Weak well orders

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Dandolo Flumini

von Horgen ZH

Leiter der Arbeit:
Prof. Dr. G. Jäger
Institut für Informatik und angewandte Mathematik

Weak well orders

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von
Dandolo Flumini
von Horgen ZH

Leiter der Arbeit:
Prof. Dr. G. Jäger
Institut für Informatik und angewandte Mathematik

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen.

Bern, 12. Dezember 2013

Der Dekan:
Prof. Dr. S. Decurtins

Acknowledgments

I would like to thank, first and foremost, my advisor Prof. Dr. Gerhard Jäger, for without his excellent guidance, motivation and patience, this thesis would never have been written. I would also like to express my gratitude to Prof. Dr. Wolfram Pohlers for co-refereeing this thesis.

I thank Philip Isler for his help in making this thesis give less evidence that English is not my native language.

Further thanks go to all members of the Logic and Theory Group for the good times and interesting discussions. In particular, I would like to thank Dieter Probst and Kentaro Sato for sharing their profound knowledge and insights.

Last but not least, I would like to thank my family for their continued patience and support.

Contents

1 Introduction **9**

1.1 About the program . . . 9
1.2 Subsystems of second order arithmetic . . . 12
1.3 Sets and classes . . . 19
1.4 The theories ZFC, NBG and MK . . . 23

2 Trees and κ-well orders **53**

2.1 κ-well orders . . . 53
2.2 Trees . . . 58

3 Weak well orders **75**

3.1 The logical complexity of weak well orders . . . 78
3.2 Elementary transfinite recursion . . . 88
3.2.1 The uniform recursion axiom . . . 107
3.3 Comparing weak well orders . . . 114
3.3.1 Comparability of well orders . . . 115
3.3.2 Order-types and comparability of weak well orders . . . 116
3.4 Bar induction and V-model reflection . . . 133

1 Introduction

1.1 About the program

The program of which this theses is taking part of, aims at investigating to what extent some of the insights that have been obtained in second order arithmetic have analog counterparts in second order set theory. In short terms, the research program can be condensed in terms of its leading question as it appears in [LL13]:

> *What happens if we replace Peano arithmetic and subsystems of arithmetic by Zermelo-Fraenkel set theory (with or without the axiom of choice) and subsystems of Morse-Kelley theory of sets and classes, respectively? Which proof-theoretic results have direct analogues and for which results do such analogues not exist?*

The term "analog" above has to be understood in the sense that it is not the purpose of the program to generalize results from arithmetic to set theory, but instead to investigate to what extent set theory resembles arithmetic. For this thesis, we employ a specific analogy, to be made more precise below, relative to which we investigate the aforementioned resemblance.

A pivotal question that characterizes the way we translate postulates from arithmetic to the realm of sets and classes, is how to make the difference between "small" and "large" objects respectively. While in arithmetic, the small objects, numbers, represent finite entities, all infinite sets of natural numbers are large. The general approach that we follow in our translation is that the small objects in set theory are sets while the large structures are to be represented by proper classes. Given that point of view, we found it very natural, though not "standard", to axiomatize our base theory NBG, *Von Neumann Bernays Gödel* *set theory*, using the principle of *limitation of size*, which in fact can be seen as a formal version of our attitude.

As mentioned above, our base theory is to be NBG, Von Neumann Bernays Gödel set theory. This system will play the role that ACA_0 plays in arithmetic, while the theory MK, *Morse Kelley set theory*, is to be the set theoretic counterpart of full second order arithmetic. There are some well known analogies between those theories and their relationship between other systems of arithmetic and set theory respectively. Both base theories are finitely axiomatizable extensions of the most prominent first order theories of numbers and sets respectively, being PA, *Peano arithmetic*, and ZFC, *Zermelo Fraenkel set theory with choice.*

Our investigations are about how far similar analogies go when different extensions of ACA_0 and NBG are considered. We want to ask ourselves which results and, yet more interestingly, which techniques from arithmetic can be mimicked in set theory. Prominent examples of such number theoretic results are: the pairwise equivalence[1] of

- Arithmetical transfinite recursion
- Comparability of well orders
- Every arithmetically defined positive operator has *some* fixed point.

or

- Π^1_1 comprehension
- Every arithmetically defined positive operator form has a (unique) *least* fixed point

and

- ω-model reflection
- bar induction.

Some of the theories mentioned above have straightforward translations to set theory (e.g. comprehension schemata), and the resulting systems behave similar

[1]Over ACA_0.

as in the arithmetical setting (see [LL13], and [JK10] where it is shown that Σ^1_1 choice over NBG can be characterized by iterated comprehension similarly as in the arithmetical case). Other theories, however, have different but equivalent formulations in arithmetic that yield theories of sets and classes that are not pairwise equivalent (e.g. see [Sat12] for set theoretically inequivalent formulations of arithmetical transfinite recursion), in these cases, the formulation to be translated has to be carefully chosen. Other theories yet have obvious translations but the resulting theories behave completely different from their arithmetical counterpart (e.g. see [Sat13] where it is shown that, in the set theoretic setting, least fixed points of monotone elementary operators can be obtained from arbitrary fixed points of monotone elementary operators).

In the course of our research, it has become apparent that an eminent source of dissimilarities between set theory and arithmetic lies in the fact that in the setting of sets and classes, the predicate that states the well foundedness of a given relation is much less expressive than it is in the arithmetical setting. In fact, in set theory, it can be expressed by an elementary formula whether or not a given relation is well founded (cf. Definition 181). In arithmetic, on the contrary, every Π^1_1 statement can be reduced to the question whether or not some specific arithmetical relation is well founded. In an effort to preserve at least some of the manifold implications of this situation for set theory, we introduce the notion of what we call a weak well order. We then establish, via a set theoretic version of Kleene's Σ^1_1 normal form lemma (cf. Theorem 172), the Π^1_1 completeness of the weak well foundedness predicate (cf. Theorem 173), and thus that there is no Σ^1_1 formula that can express the weak well foundedness of all relations (cf. Theorem 178). Despite having the notion of a weak well order and an associated normal form lemma, the situation becomes quite different from the situation in arithmetic. Weakly well founded trees, the set theoretical analog to well founded trees in arithmetic, for example, can have branches of any ordinal length and thus are much more complicated objects than the usual well founded trees familiar from arithmetic (cf. Example 139). A first evidence for this asymmetry lies in the fact that in arithmetic, the schema of arithmetical comprehension suffices to prove König's Lemma while the corresponding principle in set theory has a much greater consistency strength than NBG or ZFC (cf. Remark 155).

1.2 Subsystems of second order arithmetic

In this section, we give a very short overview of the general setup for studying subsystems of second order arithmetic. We will only quote those results that are of interest regarding our subsequent work in subsystems of set-class theory. As the purpose of this section is mainly for reference, we will be very brief and we will give reference to proofs instead of actually proving facts. Generally, for a detailed presentation of the material in this section, the reader is referred to the book [Sim98]. Obviously, none of the results presented in this section are due to the author of this thesis.

Definition 1 (Language of arithmetic)**.** In addition to all kinds of brackets and the usual logical symbols $=, \rightarrow, \leftrightarrow, \wedge, \vee, \forall, \exists$ and $\neg$, the language $\mathcal{L}^2_A$ of second order arithmetic contains two sorts of variables, *numerical variables* $x, y, z, n, m, \ldots$ (lower case letters) and *set variables* $X, Y, Z, \ldots$ (upper case letters) that are meant to range over natural numbers and sets thereof respectively. Further, the language $\mathcal{L}^2_A$ contains the relation symbols $\in$ and $<$, the (binary) function symbols $+$ and $\cdot$ as well as the constants 0 and 1.

Definition 2 (Terms and formulas)**.** The *terms* of $\mathcal{L}^2_A$ are built up inductively as follows:

1. Numerical variables and constants are terms.
2. If t and r are terms then the expressions $(t + r)$ and $(t \cdot r)$ are also terms.

Atomic formulas are expressions of the form $t < s$, $t = s$ or $t \in X$ where s, t are terms and X is a set variable. The *formulas* of arithmetic are the elements of the least set of $\mathcal{L}^2_A$ expressions that has the following closure properties.

1. Atomic formulas are formulas.
2. If φ and ψ are formulas, then $(\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \rightarrow \psi)$ and $(\neg\varphi)$ are also formulas.
3. If x is any numerical variable and φ does not contain any of the substrings $\exists x$ or $\forall x$, then the expressions $\exists x\,(\varphi)$ and $\forall x\,(\varphi)$ are also formulas.

4. If X is any set variable and φ does not contain any of the substrings $\exists X$ or $\forall X$, then the expressions $\exists X\,(\varphi)$ and $\forall X\,(\varphi)$ are also formulas.

If any variable (numerical or set) appears in a formula right next to a quantifier $(\forall, \exists)$, then this variable is called a *bound* variable of the formula. Formulas that have no bound set variables are called *arithmetical* formulas. Formulas that do not contain any second order variables are $\mathcal{L}_A^1$ *formulas.*

Definition 3. An $\mathcal{L}_A^2$ *structure* is a 7-tuple $\mathfrak{M} = (M, S, <, +, \cdot, 0, 1)$ where:

1. M is a nonempty set and S is a nonempty set of subsets of M; it is further assumed that M and S have no common elements.

2. The functions $+$ and $\cdot$ are binary operations on M.

3. The constants 0 and 1 are elements of M and $<$ is a binary relation on M.

Definition 4 (ACA_0)**.** The axioms of the theory ACA_0 of *arithmetical comprehension* are given from all the axioms of Peano arithmetic (cf. Definition IX.1.4 in [Sim98]) together with the universal closure of the following principles:

1. *Induction for sets*

$$\big(0 \in X \wedge \forall x\,(x \in X \to x + 1 \in X)\big) \to \forall x\,(x \in X).$$

2. *Arithmetical comprehension*

$$\exists X\, \forall x\, \big(x \in X \leftrightarrow \varphi(x)\big)$$

for arithmetical formulas φ that do not contain the variable X.

The following definitions and observations can all be made within the theory ACA_0.

