathfrak{m}$ for some $i \mathrel{\epsilon} I$, then $\mathcal{M}[\mathfrak{m}]$ and $\mathcal{M}_{\mathfrak{m}}$ are isomorphic ([Jec03, Exercise 14.15]). In this situation, $\mathcal{M}[\mathfrak{m}]$ is called a *forcing extension* of $\mathcal{M}$.

Remark 2.13.1. By Lemma 1.1.4 all $\mathcal{M}$-prime ideals in $\mathbb{A}$ are maximal.

Theorem 2.13.2. *Let $\mathcal{M}$ be transitive and let $\mathfrak{m}$ be an $\mathcal{M}$-prime ideal in $\mathbb{A}$.*

Then $M[\mathfrak{m}]$ is a ZFC*-universe with $M \subseteq M[\mathfrak{m}]$ and $\mathfrak{m} \in M[\mathfrak{m}]$ and $\mathbb{O}^{\mathcal{M}[\mathfrak{m}]} = \mathbb{O}^{\mathcal{M}}$.*

If $\mathcal{N}$ is another transitive ZFC*-universe with $M \subseteq \underline{\mathcal{N}}$ and $\mathfrak{m} \in \underline{\mathcal{N}}$, then $M[\mathfrak{m}] \subseteq \underline{\mathcal{N}}$.*

Proof. See [Jec03, Theorem 14.5]. One checks for instance $X = \mathbb{1}^{\mathfrak{m}}_X$ for all $X \in M$. □

Remark 2.13.3. $\mathcal{M}$-prime ideals in $\mathbb{A}$ exist, if $\mathcal{M}$ is transitive and its underlying set is countable (see [Jec03, Lemma 14.4]). Since the existence of countable transitive models of a given finite subset of ZFC is provable, one can replace the hypotheses of Theorem 2.13.2 by the assumption that $\mathcal{M}$ merely has to be a model of a finite subset of ZFC large enough to prove the theorem. With this reformulation the theorem becomes provably non-vacuous and in this modified form it can be used to obtain independence results.

Instead of "adjoining $\mathcal{M}$-prime ideals of complete boolean rings in $\mathcal{M}$" many people prefer "adjoining $\mathcal{M}$-generic filters of partially ordered sets in $\mathcal{M}$". Essentially these two approaches are the same as indicated by the observations collected below.

Forcing notions and forcing language

Definition 2.13.4. *A* forcing notion in $\mathcal{M}$ *is a partially ordered set $\langle P, \leq \rangle$ in $\mathcal{M}$.*

For $p, q \sqin P$ we say that p extends *q (or that p is* stronger than *q) if $p \leq q$.*

If p and q have a common extension, i.e. if there is $r \sqin P$ with $r \leq p$ and $r \leq q$, they are called compatible*. An $\mathcal{M}$-subset D of P is said to be* dense *if every $p \sqin P$ has an extension in D, i.e. there is $r \sqin D$ with $r \leq p$.*

Example 2.13.5 (Cohen's Forcing Notion)**.** For $\mathcal{M}$-sets I let P_I be the $\mathcal{M}$-set of $\mathcal{M}$-functions $X \twoheadrightarrow \underline{2}$ with $X \sqsubseteq I$ and $|X| < \omega$ ordered by $p \leq q \Leftrightarrow p \sqsupseteq q$.

Remark 2.13.6. If $\langle A, +, \cdot \rangle$ is a boolean ring in $\mathcal{M}$, then two non-zero $p, q \sqin A$ are compatible in the partially ordered set $\langle A \setminus [0], \leq \rangle$ iff $p \wedge q \neq 0$.