Definition 5 (Trees and sequences in arithmetic)**.** Let $\langle \cdot, \cdot \rangle$ denote a standard coding function for pairs of natural numbers as natural numbers. A finite *sequence* of natural numbers $s : \{0, 1, \ldots, n-1\} \to \mathbb{N}$, also denoted as $(s_0, \cdots, s_{n-1})$ or $(s_i)_{i<n}$, is coded as a finite set of ordered pairs (i.e. a finite set of natural

numbers) in the usual way. We write $l(s)$ to mean the *length* of a sequence s, that is, the least natural number for which the sequence is not defined. If s and t are (codes of) sequences, then we write $s \frown t$ to mean the unique sequence of length $l(s) + l(t)$ that satisfies the equations

$$s \frown t(k) = \begin{cases} s(k) & \text{if } k < l(s) \\ t(n) & \text{if } k = l(s) + n \text{ with } n < l(t) \\ \text{undefined} & \text{otherwise.} \end{cases}$$

A sequence s is an *extension* of the sequence t if a sequence r such that $t \frown r = s$ exists. We write $t \sqsubset s$ to express that s extends t - or equivalently, that t is an initial segment of s. A tree is a set of (codes of) finite sequences that is closed under initial segments. For a function $f : \mathbb{N} \to \mathbb{N}$, we write $f[n]$ for the sequence $(f(i))_{i<n}$ and we say that f is a path through a tree T if $f[n]$ is an element of T for every natural number n. A tree T is *well founded*, if no path through T exists. A tree is *finitely branching* if for every sequence $s \in T$ there are only finitely many natural numbers n such that $s \frown (n) \in T$.

Definition 6 (Well orderings in arithmetic)**.** A set R is a *linear ordering* of a set X if every element of R is (a code of) a pair of natural numbers from X and if

$$\forall x, y \in X\, (x \neq y \to \langle x, y\rangle \in R \vee \langle y, x\rangle \in R)$$

and

$$\forall x, y, z \in X\, ((\langle x, y\rangle \in R \wedge \langle y, z\rangle \in R) \to \langle x, z\rangle \in R)$$

and

$$\forall x, y \in X\, (\langle x, y\rangle \in R \to \langle y, x\rangle \notin R)$$

hold. If R is a linear ordering of X and R further satisfies that

$$\forall Y\, (\exists y\, (y \in Y \cap X) \to \exists y \in X \cap Y\, \neg\exists y' \in X \cap Y\, (\langle y', y\rangle \in R)),$$

then R is called a *well ordering* of X and we write $\mathsf{WO}((X, R))$, if X is the set of all natural numbers, then we might just write $\mathsf{WO}(R)$.

A key feature of trees in arithmetic is that for every Π^1_1 formula, that is, any formula of the form $\forall X\, \varphi(X)$ where φ is arithmetical, one can construct a tree T_φ such that $\forall X\, \varphi(X)$ holds if and only if T_φ is a well founded tree. As a consequence, we have the following theorem.

Theorem 7. *There exists no Σ^1_1 formula Ψ, that is a formula that takes the form $\Psi \equiv \exists X\, \psi(X)$ with arithmetical ψ, that satisfies $\forall Y\, (\mathsf{WO}(Y) \leftrightarrow \Psi(Y))$.*

Proof. See for example Theorem V.1.9 in [Sim98]. □

These facts give rise to many important results and widely applicable techniques from arithmetic, to name just a few: the pairwise equivalence of the theories ATR_0, CWO, FP_0 (cf. Theorem 15 and the references as given there) and more generally, all kinds of results that are obtained by applications of pseudo hierarchy arguments. The fact that in set theory, the well foundedness can be expressed with an elementary formula is a key issue when set theoretic analogs to the aforementioned theories are investigated. It is in a twofold way that dissimilarities between arithmetic and set theory arise; many arithmetic arguments cannot be carried out in set theory, and in some cases the fact that well foundedness can be expressed by an elementary formula paves the way to use a theory to prove the consistency of another theory that was equivalent in arithmetical terms. In an attempt to preserve some of the implications of Theorem 7 for set theory, as mentioned before, we will be able to state and prove a similar theorem for what we will call "weak well orders". However, since weak well orders lack some of the important properties of well orders, we will not be able to restore the as it presents itself here, in the set theoretic situation.

Lemma 8 (König's Lemma)**.** *It is provable in ACA_0 that every finitely branching infinite tree has a path.*

Proof. This is a standard result, for a proof see for example Theorem III.7.2 in [Sim98]. □

Remark 9. König's Lemma will be a first striking example of the dissimilarities of arithmetic and set theory. While the original statement can of course be formalized and proved in set theory, the *analog* statement, being that any proper class sized tree that is set branching has a path of length On, is, as we will see in the second chapter, not even even provable in the theory MK (cf. Remark 155).

The theories ATR_0, CWO and FP_0

Definition 10 (Arithmetical transfinite recursion)**.** We assign to each arithmetical formula $\varphi(X,x)$ a corresponding operator $\Gamma_\varphi : \{X \mid X \subset \mathbb{N}\} \to \{X \mid X \subset \mathbb{N}\}$ with $\Gamma_\varphi(X) = \{x \mid \varphi(X,x)\}$. The schema of *arithmetical transfinite recursion* is the statement that for every arithmetical formula $\varphi(X,x)$ and every well ordering (X,R), there exists a set $H \subset \mathbb{N}$ that satisfies for all natural numbers $b \in X$ the equation[2]

$$(H)_b = \Gamma_\varphi(\{\langle a,y\rangle \in H \mid aRb\}),$$

where $(H)_b$ is an abbreviation for $\{x \mid \langle b,x\rangle \in H\}$. The theory ATR_0 is ACA_0 together with the schema of arithmetical transfinite recursion.

Definition 11 (Comparability of well orders)**.** For two linear orderings (X,R) and (Y,S) we write $(X,R) \preceq (Y,S)$ if there exists a function $F : X \to Y$ such that

$$\forall x,y \in X\,(xRy \leftrightarrow F(x)SF(y))$$

and

$$\forall y \in Y\,\forall x \in X\,(ySF(x) \to \exists x' \in X\,F(x') = y).$$

The *comparability of well orderings principle* is the statement that for any two well orderings (X,R) and (Y,S), either $(X,R) \preceq (Y,S)$ or $(Y,S) \preceq (X,R)$ holds. The theory that consists of ACA_0 together with the comparability of well orderings principle is called CWO.

Definition 12. A formula $\varphi(X)$ in negation normal form is called *positive* in X, denoted by $\varphi(X^+)$, if it contains no implications and no subformula of the form $t \notin X$, where t ranges over terms. A formula in negation normal form is negative in X, if the negation normal form of its negation is a formula that is positive in X.

[2]Here and afterwards we write xRy to mean that $\langle x,y\rangle \in R$.

Lemma 13. *If $\varphi(X, x)$ is positive in X, then the corresponding operator $\Gamma_\varphi(X)$ is monotone in X, that is it satisfies the equation*

$$X \subset Y \to \Gamma_\varphi(X) \subset \Gamma_\varphi(Y)$$

for all sets X and Y.

Proof. By induction on the build up of $\varphi(X, x)$, proving simultaneously that operators where in the corresponding formula $\psi(X, x)$ is negative in X satisfy

$$X \subset Y \to \Gamma_\psi(Y) \subset \Gamma_\psi(X)$$

for all sets X and Y. □

Definition 14 (Fixed points of positive arithmetical operators)**.** The term FP_0 stands for the theory ACA_0 enriched by the schema

$$\exists X\, (X = \Gamma_\varphi(X))$$

where $\varphi(X, x)$ ranges over arithmetic formulas that are positive in X.

Theorem 15. *The theories ATR_0, FP_0 and CWO are all equivalent (over ACA_0).*

Proof. For the implication $\mathsf{ATR}_0 \equiv \mathsf{FP}_0$ see [Avi96] Theorem 3.1, and for the equivalence $\mathsf{ATR}_0 \equiv \mathsf{CWO}$ see Theorem V.6.8 in [Sim98]. □

The theories Π^1_1-CA_0 and LFP_0

Definition 16 (Least fixed points of positive arithmetical operators)**.** The theory LFP of *least fixed points* of positive arithmetical operators is ACA_0 augmented by the schema

$$\exists X\, \big(\Gamma_\varphi(X) = X \wedge \forall Y\, (\Gamma_\varphi(Y) = Y \to X \subset Y)\big)$$

for all arithmetical formulas φ that are positive in X.

Definition 17 (Π^1_1 comprehension)**.** The theory Π^1_1-CA_0 consists of all the axioms of ACA_0 together with the schema

$$\exists X\, \forall k\, \big(k \in X \leftrightarrow \varphi(k)\big)$$

for all Π^1_1 formulas φ in which the variable X does not occur.

Theorem 18. *The theories* Π^1_1-CA_0 *and* LFP_0 *are equivalent over* ACA_0.

Proof. This follows directly from Theorem VI.1.11 of [Bar75]. $\square$

The theories Π^1_n-BI_0 and Π^1_{n+1}-RFN_0

Definition 19 (Π^1_n bar induction)**.** Let $n > 0$ be a natural number. The theory Π^1_n-BI_0 consists of all axioms of ACA_0 together with the schema

$$\forall X \left(\mathsf{WO}(X) \to \mathsf{TI}(X, \varphi)\right)$$

where $\varphi(x)$ is a Π^1_n formula, and $\mathsf{TI}(X, \varphi)$ is an abbreviation for the formula

$$\forall n \left(\forall k \left(\langle k, n\rangle \in X \to \varphi(k)\right) \to \varphi(n)\right) \to \forall n \, \varphi(n).$$

Definition 20. For all sets X and all formulas φ, we define the relativization φ^X of φ to X on the build up of φ:

- If φ is an arithmetical formula, then $\varphi = \varphi^X$.
- For any of the connectives $\circ \in \{\to, \land \lor\}$ we define $\varphi^X = \psi_1^X \circ \psi_2^X$, if $\varphi = \psi_1 \circ \psi_2$.
- If $\varphi = \neg\psi$ then we fix $\varphi^X = \neg(\psi^X)$.
- If $\varphi = \exists Y \, \psi(Y)$, then we fix $\varphi^X = \exists k \, \psi^X((X)_k)$.
- If $\varphi = \forall Y \, \psi(Y)$, then we fix $\varphi^X = \forall k \, \psi^X((X)_k)$.

For a finite collection $\mathcal{F}$ of closed formulas, we further write $\mathcal{F}^X$ to mean the conjunction of all elements of $\mathcal{F}$ relativized to X.

Definition 21 (Π^1_{n+1} reflection)**.** Let $n > 0$ be a natural number. The theory Π^1_{n+1}-RFN_0 of Π^1_{n+1} *ω-model reflection* consists of ACA_0 augmented by the schema

$$\varphi(X) \to \exists U \left(\exists k \left(X = (U)_k\right) \land \sigma^U_{\mathsf{ACA}_0} \land \varphi^U(X)\right)$$

where φ is a Π^1_{n+1} formula with at most X as a free set variable, and σ_{ACA_0} is a finite axiomatization of ACA_0.

Theorem 22. *For all natural numbers $n > 0$ it is provable in ACA_0 that the theories Π^1_{n+1}-RFN_0 and Π^1_n-BI_0 are equivalent.*

Proof. See the main Theorem in [JS99]. $\square$

Corollary 23. *The theories Π^1_∞-RFN_0 and Π_∞-BI_0 are equivalent over ACA_0.*

1.3 Sets and classes

As we all know, in classical first order set theory there are formulas whose extensions are not sets and thus not objects of the theory under consideration. Probably the most prominent example is Russel's antinomy

$$R = \{x \mid x \notin x\},$$

which by pain of contradiction cannot be accepted as a set. Similarly, the universe

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

is also unacceptable as a set. This deficiency of ZFC is usually resolved in one of the following two ways. The most popular approach to the problem is to keep working in ZFC and treat classes as informal objects. In this approach, all proper objects are sets, while classes merely play the role of a comfortable way to speak about properties of sets (cf. the initial discussion of §9 in [Kun80]). The second approach is to work in a theory, most commonly NBG, in which formal statements about classes are possible (cf. the second paragraph in §5 of [SF10]). As it is our undertaking to investigate the impact of a variety of class existence axioms, the latter approach is a natural choice for our undertaking. In the setting of NBG, sets are the usual objects of ZFC while classes are either sets or proper classes where the latter are (some of) the collections of sets that are larger than any set.

The language of second order set theory

The language of second order set theory is a two sorted language where the first sort of variables, consisting of lower case letters, range over sets and the second kind of variables, upper case letters, range over classes (not necessarily proper classes tough). As usual, all sorts of variables may be subscripted as needed. The language $\mathcal{L}_2$ of second order set theory has only one non logical symbol, $\in$, which stands for the membership relation (of a set to a class). Also, in the formal context we will freely use well known abbreviations such as for example $x \subset Y \equiv \forall y\,(y \in x \to y \in Y)$, $x = \varnothing \equiv \forall y\,\neg(y \in x)$, etc.

Definition 24. The languages $\mathcal{L}_1$ and $\mathcal{L}_2$ are defined as follows.

1. The set $\mathcal{L}$ consists of the *logical symbols* $\forall, \exists, \neg, \wedge, \vee, \to, =, \leftrightarrow$ together with all kinds of brackets.

2. The language $\mathcal{L}_1$ of first order set theory is defined as $\mathcal{L} \cup \{v_i \mid i \in \mathbb{N}\} \cup \{\in\}$. The symbols v_i are called *set variables.*

3. The language $\mathcal{L}_2$ of second order set theory is defined as $\mathcal{L}_1 \cup \{V_i \mid i \in \mathbb{N}\}$. The symbols V_i are called *class variables.*

For $i = 1$ and $i = 2$, we will write $\mathcal{L}_i^*$ to mean the set of all finite strings with symbols in $\mathcal{L}_i$.

Definition 25. The collections of $\mathcal{L}_1$ and $\mathcal{L}_2$ *formulas* is defined inductively as follows.

1. The collection $\mathcal{F}_1$ of $\mathcal{L}_1$ formulas is the least subset of $\mathcal{L}_1^*$ that contains all expressions of the form $x = y$ and $x \in y$ for any set variables x, y and further has the following closure properties:

   a) If φ and ψ are $\mathcal{L}_1$ formulas, then so are the expressions $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$, $(\varphi \leftrightarrow \psi)$ and $\neg(\varphi)$.

   b) If φ is an $\mathcal{L}_1$ formula that does not contain any of the two substrings "$\exists x$" or "$\forall x$" respectively, then the expressions $\forall x\,(\varphi)$ and $\exists x\,(\varphi)$ are $\mathcal{L}_1$ formulas as well.

2. The collection $\mathcal{F}_2$ of $\mathcal{L}_2$ formulas is the least subset of $\mathcal{L}_2^*$ that contains all $\mathcal{L}_1$ formulas and all expressions of the form $x \in X, x = X$ and $X = Y$ for set variables x and class variables X and Y, and further satisfies the following closure properties:
   a) If φ and ψ are $\mathcal{L}_2$ formulas, then so are the expressions $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$, $(\varphi \leftrightarrow \psi)$ and $\neg(\varphi)$.
   b) If φ is a $\mathcal{L}_2$ formula that does not contain any of the two substrings "$\exists x$" or "$\forall x$" respectively, then the expressions $\forall x\,(\varphi)$ and $\exists x\,(\varphi)$ are $\mathcal{L}_2$ formulas as well.
   c) If φ is a $\mathcal{L}_2$ formula that does not contain any of the two substrings "$\exists X$" or "$\forall X$" respectively, then the expressions $\forall X\,(\varphi)$ and $\exists X\,(\varphi)$ are $\mathcal{L}_2$ formulas as well.

The *scope* of a quantifier in a formula φ is that subformula of φ that is confined by the first pair of brackets that opens after the quantifier under consideration. In a formula φ, a variable x is bound by the quantifier $\exists x$ ($\forall x$) if x is in the scope of that quantifier. Formulas that do not contain quantified second order variables at all are called *elementary* formulas. The *free set- and class-variables* of a formula are defined inductively as follows:

1. In quantifier free formulas, all occurring set-variables are free set-variables and all occurring class-variables are free class-variables.

2. The free set-(class-)variables of any formula of the form $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$ or $(\varphi \leftrightarrow \psi)$ are exactly those variables that are free (class-) set-variables of φ or of ψ.

3. The free (class-) set-variables of $\neg\varphi$ are those of φ.

4. The free set-variables of a formula of the form $\exists x\,\varphi$ and $\forall x\,\varphi$ are those free set-variables of φ that differ from x. The free class-variables of $\exists x\,\varphi$ and $\forall x\,\varphi$ are those of φ.

5. The free class-variables of a formula of the form $\exists X\,\varphi$ or $\forall X\,\varphi$ are those free class-variables of φ that differ from X. The free set-variables of $\exists X\,\varphi$ and $\forall X\,\varphi$ are those of φ.

Formulas with no free variables are called *closed* formulas. Let φ be a formula and let X, Y and x, y be free class- and set-variables respectively. The expression $(\varphi(\frac{Y}{X}))$ $\varphi(\frac{y}{x})$ stands for the result of simultaneously replacing all occurrences of (X) x in φ by (Y) y.

Remark 26. Whenever readability can be improved without obfuscating meaning, we will simplify sequences of variables by the vector notation, i.e. we will write $\varphi(\vec{X}, \vec{x})$ instead of $\varphi(X_1, \ldots, X_n, x_1, \ldots, x_k)$ and similarly, for a quantifier Q, $Q\vec{X}$ and $Q\vec{x}$ instead of $QX_1, \ldots X_n$ and $Qx_1, \ldots, x_k$.

Definition 27. We use a Hilbert style axiomatization of first order logic with equality as it is presented in [Hal11], I.3, and outlined in [Kun80], Chapter I §2 and §4. We use the symbol $\vdash$ to denote formal derivability. We use the term *theory* interchangeably with collection of formulas.

Definition 28. Let T be any theory. The classes $\Delta^k_n(T)$, Σ^k_n and Π^k_n for all natural number n and $k \in \{0, 1\}$ are inductively defined as follows.

1. The collections $\Pi^0_0 = \Sigma^0_0 = \Delta^0_0$ consist of all quantifier free formulas. form $x \in Y$, $x \in y$, $x = y$ and $X = Y$. The collections $\Pi^1_0 = \Sigma^1_0 = \Delta^1_0$ consist of all elementary formulas.

2. The formulas in Σ^0_{n+1} are those formulas that are of the form $\exists x\, \psi(x)$ with $\psi \in \Pi^0_n$. The formulas in Σ^1_{n+1} are those formulas that are of the form $\exists X\, \psi(X)$ with $\psi \in \Pi^1_n$.

3. The formulas in Π^0_{n+1} are those formulas that are of the form $\forall x\, \psi(x)$ with $\psi \in \Sigma^0_n$. The formulas in Π^1_{n+1} are those formulas that are of the form $\forall X\, \psi(X)$ with $\psi \in \Sigma^1_n$.

4. A formula φ is an element of $\Delta^k_n(T)$ if there are formulas $\pi \in \Pi^k_n$ and $\sigma \in \Sigma^k_n$ such that $T \vdash \varphi \leftrightarrow \sigma \wedge \varphi \leftrightarrow \pi$. We write Δ^k_n to mean $\Delta^k_n(\mathsf{NBG})$.

Definition 29. A *structure* for the language $\mathcal{L}_2$ is a triple $\mathfrak{M} = (M, S, \epsilon)$ that consists of a set M, a set S of subsets of M and a binary relation ϵ, i.e. a set of pairs (x, y) where $x \in M$ and $y \in M \cup S$. We will write (M, S) to mean the structure $(M, S, \in)$. The satisfaction relation for $\mathcal{L}_2$ formulas is defined as usual by interpreting $\in$ as ϵ. A *model* of a set $\mathcal{F}$ of $\mathcal{L}_2$ formulas is a structure for the language $\mathcal{L}_2$ that satisfies all formulas in $\mathcal{F}$.

1.4 The theories ZFC, NBG and MK

The axioms of ZFC

In the literature, the term NBG does not always refer to exactly the same theory. Aside from slightly different formulations of the standard axioms, a noteworthy difference in the expositions lies in whether the axiom of choice is used in its local (set theoretic) form or as the existence of a global choice function. However, as it is presented in [Fel71], the resulting theories prove the same set theoretic formulas.

Although we assume some basic familiarity with naive set theory and corresponding notations, we want to make some abbreviations explicit to avoid any misinterpretation.

Definition 30. Let X, F and Y be classes, let x, y be sets and let $\varphi(x)$ be any formula.

1. We write $\exists! x\, \varphi(x)$ to mean
$$\exists x\, \varphi(x) \wedge \forall x, y\, (\varphi(x) \wedge \varphi(y) \to x = y).$$

2. The term $\langle x, y\rangle$ stands for the *Kuratowski pair* $\{\{x\}, \{x, y\}\}$ of x and y. This shorthand notation is justified from the axioms that are presented in the following, in particular, the Kuratowski pair of any two sets always exists and is unique.

3. The expression $\forall x \in R\, \exists y, z\big(x = \langle y, z\rangle\big)$ is abbreviated by $\mathsf{rel}(R)$. If $\mathsf{rel}(R)$ holds of some class R, we say the R is a *(binary) relation.*

4. The term $\mathsf{fun}(F)$ stands for
$$\mathsf{rel}(F) \wedge \forall x, y, z\, \big(\langle x, y\rangle \in F \wedge \langle x, z\rangle \in F \to y = z\big).$$

If $\mathsf{fun}(F)$ holds of some class F, then we say that F is a *function.*

5. The term $F : X \to Y$ is meant to abbreviate

$$\mathsf{fun}(F) \wedge \forall x \in X \, \exists y \in Y \, (\langle x, y \rangle \in F) \wedge \forall x, y \, (\langle x, y \rangle \in F \to x \in X).$$

If $F : X \to Y$ holds and x is some element of X, then we write $F(x)$ to mean the unique element y of Y that satisfies $\langle x, y \rangle \in F$. If $F : X \to Y$ holds and in addition $\forall x, y \in X \, (x \neq y \to F(x) \neq F(y))$ holds, then F is an *injective* or *one-to-one* function. The function $F : X \to Y$ is said to be *onto* Y if $\forall y \in Y \, \exists x \in X \, (F(x) = y)$. An injective function $F : X \to Y$ that maps X onto Y is called a *bijective* function of X to Y.