Theorem 2.13.7. *For every partially ordered set $\langle P, \leq \rangle$ in $\mathcal{M}$ there is a complete boolean ring $\langle A, +, \cdot \rangle$ in $\mathcal{M}$ and a homomorphism $\iota \colon \langle P, \leq \rangle \twoheadrightarrow \langle A \setminus [0], \leq \rangle$ with dense image that* reflects compatibility*, i.e. p and q are compatible in P iff $\iota(p) \wedge \iota(q) \neq 0$.*

Moreover, $\langle \langle A, +, \cdot \rangle, \iota \rangle$ is up to isomorphism uniquely determined. We refer to it as a boolean completion *of $\langle P, \leq \rangle$.*

Proof. See [Jec03, Corollary 14.12]. $\square$

Example 2.13.8 (Cohen Algebra)**.** Consider $\langle P_I, \leq \rangle$ as in Example 2.13.5 and view $[I \twoheadrightarrow \underline{2}] = \ast_{i \sqin I} \underline{2}$ as a topological space in $\mathcal{M}$ equipped with the product topology where $\underline{2}$ carries the discrete topology. Then the $\mathcal{M}$-set A_I of regular $\mathcal{M}$-subsets of $[I \twoheadrightarrow 2]$ becomes a complete boolean ring $\langle A_I, +, \cdot \rangle$ in $\mathcal{M}$ with the operations described in Exercise 4 of Problem Set 1.

Let $\iota \colon P_I \twoheadrightarrow A_I \setminus [0]$ be given by $\iota(p) = \left[f \sqin [I \twoheadrightarrow \underline{2}] : p \sqsubseteq f \right]$.

It is not hard to see that $\langle \langle A_I, +, \cdot \rangle, \iota \rangle$ is a boolean completion of $\langle P_I, \leq \rangle$. Indeed, for all $p, q \sqin P$ we have $p \sqsupseteq q \Rightarrow \iota(p) \sqsubseteq \iota(q)$. Furthermore, the image of ι is dense in $A_I \setminus [0]$ because it is a basis of the topology of $[I \twoheadrightarrow \underline{2}]$ that consists of closed and open (so in particular regular) $\mathcal{M}$-subsets of $[I \twoheadrightarrow \underline{2}]$. Finally, two $\mathcal{M}$-functions $p, q \sqin P$ agree on their common domain $\mathrm{dom}(p) \sqcap \mathrm{dom}(q)$ iff $\iota(p) \sqcap \iota(q)$ is non-empty.

Convention 2.13.9. Fix a forcing notion $\langle P, \leq \rangle$ in $\mathcal{M}$ and moreover $\langle A, +, \cdot \rangle$ and ι as in Theorem 2.13.7. We then continue to use the notation of Convention 2.12.1 and consider $\mathbb{P} = \sqin^{-1}(P)$ as a partially ordered set under $\hat{\leq}$.

We can now introduce the *forcing relation* $\Vdash$. Some of its important properties are collected in [Jec03, Theorem 14.7] and [Bel11, Theorem 2.5]. All of them are evident from the calculation rules for boolean rings and the fact that $\mathcal{M}_{\mathbb{A}}$ has witnesses.

Definition 2.13.10. *Let* $\sigma = \pi^{(\mathcal{M}_{\mathbb{A}})^{\vec{x}}_{\vec{a}}}$ *where* π *is an* S*-formula and* $\vec{x} = (x_1, \ldots, x_r)$ *a tuple of pairwise distinct symbols in* S_V *and* $\vec{a} = (a_1, \ldots, a_r) \in \mathbb{B}^r$.

For $p \sqsubseteq P$ *one says that* p forces σ *if* $\iota(p) \leq \sigma$. *In this case we write* $p \Vdash \sigma$.

Lemma 2.13.11. $p \Vdash \sigma$ *iff* $\mathcal{M}/\mathfrak{m} \vDash \pi[\vec{x}/\vec{a}]$ *for every* $\mathfrak{m} \in \mathrm{Spec}(\mathbb{A})$ *with* $\iota(p) \notin \mathfrak{m}$.