6. The expression $|X| = |Y|$ stands for the formula

$$\exists F \big(F : X \to Y \wedge \forall y \in Y \, \exists ! x \in X \, (\langle x, y \rangle \in F) \big).$$

That is, the expression $|X| = |Y|$ stands for the statement that a bijective function from X to Y exists.

Now we are ready to present the axioms of ZFC set theory. These principles represent our basic *set theoretic* axioms, and they will hold in every theory under consideration. Note that the collection of axioms that we use to axiomatize ZFC (and later NBG) is redundant in the sense that there are axioms which could be derived from the other axioms. For example, we will introduce the so called axiom of separation, a formal justification for the use of terms in the style of $\{x \in y \mid \varphi(x)\}$. The reason to keep these axioms in our formulation of ZFC is twofold: first, it is an axiomatization that is often used in the literature (cf. [Jec03] and [Kun80]) and second, principles such as the axiom of separation reflect characteristic properties of how we think of sets; as such they qualify as axioms.

The axiom of *extensionality*, is the statement that sets with the same extension are equal. Formally, it is

$$\forall x, y \, (x = y \leftrightarrow \forall z \, (z \in x \leftrightarrow z \in y)). \tag{EXT}$$

In particular, the axiom of extensionality implies that there are no two distinct sets with the same members. We will use this fact constantly, and without explicitly quoting the axiom of extensionality, to name sets that are given from a collection of elements whose existence is guaranteed by some further axioms. For example, we will use the symbol $\varnothing$ to mean the unique set $\{x \mid x \neq x\}$, whose existence is guaranteed by the *axiom of the empty set*

$$\exists x \, \forall y \, (y \notin x). \tag{NUL}$$

The *axiom of pairs* is the statement for any two sets x and y there is also a set $\{x, y\}$ that contains x and y and nothing else. Formally, this is the statement

$$\forall x, y \, \exists z \, \forall w \, (w \in z \leftrightarrow w = x \vee w = y). \tag{PAIR}$$

The axiom PAIR, thus guarantees that for all sets x and y, the aforementioned (Kuratowski-) pair $\langle x, y\rangle = \{\{x\}, \{x, y\}\} = \{\{x, x\}, \{x, y\}\}$ exists.

The *axiom of regularity*,

$$\forall x \, \big(\exists y \, (y \in x) \to \exists y \in x \, \forall z \in x \, (z \notin y)\big), \tag{REG}$$

is the statement that the elementhood relation is well founded on sets.

The *axiom of union*, denoted by U, is the statement that for every set x there is a set y that contains exactly the sets that are an element of some element of x. Formally, corresponds to

$$\forall x \, \exists y \, \forall z \, (z \in y \leftrightarrow \exists u \in x \, (z \in u)). \tag{U}$$

In the context of any theory that contains EXT and U, and thus in every theory under consideration of this work, the above set y exists and is uniquely determined from x. Hence it is justified to write $\cup x$ to mean that specific set y. Similarly, we will write $x \cup y$ and $\bigcup_{i \in I} x_i$ to denote the sets $\cup\{x, y\}$ and $\cup\{x_i \mid i \in I\}$ respectively.

Definition 31. A class W is *inductive* if $\forall x \, (x \in W \to x \cup \{x\} \in W)$ and $\varnothing \in W$.

The *axiom of infinity* is the assertion that inductive sets exist.

$$\exists w \left(\forall x \left(\forall y \left(y \notin x \right) \to x \in w \right) \right. \\ \left. \wedge \forall z \left(z \in w \to \exists a \in w \left(\forall b \left(b \in a \leftrightarrow b = z \vee b \in z \right) \right) \right) \right). \tag{INF}$$

However, in the context of theories that comprise the axioms $\mathsf{PAIR}, \mathsf{U}$ and NUL the above formula can be expressed more naturally as

$$\exists w \left(\varnothing \in w \wedge \forall x \in w \left(x \cup \{x\} \in w \right) \right).$$

The *axiom of the power set* is the statement that for any set x there exists a set that contains exactly the subsets of x as its members. As a formula, this is

$$\forall x \, \exists y \, \forall z \left(z \in y \leftrightarrow z \subset x \right). \tag{POW}$$

In any context where EXT and POW are available, the above set y exists and is uniquely determined from x. We shall call the set y the *power set* of x and denote it by $\mathcal{P}(x)$.

Given any set x and any $\mathcal{L}_1$ formula $\varphi(z)$, the *axiom of separation* guarantees the existence of the set $\{z \in x \mid \varphi(z)\}$. Hence, the axiom of separation is (the universal closure of) the schema

$$\forall x \, \exists y \, \forall z \left(z \in y \leftrightarrow z \in x \wedge \varphi(z, \vec{p}) \right), \tag{SEP}$$

where φ is meant to range over $\mathcal{L}_1$ formulas with free variables different from y and x.

The *axiom of replacement* is (the universal closure of) the schema

$$\forall u \left(\forall x \in u \, \exists! y \, \varphi(x, y, \vec{p}) \;\to\; \exists v \, \forall x \in u \exists y \in v \, \varphi(x, y, \vec{p}) \right) \tag{REP}$$

where $\varphi(x, y, \vec{p})$ is meant to range over $\mathcal{L}_1$ formulas with free variables different from y. In particular, if f is a function and a set, then the axiom of replacement guarantees for every set x that the class $f[x]$ is a set as well. We will see in Lemma 42 that the restriction on f to be a set can be dropped.

The *axiom of choice* is the statement that for each set x of nonempty sets, there exists a choice function. As a formula, this is

$$\forall x \left(\varnothing \notin x \to \exists f\, (f : x \to \cup x \wedge \forall y \in x\, (f(y) \in y)\right). \tag{C}$$

The axioms of NBG and MK

The theory NBG, Von Neumann Bernays Gödel set theory, consists of all axioms of ZFC together with the following axioms for classes. This theory will serve as our "base theory" in the sense that all theories under later investigation will be extensions thereof. As opposed to NBG, the theory MK, Morse Kelley set theory, is strong enough to include (almost) all the theories under consideration.

Our first step in formulating NBG set theory is to restate the axiom of extensionality so that also classes with the same extension are treated as one object. Formally, this is

$$\forall \mathcal{X}, Y\, (\mathcal{X} = Y \leftrightarrow \forall z\, (z \in \mathcal{X} \leftrightarrow z \in Y)), \tag{EXTC}$$

where $\mathcal{X}$ stands for either X or x respectively. We need the variable $\mathcal{X}$ as a class variable and as a set variable so that we can prove that if a class and a set are extensionally equal, then they denote the same object. This is necessary to prove that every set is also a class (cf. Fact 36). The *axiom of elementary comprehension* guarantees that for every elementary formula $\varphi(x)$ there exists a class that contains exactly those sets x that satisfy $\varphi(x)$. It is the universal closure of

$$\exists X\, \forall y\, (y \in X \leftrightarrow \varphi(y, \vec{p})), \tag{ECA}$$

where $\varphi(x, \vec{p})$ is meant to range over all elementary formulas with free variables different from X. Note that elementary comprehension gives rise to the class $V = \{x \mid x = x\}$ of all sets. Also note that elementary comprehension and extensional equality of classes allow us to extend the powerset operator to form powerclasses in the sense that $\mathcal{P}(X)$ stands for the class $\{x \mid x \subset X\}$. The powerset axiom from ZFC can now be expressed by requiring powerclasses of sets to be sets again. Furthermore, we can use the axiom ECA to form the *cartesian product* $X \times Y = \{\langle x, y\rangle \mid x \in X \wedge y \in Y\}$ for arbitrary classes X and Y.

The last axiom of NBG is the aforementioned *limitation of size axiom.* It is the formal version of our paradigm that proper classes are exactly those collection of sets whose extension is larger than any set. The formal statement is

$$\forall X \big(\forall x\,(x \neq X) \leftrightarrow |X| = |V|\big). \tag{LIM}$$

Finally, Morse Kelley set theory is obtained from NBG by replacing the axiom ECA by full comprehension, that is, in the schema of elementary comprehension, the elementarity condition on the formula φ is dropped.

Meta mathematical aspects of $\mathsf{NBG}, \mathsf{ZFC}$ and MK

During the development of the foundations of mathematics, the meta mathematics of theories such as $\mathsf{ZFC}, \mathsf{NBG}$ and MK has been thoroughly studied. It goes far beyond the scope of this work to list all the insights that where made in that respect. It is our aim here, however, to just mention very briefly the sort of fundamental results that substantiate our viewpoint that PA relates to ACA_0 similarly as ZFC relates to NBG, and that MK set theory is the counterpart of full second order arithmetic Π^1_∞-CA_0. If we informally use the fraction line to express the "relates to" relation, we can summarize this point of view as

$$\frac{\mathsf{PA}}{\mathsf{ACA}_0} = \frac{\mathsf{ZFC}}{\mathsf{NBG}}$$

and in a wider sense also

$$\frac{\mathsf{PA}, \mathsf{ACA}_0}{\Pi^1_\infty\text{-}\mathsf{CA}_0} = \frac{\mathsf{ZFC}, \mathsf{NBG}}{\mathsf{MK}}.$$

In arithmetic, the first order base theory PA has infinitely many axioms[3] and has a conservative (cf. Theorem IX.1.5 and Remark IX.1.7 in [Sim98]) second order extension ACA_0, that in turn is finitely axiomatizable (cf. Lemma VIII.1.5 in [Sim98]).

[3] The instances of the schema of mathematical induction.

The situation is analog in set theory; while ZFC cannot be finitely axiomatized (cf. Corollary 7.7 in [Kun80] and the explanation and references given in Fact 32 below), there are finitely many instances of ECA from which all other instances are derivable (a finite axiomatization can be found in [Men97]).

Fact 32. The theory ZFC has no finite axiomatization.

Proof. This is a direct consequence of the reflection principle (as provable in ZFC) and of Gödels second incompleteness theorem. For more details see also Theorem I.12.14 and the subsequent discussion in [Jec03]. □

Fact 33. The theories NBG and ACA_0 can be finitely axiomatized.

Proof. For a detailed exposition of a finite axiomatization of NBG, the reader is referred to 4.1. in [Men97][4]. To see that ACA_0 is finitely axiomatizable, the reader is referred to Lemma VIII.1.5 in [Sim98]. □

A second similarity between the relationship of PA to ACA_0 on one side and ZFC to NBG on the other side, is that the second order theories are conservative extensions of the respective first order theories.

Fact 34. For all formulas $\varphi \in \mathcal{L}_1$, we have that

$$\mathsf{ZFC} \vdash \varphi \Leftrightarrow \mathsf{NBG} \vdash \varphi.$$

For all formulas $\psi \in \mathcal{L}_A^1$, we have that

$$\mathsf{PA} \vdash \psi \Leftrightarrow \mathsf{ACA}_0 \vdash \psi,$$

where PA stands for Peano arithmetic.

Proof. The first part is a direct consequence of Theorem 1 in [Fel71]. For the second part, the reader is referred to Theorem IX.1.5 and Remark IX.1.7 in [Sim98]. □

[4] To obtain exactly the same theory as introduced here, one has to keep our axiom of limitation of size or alternatively add an axiom of global choice to Mendelson's variant of NBG.

Since the theory MK is neither finitely axiomatizable nor is it conservative over NBG and ZFC, the analogy is also preserved if we consider to extend the second order theories by full comprehension schemata.

Basic observations and definitions

We will now list some of the most basic facts about NBG and introduce the basic concepts on which the remainder of this text relies. Many of the proofs are standard and can be found in the literature. However, in many cases we found it more cumbersome to point out appropriate sources instead of giving an actual proof ourselves. In these cases, we just wrote down the proofs, but of course we do not claim any of the results. In fact, we assume that all statements in this section are "folklore".

Our first observation is that in NBG, the terminus of proper classes[5] is not void.

Fact 35. Not every class is a set. Classes that are not sets are called *proper classes.*

Proof. If V was a set, then also $\{V\}$ would be set. As the existence of such a set contradicts the axiom of regularity, this is not possible. $\square$

Fact 36. Every set is a class.

Proof. Let x be any set. The axiom of elementary comprehension yields a class $X = \{z \mid z \in x\}$ with the same extension as x. It follows from extensional equality of classes that $X = x$. $\square$

Fact 37. Every subclass of a set is a set.

Proof. This is immediate from the limitation of size axiom. $\square$

Fact 38. There are no cycles in the elementhood relation, i.e. there are no sets $x_0, x_1, \ldots, x_n$ such that $x_0 \in x_1 \in \cdots \in x_n \in x_0$.

[5]Proper classes are classes that are not sets.

Proof. Let $x_0, \dots, x_n$ be sets. By the axioms of union and pairing, the collection $\{x_0, \dots, x_n\}$ is a set. Applying the axiom of regularity, we conclude that there is a set $z \in \{x_0, \dots, x_n\}$ with $z \cap \{x_0, \dots, x_n\} = \varnothing$. Hence, z cannot be part of the proposed cycle. □

Definition 39. A class X is *transitive* if $\forall y\,(y \in X \to y \subset X)$ holds. We denote the transitivity of a class X by $\mathsf{tr}(X)$. The class X is *swelled* whenever $\forall x\,(x \in X \to \mathcal{P}(x) \subset X)$ holds. If X is both, swelled and transitive, then it is called *supertransitive*.

Fact 40. V is supertransitive.

Proof. This is immediate from the definition of V. □

Definition 41.
1. The *restriction* $X \restriction Y$ of a class X to a class Y is the collection $\{\langle x, y\rangle \in X \mid x \in Y\}$.

2. Let F be a function. The class

$$dom(F) = \{x \mid \exists y\,(\langle x, y\rangle \in F)\}$$

is called the *domain* of F. The *image* of a class X under the function F is defined from $F[X] = \{F(x) \mid x \in X \cap dom(F)\}$. The *range* of F is the class $rng(F) = F[dom(F)]$. The *preimage* of a class X under F is the class $F''X = \{y \mid F(y) \in X\}$. For any set x, we write $F'x$ to mean the *fiber* $F''\{x\}$ over x.

3. For any two functions F and G with $rng(F) \subset dom(G)$, we define the *composition* $G \circ F : dom(F) \to rng(G)$ from $G \circ F(x) = G(F(x))$.

Lemma 42 (Replacement principle)**.** *If F is a function and u is a set, then also $F[u]$ is a set.*

Proof. Let F be any function and let u be a set. From instantiating the axiom REP with the $\mathcal{L}_1$ formula $\varphi(x, y, f) \equiv \langle x, y\rangle \in f$, we get that $f[u]$ is a set for every function f. Since $F[u] = (F \restriction u)[u]$, it is enough to show that $F \restriction u$ is a set. This, however follows from LIM since $F \restriction u$ can be put into one-to-one correspondence with u. □

Definition 43. A pair $(A, <_A)$ of two classes A and $<_A$ where $<_A$ is a binary relation on A is a *transitive relation* if $(a <_A b \wedge b <_A c \rightarrow a <_A c)$ holds for all $a, b, c \in A$. We will write $\mathsf{tran}(A, <_A)$ to mean that $(A, <_A)$ is a transitive relation. A *linear order* is a transitive relation that additionally fulfills the following[6] requirements for all $a, b \in A$:

1. $a <_A b \vee b <_A a \vee a = b$
2. $(a <_A b \rightarrow \neg(b <_A a))$.

We will write $\mathsf{lo}(A, <_A)$ to mean that $(A, <)$ is a linear order.

Remark 44. If $(A, <_A)$ is a linear order, then we say that $<_A$ is a linear ordering of A. If we just say that $<_A$ is a linear ordering, then we mean to express that $(\{x \mid \exists y\,(x <_A y \vee y <_A x)\}, <_A)$ is a linear order. Moreover, if $(A, <_A)$ is a linear order and $X \subset A$, then we mean $(X, (X \times X) \cap <_A)$ when we write $(X, <_A)$.

Definition 45. We introduce the following shorthand notations for linear orders $(A, <_A)$ and arbitrary subclasses X and Y of A.

1. For any elements a and b of A, we write $a \leq_A b$ to mean $a <_A b \vee a = b$.
2. If $a \in A$, then:
   a) The term $A_{<_A a}$ denotes the class $\{b \in A \mid b <_A a\}$.
   b) The expression $a <_A X$ is a shorthand notation for the statement $\forall x \in X\,(a <_A x)$.
   c) The class $A_{<_A X}$ is defined as $\{b \in A \mid b <_A X\} = \bigcap_{b \in X} A_{<_A b}$. Similar notions like $A_{\leq_A X}$ or $A_{_A > X}$ are interpreted accordingly.
   d) We also use $\overline{X}$ to mean the relative complement $A \setminus X$ of X to the field of the relation at hand.
3. The class X is an *initial* subclass of Y (relative to $(A, <_A)$), in symbols $X \prec Y$, if

$$X \subset Y \wedge \forall a, b \in Y\,\big(a <_A b \wedge b \in X \rightarrow a \in X\big)$$

   holds, it is called *coinitial* in Y if $Y \setminus X \prec Y$.

[6]From here on out, we write $x <_A y$ to mean $\langle x, y\rangle \in <_A$.

4. The class X is called *minimal* in $Y \subset A$ if

$$X \subset Y \wedge \forall a \in Y\, \exists b \in X\, (b \leq_A a)$$

holds.

Definition 46. If R is any binary relation on a class A and $X \subset A$ is a subclass of A, then we say that $a \in X$ is *a minimal element* in X (with respect to R) if

$$\forall b \in X\, \neg(bRa)$$

holds.

Remark 47. Note that for linear orderings $(A, <_A)$, we have that any $a \in A$ is exactly then a minimal element of some class $X \subset A$ if the set $\{a\}$ is a minimal subset of X. Further, we have that $\varnothing$ is minimal in some $X \subset A$ if and only if $X = \varnothing$, and similarly, that $A_{<_A \varnothing} = A$.

Definition 48. A set x is called an *ordinal* if it is transitive and all its members are also transitive. Formally, we write

$$\mathsf{ord}(x) \;\equiv\; \forall y\, (y \in x \cup \{x\} \to \mathsf{tr}(y)).$$

We will use lower case Greek letters to range over ordinals. Whenever we write $\alpha < \beta$ for two ordinals α and β, we mean, unless otherwise stated, that $\alpha \in \beta$. We write On to denote the class of all ordinals.

Remark 49. Note that a set is exactly then an ordinal if it is transitive and all its elements are ordinals. That is, $\mathsf{ord}(x) \leftrightarrow \mathsf{tr}(x) \wedge \forall y \in x\, (\mathsf{ord}(y))$ holds. In particular, On is a transitive class of transitive sets.

Proof. Let α be any ordinal; we prove that every element of α is an ordinal as well. If $x \in \alpha$, then x is transitive because so are all the elements of an ordinal. For any $y \in x$ we have, by transitivity of α, that y is an element of α and thus is also transitive. On the other hand, it is obvious that every transitive set of ordinals is an ordinal itself. $\square$

Remark 50. Intersections and unions of transitive sets are transitive. Hence, we get as a consequence of the previous remark and the limitation of size axiom, that the intersection of any nonempty class or the union of any set of ordinals is itself an ordinal.

Lemma 51. *Every nonempty class X of ordinals contains minimal elements with respect to the elementhood relation.*

Proof. Applying the axiom REG to a nonempty set x of ordinals yields

$$\exists \mu \in x \, \forall \alpha < \mu \, (\alpha \notin x).$$

In other words, every nonempty set of ordinals has minimal elements. Now let α be an element of any class X of ordinals. If $\alpha \cap X = \varnothing$, then α is a minimal element of X and we are done. Assume that $x = X \cap \alpha \neq \varnothing$. Since x is a nonempty set of ordinals, it follows from our previous observation that x must contain a minimal element μ. To see that μ is also minimal in X, assume that $\beta < \mu$ for some ordinal β. Since $\mu \in \alpha$ and α is transitive, this implies that $\beta \in \alpha$. Thus, by the minimality of μ in x it follows that $\beta \notin x$, the latter implies that β cannot be an element of X. □

Remark 52. The empty set, as an ordinal also called 0, is the unique minimal element of On.

Corollary 53 (Induction schema for On)**.** *For every elementary formula $\varphi(x)$ with a distinct free variable x it is provable in* NBG *that the following holds:*

$$\forall \alpha \left(\forall \beta < \alpha \, \varphi(\beta) \rightarrow \varphi(\alpha) \right) \rightarrow \forall \alpha \, \varphi(\alpha).$$

Proof. As φ is an elementary formula, we can use comprehension to form the class $X = \{\alpha \mid \neg\varphi(\alpha)\}$. If φ satisfies the precondition of the induction schema, then X cannot have any minimal element and thus must be empty by Lemma 51. □

Corollary 54 (Class induction for On)**.** *It is provable in* NBG *that*

$$\forall X \left(\forall \alpha \, (\alpha \subset X) \rightarrow \alpha \in X) \rightarrow On \subset X \right)$$

holds.