Proof. This follows readily from Corollary 2.12.12 and Lemma 1.3.9. □

Forcing extensions by filters

Definition 2.13.12. *A* filter in $\mathbb{P}$ *is a non-empty subset of* $\mathbb{P}$ *with* $q \geq p \in F \Rightarrow q \in F$ *for all* $p, q \in \mathbb{P}$ *and such that for all* $p, q \in F$ *there is* $r \in F$ *with* $r \leq p$ *and* $r \leq q$.

A filter in $\mathbb{P}$ *is* $\mathcal{M}$-generic *if it intersects every dense* $\mathcal{M}$*-subset of* $\mathbb{P}$ *non-trivially.*

Lemma 2.13.13. *The assignment* $F \mapsto \mathfrak{m}_F = \{x \in \mathbb{A} : x < \iota(p) \text{ for every } p \in F\}$ *maps* $\mathcal{M}$*-generic filters in* $\mathbb{P}$ *to* $\mathcal{M}$*-prime ideals in* $\mathbb{A}$.

Proof. See [Jec03, Lemma 14.13 and Exercise 14.10]. □

If $\mathcal{M}$ is *transitive* and F is an $\mathcal{M}$-generic filter in $\mathbb{P}$ one can define again a transitive $\equiv$-respecting S-structure $\mathcal{M}[F]$ with underlying set $M[F] = \{X^F : X \in \mathbb{B}\}$ where

$$X^F = \{U^F : U \in \mathrm{dom}(X) \text{ such that } X(U) \in F\}.$$

It turns out that $F \in M[F] = M[\mathfrak{m}_F]$. Moreover, $\mathcal{M}[F]$ has the universal property that $M[F] \subseteq \underline{N}$ for all transitive ZFC-universes $\mathcal{N}$ with $M \subseteq \underline{N}$ and $F \in \underline{N}$.

2.14 Independence of GCH

We will now apply the theory of boolean-valued sets to construct ZFC-universes $\mathcal{N}$ where the cardinality of $\mathsf{P}^{\mathcal{N}}(\omega^{\mathcal{N}})$ is $\aleph_{\underline{2}}^{\mathcal{N}}$. Obviously, such $\mathcal{N}$ will violate GCH.

Convention 2.14.1. Fix an $\mathcal{M}$-ordinal α such that $\aleph_\alpha^\omega = \aleph_\alpha$ and let $I = \omega * \aleph_\alpha$. With this choice of I, let $\langle P, \leq \rangle = \langle P_I, \leq \rangle$ be Cohen's Forcing Notion from Example 2.13.5 and take $\langle A, +, \cdot \rangle$ to be its boolean completion described in Example 2.13.8.

Lemma 2.14.2. $\langle A, +, \cdot \rangle$ *is ccc and* $|A| = \aleph_\alpha$.

Proof. See [Bel11, Corollary 2.11] and use the assumption $\aleph_\alpha^\omega = \aleph_\alpha$. □

The choice of I above provides us with "many $\mathbb{A}$-valued $\mathcal{M}$-subsets of $\mathbb{1}_\omega$" that do not arise from $\mathcal{M}$-subsets of ω. Namely, we will consider the following ones:

Definition 2.14.3. *Define* $\langle U_\gamma \rangle_{\gamma < \aleph_\alpha}$ *to be the family of* $\mathbb{A}$*-valued* $\mathcal{M}$*-sets given by* $\mathrm{dom}(U_\gamma) = \mathrm{dom}(\mathbb{1}_\omega) = [\mathbb{1}_\sigma : \sigma < \omega]$ *for* $\gamma < \aleph_\alpha$ *such that for* $\sigma < \omega$

$$U_\gamma(\mathbb{1}_\sigma) = \left[f \sqsubseteq [I \rightharpoonup \underline{2}] : f(\langle \sigma, \gamma \rangle) = \underline{1}\right].$$

Observe that $U_\gamma(\mathbb{1}_\sigma)$ is the preimage of $\underline{1}$ under the canonical projection $\bigast_{i\in I}\underline{2} \twoheadrightarrow \underline{2}$ onto the $\langle\sigma,\gamma\rangle$-th component. Hence, it is open and closed and thus regular, so it belongs to A. The closedness also shows that $\neg U_\gamma(\mathbb{1}_\sigma) = A \setminus U_\gamma(\mathbb{1}_\sigma)$.

Lemma 2.14.4. *For all $\gamma < \delta < \aleph_\alpha$ and $p \in P$ and $\langle\sigma,\gamma\rangle \in \mathrm{dom}(p)$ we have:*

(i) $U_\gamma \Subset \mathbb{1}_\omega = 1$.

(ii) $U_\gamma(\mathbb{1}_\sigma) = \mathbb{1}_\sigma \in U_\gamma$.

(iii) $\iota(p) \le \mathbb{1}_\sigma \in U_\gamma \Leftrightarrow p(\langle\sigma,\gamma\rangle) = \underline{1}$ *and* $\iota(p) \le \neg(\mathbb{1}_\sigma \in U_\gamma) \Leftrightarrow p(\langle\sigma,\gamma\rangle) = \underline{0}$.

(iv) $U_\gamma \approx U_\delta = 0$.

Proof. (i) $U_\gamma \Subset \mathbb{1}_\omega = \bigwedge_{\tau<\omega}(U_\gamma(\mathbb{1}_\tau) \to \mathbb{1}_\tau \in \mathbb{1}_\omega) = 1$ since $\mathbb{1}_\tau \in \mathbb{1}_\omega = 1$ for $\tau < \omega$.

(ii) $\mathbb{1}_\sigma \in U_\gamma = \bigvee_{\tau<\omega}(\mathbb{1}_\sigma \approx \mathbb{1}_\tau \wedge U_\gamma(\mathbb{1}_\tau)) = U_\gamma(\mathbb{1}_\sigma)$ since $\mathbb{1}_\sigma \approx \mathbb{1}_\tau = 0$ for $\tau \ne \sigma$.

(iii) Let $f \in [I \twoheadrightarrow \underline{2}]$. We must show $f \sqsubseteq \iota(p) \Rightarrow f \sqsubseteq U_\gamma(\mathbb{1}_\sigma)$ iff $p(\langle\sigma,\gamma\rangle) = \underline{1}$ and $f \sqsubseteq \iota(p) \Rightarrow f \not\sqsubseteq U_\gamma(\mathbb{1}_\sigma)$ iff $p(\langle\sigma,\gamma\rangle) = \underline{0}$. This is clear by definition of ι and U_γ.

(iv) Assume $U_\gamma \approx U_\delta \ne 0$. Since the image of ι is dense in $A \setminus [0]$ there is $q \sqsubseteq P$ such that $\iota(q) \le U_\gamma \approx U_\delta$. Because of $|\mathrm{dom}(q)| < \omega$ we can therefore choose $\sigma < \omega$ such that $\langle\sigma,\tau\rangle \not\sqsubseteq \mathrm{dom}(q)$ for all $\tau < \aleph_\alpha$. Then we can pick any $p \sqsubseteq P$ with $p \le q$ (i.e. extending q) and $p(\langle\sigma,\gamma\rangle) = \underline{1}$ and $p(\langle\sigma,\delta\rangle) = \underline{0}$. Using (iii) we can conclude $\iota(p) \le \mathbb{1}_\sigma \in U_\gamma \wedge \neg(\mathbb{1}_\sigma \in U_\delta) \le \neg(U_\gamma \approx U_\delta)$. But we also have $\iota(p) \le \iota(q) \le U_\gamma \approx U_\delta$. Combining these two inequalities yields the absurd $\iota(p) = 0$. $\square$

Denote by InjectiveFunctionFromTo(u,x,y) some fixed S-formula π such that for all ZFC-universes $\mathcal{N}$ and $\mathcal{N}$-sets U,X,Y we have $\mathcal{N} \vDash \pi[(u,x,y)/(U,X,Y)]$ if and only if U is an injective $\mathcal{N}$-function with domain X and image contained in Y.