Proof. This corresponds to the instance of the induction schema for On where $\varphi(x)$ is $x \in X$. $\square$

Lemma 55. *The class On is linearly ordered by the elementhood relation.*

Proof. Since we have the axiom of regularity and since all ordinals are transitive, it is enough to prove that the elementhood relation is total on On. That is, we have to show that the class

$$X = \{\alpha \mid \exists \beta \, (\alpha \notin \beta \wedge \beta \notin \alpha \wedge \alpha \neq \beta)\}$$

is the empty set. By way of contradiction, assume that $X \neq \varnothing$. Let α_0 be a minimal element of X and let β_0 be a minimal element of the class $\{\beta \mid \alpha_0 \notin \beta \wedge \beta \notin \alpha_0 \wedge \alpha_0 \neq \beta\}$. If μ is any element of α_0, then by the minimality of α_0 in X, we have that $\mu \in \beta_0 \vee \beta_0 \in \mu \vee \beta_0 = \mu$. Since the latter two cases cannot hold, it is clear that $\mu \in \beta_0$. Hence, $\alpha_0 \subset \beta_0$. If, on the other hand, μ is an element of β_0, then it follows from the minimality of β_0 that either $\mu \in \alpha_0$ or $\mu = \alpha_0$ or $\alpha_0 \in \mu$. The last two cases are not possible since they would imply that $\alpha_0 \in \beta_0$, contradicting our choice of β_0. Hence, we have that $\mu \in \alpha_0$ and therefore that $\beta_0 = \alpha_0$. This is in contradiction to our choice of β_0. $\square$

Corollary 56. *If α and β are ordinals, then either $\alpha \subset \beta$ or $\beta \subset \alpha$ holds.*

Corollary 57. *For ordinals α, β, we have the equivalence*

$$\alpha \subsetneq \beta \Leftrightarrow \alpha \in \beta.$$

Corollary 58. *Every nonempty class of ordinals has a unique least element with respect to the natural ordering.*

Proof. This is a direct consequence of Lemma 51 and Lemma 55. $\square$

Definition 59. Let X be a nonempty class of ordinals; we use the term $\min(X)$ to denote the unique *least element* of X with respect to the natural orderings on the ordinals.

Definition 60. For any set x, the set $\{x\} \cup x$ is the *successor* of x. We write $\alpha + 1$ to mean the successor of an ordinal α. An ordinal is called a *successor ordinal* if it is the successor of some ordinal and *limit ordinal* otherwise. We write $\mathsf{lim}(\alpha)$ to mean that α is a limit ordinal.

Lemma 61. *For any ordinal α, the successor $\alpha + 1$ is an ordinal as well.*

Proof. We have to show that $\alpha + 1$ is transitive. Let $\beta \in \alpha + 1$. It is obviously the case that $\alpha \subset \alpha + 1$, thus the case where $\beta = \alpha$ is done. If $\alpha \neq \beta$, then the transitivity of α implies that $\beta \subset \alpha \subset \alpha + 1$. □

Corollary 62. *The class On is not a set.*

Proof. If *On* was a set, it would be a transitive set of ordinals and thus be itself an ordinal; therefore, it would be the case that $On \in On$, which contradicts the fact that the elementhood relation is cycle free. □

Corollary 63. *For any set x, there are ordinals λ that are greater than any ordinal that is a member of x. Particularly, every set of ordinals has a supremum given from* $\sup(x) = \min\{\alpha \mid \forall \beta \in x\,(\alpha > \beta)\}$.

Proof. Let x be any set; we know from the axiom of union that $y = \cup\{\alpha \mid \alpha \in x\}$ is a set, as such it must be an ordinal. The ordinal $y + 1$ is a strict upper bound for $x \cap On$. □

Lemma 64. *A nonzero ordinal is a limit ordinal if and only if it is inductive.*

Proof. Recall that a class X with $0 \in X$ is inductive if it satisfies the formula

$$\forall x\,(x \in X \to x \cup \{x\} \in X).$$

Let $0 \neq \lambda$ be an ordinal. Obviously, if λ is a successor ordinal, then it is not inductive. The fact that nonzero limit ordinals are inductive follows from the observation that for any ordinal α, the axiom of regularity implies that the class $\{\mu \mid \alpha < \mu < \alpha + 1\}$ is empty. □

Corollary 65. *Limit ordinals are exactly those ordinals that equal their union.*

Definition 66. A class X is *closed under chain unions* if

$$\forall x \subset X\,(\forall y, z \in x\,((z \subset y) \vee (y \subset z)) \to \cup x \in X)$$

is satisfied. An inductive class X that is closed under chain unions is called *superinductive.*

Lemma 67. *A set x is an ordinal if and only if it is an element of every superinductive class. That is, the class On of all ordinals is the intersection of all superinductive classes.*

Proof. Assume that X is some superinductive class which does not contain all of On. Since the intersection of ordinals is always an ordinal, it follows that $\alpha_0 = \cap\{\mu \mid \mu \notin X\}$ is an ordinal. If $\alpha_0 = \beta + 1$ is a successor ordinal, then, since β is a proper subset of α_0, β cannot be an element of the class $\{\mu \mid \mu \notin X\}$. Hence, we have that $\beta \in X$. This is a contradiction to the assumption that X is inductive. Thus, α_0 must be a limit ordinal. Since all elements of α_0 are members of X and $\alpha_0 = \bigcup_{\mu<\alpha_0} \mu$, we have a contradiction to the assumption that X is closed under chain unions. Since On is clearly inductive, it remains to be shown that On is closed under chain unions. Let x be any set of ordinals. We have to verify that $\cup x$ is an ordinal. If $a \in \cup x$, then there is a set y such that $a \in y \in x$ and thus, by transitivity of y, we have that $a \subset y \subset \cup x$. Hence, $\cup x$ is a transitive set of ordinals and thus is an ordinal itself. □

Note that the last lemma corresponds to the definition of ordinals as it is presented in [SF10]. We might refer to that work in subsequent proofs about ordinals.

Definition 68. Let X be any class and let y be a set. We define the y-th section of X as follows:

$$(X)_y = \{x \mid \langle y, x\rangle \in X\}.$$

Theorem 69. *It is provable from* NBG *that*

$$\forall Z\, \exists F\, \big(\mathsf{fun}(F) \wedge \forall x\,(((Z)_x = \varnothing \wedge F(x) = \varnothing) \vee F(x) \in (Z)_x)\big)$$

holds.

Proof. We have noticed before that the class of all ordinals is transitive; therefore, if it was a set, it would contain itself as an element and thus contradict the axiom of regularity (cf. Fact 38). Hence, we can apply the limitation of size axiom to obtain a bijective function $H : V \to On$. Now we can use Corollary 58 and elementary comprehension in order to define F such that

$$\langle x, y\rangle \in F \Leftrightarrow (y = \varnothing \wedge y = (Z)_x) \vee H(y) = \min(H[(Z)_x])$$

holds. □

Corollary 70 (Global choice function)**.** *There exists a class C that satisfies the following property:*

$$C : V \to V \wedge \forall x \big((x = \varnothing \wedge C(x) = x) \vee C(x) \in x\big).$$

Proof. Applying the previous theorem to the class $Z = \{\langle x, y\rangle \mid y \in x\}$ yields a global choice function. □

Corollary 71. *For every elementary formula $\varphi(x, y)$, it is provable in* NBG *that*

$$\exists F \big(F : V \to V \wedge \forall x \big(\exists y\, \varphi(x, y) \to \varphi(x, F(x))\big)\big)$$

holds.

Proof. Given any elementary formula $\varphi(x, y)$, we can apply the previous theorem to the class $Z = \{\langle x, y\rangle \mid \varphi(x, y)\}$ to obtain a suitable function F. □

Corollary 72 (Choice schema)**.** *For every elementary formula $\varphi(x, y)$, it is provable in* NBG *that*

$$\exists F \big(F : V \to V \wedge \forall x\, \exists y\, \varphi(x, y) \to \forall x\, \varphi(x, F(x))\big)$$

holds.

Proof. This is a special case of the previous corollary. □

Theorem 73 (Transfinite recursion)**.** *Let $G : V \to V$ be a function. There exists a unique function $F : On \to V$ that satisfies the following equation for all ordinals α:*

$$F(\alpha) = G(F \restriction \alpha)$$

Proof. For the uniqueness, assume that there are functions F and H that both satisfy the above equation of F. Let $X = \{\alpha \mid F(\alpha) \neq H(\alpha)\}$. Since X cannot have a least element, it follows from Corollary 58 that X must be empty and thus F and H must take the same values on all ordinals. For the existence, let

$$P(f, \alpha) \equiv f : \alpha \to V \wedge \forall \mu < \alpha \, (f(\mu) = G(f \restriction \mu)).$$

We prove that the class $X = \{\alpha \mid \exists f \, P(f, \alpha)\}$ is superinductive and thus contains all ordinals. Since $P(\varnothing, 0)$ trivially holds, we have that $0 \in X$. If α is in X, then there is a function f with $P(f, \alpha)$. Stipulating $\hat{f} = f \cup \{\langle \alpha, G(f) \rangle\}$, we get that $P(\hat{f}, \alpha + 1)$ and thus that $\alpha + 1 \in X$. Now assume that y is a chain in X. Since the union of any set of ordinals is an ordinal, we can fix $\gamma = \cup y$. Essentially the same argument as in the uniqueness part of this proof yields that $\hat{f} = \cup\{f \mid \exists \alpha \in y \, (P(f, \alpha))\}$ is a function $\hat{f} : \gamma \to V$. If $\alpha < \gamma$, then $\alpha \in \cup y$; hence, there exists a $\beta \in y$ such that $\alpha \in \beta$. Since $y \subset X$, we can pick a function f_β with $P(f_\beta, \beta)$. Now consider

$$\hat{f}(\alpha) = f_\beta(\alpha) = G(f_\beta \restriction \alpha) = G(\hat{f} \restriction \alpha).$$

Now that we know that X is superinductive, we can apply Lemma 67 to conclude that $X = On$. Applying exactly the same reasoning one more time yields that $F = \cup\{f \mid \exists \alpha \, P(f, \alpha)\}$ is a function with the desired properties. □

Corollary 74. *It is provable in* NBG *that for every elementary formula φ,*

$$\exists F \left(F : On \to V \wedge \forall \alpha \left(\exists x \, \varphi(F \restriction \alpha, x) \to \varphi(F \restriction \alpha, F(\alpha))\right)\right).$$

holds.