We will use the non-surprising notation $\langle X,Y\rangle^{\mathbb{A}}$ for $\big[X,[X,Y]^{\mathbb{A}}\big]^{\mathbb{A}}$ where $X,Y \in \mathbb{B}$.

Theorem 2.14.5. $\mathsf{P}^{\mathbb{A}}\Big(\aleph^{\mathbb{A}}_{\mathbb{1}_{\underline{0}}}\Big) \simeq \aleph^{\mathbb{A}}_{\mathbb{1}_\alpha}$.

Proof. Abbreviate $\mathsf{P}_\omega = \mathsf{P}^{\mathbb{A}}(\mathbb{1}_\omega)$. By Lemmas 2.12.14, 2.12.17 and 2.14.2 we have $\mathbb{1}_\omega \approx \aleph^{\mathbb{A}}_{\mathbb{1}_{\underline{0}}} = 1$ and $\mathbb{1}_{\aleph_\alpha} \simeq \aleph^{\mathbb{A}}_{\mathbb{1}_\alpha}$. So it is sufficient to show $\mathsf{P}_\omega \simeq \mathbb{1}_{\aleph_\alpha}$.

On the one hand, using once more Lemmas 2.12.14, 2.12.17 and 2.14.2, we have the inequality $|\mathrm{dom}(\mathsf{P}_\omega)| \le |[\omega \twoheadrightarrow A]| = |A|^\omega = \aleph_\alpha^\omega = \aleph_\alpha$, so $\mathsf{P}_\omega \preceq \mathbb{1}_{\mathrm{dom}(\mathsf{P}_\omega)} \preceq \mathbb{1}_{\aleph_\alpha}$.

On the other hand, in view of Lemma 2.14.4 (i) and (iv), it is possible to show that InjectiveFunctionFromTo$(U, \mathbb{1}_{\aleph_\alpha}, \mathsf{P}_\omega) = 1$ where U is defined as the $\mathbb{A}$-valued $\mathcal{M}$-set with $\mathrm{dom}(U) = [\langle\mathbb{1}_\gamma, U_\gamma\rangle^{\mathbb{A}} : \gamma < \aleph_\alpha]$ taking constant value 1 (for a more detailed argument see the proof of [Bel11, Theorem 2.12]). Hence, $\mathbb{1}_{\aleph_\alpha} \preceq \mathsf{P}_\omega$, too. $\square$

Corollary 2.14.6. *GCH is independent of* ZFC*, i.e. if there is any* ZFC*-universe at all, then some* ZFC*-universes satisfy GCH and some do not.*

Proof. By the results of § 2.11 we may replace $\mathcal{M}$ by $\mathcal{M}|_{\mathbb{L}}$ to make sure that $\mathcal{M}$ satisfies GCH. Then $\underline{2}$ is a valid choice for α because of $\aleph_{\underline{2}}^\omega = \aleph_{\underline{2}}$ by Lemmas 2.10.26 and 2.10.38. Assuming $\alpha = \underline{2}$ Theorem 2.14.5 shows that $\mathcal{M}_{\mathbb{A}}$ does not satisfy GCH since $\aleph^{\mathbb{A}}_{\mathbb{1}_{\underline{2}}} \not\simeq \aleph^{\mathbb{A}}_{\mathbb{1}_{\underline{1}}}$ according to Lemmas 2.12.14, 2.12.17 and 2.14.2. $\square$

References

[Bel11] J. L. Bell, *Set Theory*. Oxford University Press, 2011.

[Ges10] S. Geschke, *Einführung in die Logik und Modelltheorie*. Bonn, 2010.

[Jec03] T. Jech, *Set Theory*. Springer-Verlag, Berlin, 2003.

[Kun11] K. Kunen, *Set Theory*. College Publications, London, 2011.

[Ros69] J. B. Rosser, *Simplified Independence Proofs. Boolean Valued Models of Set Theory*. Academic Press, New York–London, 1969.