Proof. From Corollary 71, we get a function $G : V \to V$ with the property that

$$\forall x\,(\exists y\,\varphi(x,y) \to \varphi(x, G(x))).$$

Now we apply transfinite recursion to obtain a function $F : On \to V$ that satisfies the equation $F(\alpha) = G(F \restriction \alpha)$ for every ordinal α. Obviously, F has the required properties. $\square$

Corollary 75 (Dependent choice)**.** *It is provable from* NBG *that for any elementary formula φ, the following holds:*

$$\forall x\,\exists y\,\varphi(x,y) \to \exists F\,\big(F : On \to V \wedge \forall \alpha\,\varphi(F \restriction \alpha, F(\alpha))\big)$$

Proof. This is a special case of the previous corollary. $\square$

Definition 76. We define the set ω as $\bigcap\{x \mid x \text{ is inductive}\}$. The elements of ω are called *natural numbers*. We call a set x *finite* if there is a bijective mapping $f : x \to n$ for some $n \in \omega$. We write $|x| \in \omega$ to mean that a set x is finite.

Remark 77. Note that the axiom of infinity implies that ω is indeed a set. Moreover, since being inductive is preserved under intersections, ω is an inductive set.

Lemma 78. *The set ω is the least nonzero limit ordinal.*

Proof. In view of Lemma 64, it is enough to show that ω is an ordinal. From the axiom of infinity and from Lemma 67, we already know that ω is a set of ordinals. Let α_0 be the least ordinal that is not a subset of ω and let $\beta \in \alpha_0 \backslash \omega$ be arbitrarily chosen. Since $\beta \notin \omega$, we know that $\beta + 1 \not\subset \omega$. Hence, it follows from the minimality of α_0 that $\alpha_0 = \beta + 1$. If we show that $\omega = \beta$, then we are done. The fact that $\beta \subset \omega$ follows immediately from the minimality of α_0. For the converse implication, it is enough to prove that β is inductive. By definition, we have that $\varnothing \in \omega$, thus our choice of β implies that $\beta \neq \varnothing$ hence $\varnothing \in \beta$. If there was an ordinal $\lambda \in \beta$ such that $\beta = \lambda + 1$, then we would already have that $\lambda \in \omega$, and since ω is inductive also that $\beta \in \omega$, contradicting our choice of β. $\square$

Lemma 79 (Mathematical induction)**.** *Let $\varphi(x)$ be any elementary formula with a distinct free variable x. It is provable in* NBG *that the following holds.*

$$\varphi(0) \land \forall n \in \omega\, (\varphi(n) \to \varphi(n+1)) \to \forall n \in \omega\, \varphi(n)$$

Proof. Let $\varphi(x)$ be an elementary formula that satisfies the antecedent of the claim. Let $\psi(\alpha) \equiv \varphi(\alpha) \lor \alpha \notin \omega$. Now assume that $\psi(\beta)$ holds for all ordinals β below some α. If α is a natural number, then it is a successor ordinal or 0, in any case it follows from our assumption that $\varphi(\alpha)$ and thus also $\psi(\alpha)$ holds. In case where α is greater or equal to ω it trivially satisfies $\psi(\alpha)$. Using transfinite induction on the ordinals, we get that $\forall \alpha \psi(\alpha)$ and thus we also have $\forall n \in \omega\, \varphi(n)$. □

Lemma 80 (Recursion along ω)**.** *For every set x and any function $G : V \to V$, there exists a unique function $f : \omega \to V$ that satisfies the following equations*

$$\begin{aligned} f(0) &= x \\ f(n+1) &= G(f(n)). \end{aligned}$$

Proof. This follows directly from Theorem 73 and the fact that 0 is the only limit ordinal below ω. □

Definition 81. Let x be any set. We define the *transitive closure* of x by

$$\mathsf{trcl}(x) = \cap\{y \mid x \subset y \land \mathsf{tr}(y)\}.$$

Lemma 82. *If x is any set, then we have that $\mathsf{trcl}(x) = \bigcup_{n\in\omega} x_n$ where the x_n's are defined by recursion along ω as follows:*

$$\begin{aligned} x_0 &= x \\ x_{n+1} &= (\cup x_n) \cup x_n. \end{aligned}$$

Proof. Let x be any set and $y = \bigcup_{n\in\omega} x_n$ as is defined before. We prove that y is transitive and that every transitive set that contains x also contains y. To see that y is transitive, consider that if z is any element of y, then it is an element of some x_n; therefore it is a subset $\cup x_n$ and thus a subset of x_{n+1}, which is a subset of y. Let now z be any transitive superset of x. We prove by induction on ω that $\forall n \in \omega\,(x_n \subset z)$. By assumption, we already know that $x_0 = x \subset z$. From the transitivity of z, it follows that if $x_n \subset z$, then also $x_{n+1} = (\cup x_n) \cup x_n \subset (\cup z) \cup z = z$. □

Corollary 83. *The transitive closure of any set is a transitive set; particularly, every set is contained in a transitive set.*

Proof. This follows from Lemma 82 and the axiom of union. □

Lemma 84 (Regularity for classes)**.** *The axiom of regularity can be extended to classes i.e. the formula*

$$\forall X\,\big(\exists x\,(x \in X) \to \exists x \in X\,(X \cap x = \varnothing)\big)$$

is provable in NBG*.*

Proof. Let X be any nonempty class. Pick $x \in X$. If $x \cap X = \varnothing$, then we are done. Otherwise, $X \cap \mathsf{trcl}(x)$ is a nonempty set. Applying the axiom of regularity, we can pick a set $y \in X \cap \mathsf{trcl}(x)$ such that $y \cap (X \cap \mathsf{trcl}(x)) = \varnothing$. Since $y \subset \mathsf{trcl}(x)$, this implies that we have found a set $y \in X$ with the property that $y \cap X = \varnothing$. □

Theorem 85 ($\in$ induction for classes)**.** *The following formula is provable in* NBG*.*

$$\forall X\,\big(\forall x\,(x \subset X \to x \in X) \to X = V\big)$$

Proof. Assume that X is a class that satisfies $\forall x\,(x \subset X \to x \in X)$. Let $Y = V \setminus X$. Assuming that $Y \neq \varnothing$, we can find a set $y \in Y$ such that $y \cap Y = \varnothing$, hence $y \subset X$. This contradicts our choice of X. □

Corollary 86 ($\in$ induction schema)**.** *For every elementary formula $\varphi(x)$ with a distinct free variable x, the following is provable in* NBG*:*

$$\forall x\,\big(\forall y \in x\,\varphi(y) \to \varphi(x)\big) \to \forall x\,\varphi(x).$$

Proof. This follows immediately from applying $\in$ induction to $\{x \mid \varphi(x)\}$. □

Definition 87 (Cumulative hierarchy)**.** We apply transfinite recursion to define for each ordinal α a set V_α as follows:

$$\begin{aligned} V_0 &= \varnothing \\ V_{\alpha+1} &= \mathcal{P}(V_\alpha) \\ V_\lambda &= \bigcup_{\alpha<\lambda} V_\alpha \qquad \text{for limit ordinals } \lambda > 0. \end{aligned}$$

Lemma 88. *Every set of the form V_α is transitive.*

Proof. This follows from the fact that transitivity is preserved under union and the powerset operator. □

Lemma 89. *For any two ordinals α and β, we have that $\alpha < \beta \to V_\alpha \in V_\beta$.*

Proof. We prove $\forall \beta < \alpha\,(V_\beta \in V_\alpha)$ by induction on α. If $\alpha = 0$, then the statement is clearly satisfied. Let $\beta < \alpha$ be chosen arbitrarily. If $\alpha = \gamma + 1$ is a successor ordinal, then $V_\gamma \subset V_\gamma$ implies that $V_\gamma \in V_\alpha$ and thus that either $V_\beta = V_\gamma \in V_\alpha$ or we apply the induction hypothesis to γ and therefore obtain that $V_\beta \in V_\gamma \in V_\alpha$. In that case, the claim follows from the transitivity of V_α. If α is a limit ordinal, then we can find some ordinal γ such that $\beta < \gamma < \alpha$. We use the induction hypothesis on γ and the transitivity of V_α to obtain that $V_\beta \in V_\gamma \subset V_\alpha$ as desired. □

Lemma 90. *For every set x, there is an ordinal α such that $x \in V_\alpha$. In other words, $\bigcup_\alpha V_\alpha = V$.*

Proof. We proceed by $\in$ induction. Let x be any set such that for every member y of x, there is an ordinal α_y with $y \in V_{\alpha_y}$. By the replacement axiom, we know that $\{\alpha_y \mid y \in x\}$ is a set of ordinals. Hence, $\alpha_x = \cup\{\alpha_y \mid y \in x\}$ is itself an ordinal. It follows from the monotonicity of the cumulative hierarchy that $x \subset \bigcup_{y\in x} V_{\alpha_y} \subset V_{\alpha_x}$ and thus that $x \in V_{\alpha_x+1}$. □

Definition 91. We define for any set x its *rank* as $rk(x) = \min\{\alpha \mid x \subset V_\alpha\}$.

Corollary 92. *Every set has a rank.*

Lemma 93. *For any two sets x and y, we have that $x \in y \to rk(x) < rk(y)$.*

Proof. Let $x \in y$ be any two sets and let $\alpha = rk(y)$. From $x \in y \subset V_\alpha$, we know that $x \in V_\alpha$, thus there must exist an ordinal $\beta < \alpha$ such that $x \subset V_\beta$. $\square$

Lemma 94. *For any ordinal α, we have that $rk(\alpha) = \alpha$.*

Proof. Clearly, for all α it is the case that $\alpha \subset V_\alpha$ and thus that $rk(\alpha) \leq \alpha$. By way of contradiction, we assume that α is the least ordinal such that $rk(\alpha) < \alpha$. By Lemma 93, we conclude that $rk(rk(\alpha)) < rk(\alpha)$, contradicting the minimality of α. $\square$

Definition 95. Given a linear ordering $(A, <_A)$, we call a function $f : \alpha \to A$ a *descending chain of length* α in $(A, <_A)$ if $\forall \beta, \gamma < \alpha\, (\beta < \gamma \to f(\gamma) < f(\beta))$ holds. A descending chain f with $dom(f) \geq \omega$ is called an *infinitely descending chain.*

Lemma 96. *Let $(A, <_A)$ be a linear order and let $f : \alpha \to A$ be some function. The following two assertions are equivalent:*

1. *f is a descending chain of length α.*
2. *For all ordinals μ with $\mu + 1 < \alpha$ and for all limit ordinals $\lambda < \alpha$, the following two assertions hold:*
   a) $f(\mu + 1) <_A f(\mu)$
   b) $f(\lambda) \leq_A \{f(\nu) \mid \nu < \lambda\}$

Proof. Clearly, *1.* above implies *2.* above. For the converse direction, assume that $f : \alpha \to A$ satisfies *a*) and *b*). We prove by induction on β that

$$\beta \geq \alpha \vee \forall \mu < \beta\, (f(\beta) <_A f(\mu)).$$

Let $\mu < \beta < \alpha$ be arbitrary. If β is a successor ordinal $\delta + 1$, then, by induction hypothesis and *a*), $f(\beta) <_A f(\delta) \leq_A f(\mu)$. If β is a limit, then $\mu + 1 < \beta$ and thus, from the induction hypothesis and from *b*) we can conclude that $f(\beta) \leq_A f(\mu + 1) <_A f(\mu)$, as desired. $\square$

Definition 97. A linear ordering $(A, <)$ is called a *well order* if one (and thus any) of the following pairwise equivalent conditions is satisfied.

1. *The minimum principle*: $\forall X \subset A\,(\varnothing \neq X \to \exists a \in X\,\forall b \in X\,(a \leq b))$
2. *The minimum principle for sets*: $\forall x \subset A\,(\varnothing \neq x \to \exists a \in x\,\forall b \in x\,(a \leq b))$
3. *The chain condition*: $\forall f\,((f : \omega \to A) \to \exists n, m \in \omega\,(n < m \wedge f(n) \leq f(m))$
4. *Transfinite induction:* $\forall X\,(\forall a \in A\,(A_{<a} \subset X \to a \in X) \to A \subset X)$

We write $\mathsf{wo}(A, <)$ to mean that $(A, <)$ is a well order. If $<$ is a linear ordering of V we might also just write $\mathsf{wo}(<)$ to mean $\mathsf{wo}(V, <)$.

Proof of the equivalence. $1 \Rightarrow 2$: Immediate.

$2 \Rightarrow 3$: Let $f : \omega \to A$ be any function. Applying the minimum principle, we get the least element (with respect to $(A, <)$) a_0 of $f[\omega] \subset A$, which by the axiom of replacement is a set . Let $n \in \omega$ be any natural number such that $f(n) = a_0$. It follows that $f(n) \leq f(n+1)$, as desired.

$1 \Rightarrow 4$: By contrapositive, assume that there is a class X that satisfies the premise of transfinite induction, but still does not contain all of A. Stipulating Y for $A \backslash X$, we get a nonempty class that satisfies

$$A_{<a} \cap Y = \varnothing \to a \notin Y$$

for all $a \in A$. Obviously, such a class Y has no least element.

$4 \Rightarrow 1$: Let X be a subclass of A with no least element. Clearly, $Y = A \backslash X$ satisfies the premise of transfinite induction, thus we have that $Y = A$ and therefore that the only subclass of A with no least element is the empty set.

$3 \Rightarrow 1$: By contrapositive, assume that there is a nonempty class $X \subset A$ that has no least element. We apply Corollary 74 to the formula

$$\psi(f, x) \equiv \mathsf{fun}(f) \wedge rng(f) \subset X \wedge x \in X \wedge x < rng(f)$$

to obtain a function $F : On \to V$ with the property

$$\forall \alpha\,\big(\exists x\,\psi(F \restriction \alpha, x) \to \psi(F \restriction \alpha, F(\alpha))\big).$$

This implies that

$$\forall\alpha\left(\exists x \in X\,(x < rng(F \restriction \alpha)) \rightarrow F(\alpha) \in X_{<rng(F\restriction\alpha)}\right).$$

Since X has no least element, the premise of the above implication is indeed satisfied for all $\alpha < \omega$, hence we have that

$$\forall n \in \omega\left(F(n) \in X_{<rng(F\restriction n)}\right)$$

and thus that $F \restriction \omega$ is a descending chain of length ω. □

Lemma 98. *Every proper initial segment of a well order $(A, <_A)$ is of the form $A_{<_A a}$ for some $a \in A$. Moreover, if a linear order has this property, then it is already a well order.*

Proof. Let $X \subsetneq A$ be an initial segment. We chose $a = \min_{<_A}(A \backslash X)$ and thus get $X = A_{<a}$. If $(A, <_A)$ is a linear order all of whose proper initial segments are of the form $A_{<_A a}$ for some $a \in A$, then all coinitial subclasses of $(A, <_A)$ have a least element. Therefore, if X is any nonempty subclass of A, then

$$Z = \{x \in A \mid \exists y \in X\,(y \leq_A x)\}$$

is a nonempty coinitial segment of $(A, <_A)$ and thus we can find a minimal element a of X from stipulating $a = \min_{<_A}(Z)$. □

Remark 99. Any restriction of a well ordering is a well ordering.

Proof. This follows immediately from the observation that any possible infinite descending chain in the restriction of a linear ordering is also an infinite descending chain in the original ordering. □

Lemma 100. *Every class can be well ordered.*

Proof. Let X be any class. Since the class of all ordinals is not a set, there exists a bijective function $F : V \rightarrow On$. Since On is well ordered by $\in$, so is X by

$$<_X = \{\langle a, b\rangle \mid a, b \in X \wedge F(a) \in F(b)\}.$$ □

Lemma 101. *Ordinals are exactly those sets that are transitive and well ordered by the elementhood relation. The class On is well ordered by the elementhood relation.*

Proof. It is immediate from Corollary 58 that every ordinal is a transitive set that is well ordered by $\in$. We have to show that for any transitive set x, the assumption $\mathsf{wo}(x, \in)$ entails that all elements of x are transitive. Assume that this is not the case and let $y_0 = \min_{\in}\{y \in x \mid \neg\mathsf{tr}(y)\}$. If there are some sets a, b with $a \in b \in y_0$, then by the observation made in Fact 38 it cannot be the case that $y_0 \in a$ nor that $y_0 = a$. Thus, since $\in$ is total on x, it has to be the case that $a \in y_0$. This contradicts our choice of y_0. $\square$

Lemma 102. *Let $(A, <_A)$ and $(B, <_B)$ be well orders with disjoint fields. The sum of $<_A$ and $<_B$*

$$<_{(A,<_A)\oplus(B,<_B)} = \{\langle x, y\rangle \mid x <_A y \vee x <_B y \vee (x \in A \wedge y \in B)\}$$

is a well ordering of the union $A \cup B$. Given two well orders $(A, <_A)$ and $(B, <_B)$ with disjoint fields, we wil write $(A, <_A) \oplus (B, <_B)$ to mean

$$(A \cup B, <_{(A,<_A)\oplus(B,<_B)}).$$

Proof. If there is a descending $<_{A\oplus B}$-chain $f : \omega \to A \cup B$, then either $f[\omega] \subset B$ or there is a natural number n such that $f[\omega_{>n}] \subset A$. In any case, at least one of the orders $(A, <_A)$ or $(B, <_B)$ is not a well order. $\square$

Definition 103. Let $(A, <_A)$ and $(B, <_B)$ be linear orders. By an *(order) isomorphism* from $(A, <_A)$ to $(B, <_B)$ we mean a function $F : A \to B$ that satisfies the following conditions:

1. The function F is *surjective* on B, i.e. $F[A] = B$.

2. The function F is *(strictly) order preserving*, i.e.

$$\forall x, y \in A \big(x <_A y \leftrightarrow F(x) <_B F(y)\big).$$

If there is an (order) isomorphism from a linear order $(A, <_A)$ into an initial segment of some linear order $(B, <_B)$, then we say that $(A, <_A)$ *embeds* into $(B, <_B)$. Consequently, we call an isomorphism $F : A \to Y$ from A onto some initial segment Y of $(B, <_B)$ an *embedding* of $(A, <_A)$ into $(B, <_B)$.

Theorem 104 (Global well orderings)**.** *Every well ordering can be extended to a well ordering of V. Moreover, for every well order $(A, <_A)$, there is a well orderings $<'$ of V, such that $(A, <_A)$ is isomorphic to an initial segment of $(V, <')$.*

Proof. Let $(A, <_A)$ be a well order. As we have observed in Lemma 100, there is a well ordering $<_X$ of $X = V \backslash A$. As a consequence of Lemma 102, we find the desired well ordering of V in $<_{A \oplus X}$. $\square$

Remark 105. The fact that any well ordering can be extended to a global well ordering, will be important in Proposition 251 and Proposition 254.

Definition 106. We will write $\lhd$ to mean a well ordering of V. We will refer to $\lhd$ as "the" *global well ordering* of V.

The following two theorems[7] are essentially a summary of chapter 6 §2 of the book [SF10]. We refer to the cited book for a more detailed presentation.

Theorem 107. *Let $(A, <_A)$ and $(B, <_B)$ be linear orders.*

1. *The class $F^{-1} = \{\langle y, x\rangle \mid \langle x, y\rangle \in F\}$ is an order isomorphism from $(B, <_B)$ to $(A, <_A)$ for any order isomorphism $F : A \to B$.*

2. *If there is an order isomorphism between two linear orders, then either both or none of the orders are well orders.*

3. *If $(A, <_A)$ is a well order and $F : A \to A$ is an order isomorphism, then for all $a \in A$ we have that $a \leq_A F(a)$.*

4. *If $H : A \to B$ and $F : A \to B$ are order isomorphisms and $(A, <_A)$ is a well order, then $F = H$.*

[7] The proofs of the theorems, however, are different from the exposition in [SF10].

5. *If there is an order isomorphism between two well orders* $(A, <_A)$ *and* $(B, <_B)$*, then neither of the two well orders* $(A, <_A)$ *and* $(B, <_B)$ *can be a proper initial segment of the other.*

6. *If* $F : A \to B$ *is a bijective function, then* $(B, \{\langle x, y\rangle \mid F'x <_A F'y\})$ *is a well order and* F *is an order isomorphism from* $(A, <_A)$ *to* $(B, <_B)$*.*

Proof. 1. This is a direct consequence of the definition of order isomorphisms.

2. In view of the first item, it is enough to prove that whenever there is an isomorphism from a well order to some linear order, then the latter is also a well order. Let F be an order isomorphism from A to B. Assume that $(A, <_A)$ is a well order and let $X \subset B$ be any class. Since F is bijective, we have that $X = F[F''X]$ and thus that $F(\min_{<A}(F''X))$ is a minimal element of X.

3. If $F : A \to A$ is an order isomorphism, then the class $\{a \in A \mid F(a) <_A a\}$ has no least element; thus, if $(A, <_A)$ is a well order, it must be empty.

4. By way of contradiction, assume that $X = \{a \in A \mid F(a) \neq H(a)\}$ is a nonempty class and let a_0 be its minimal element. Without loss of generality, we can assume that $F(a_0) < H(a_0)$. As $H \restriction A_{<a_0}$ and $F \restriction A_{<_A a_0}$ coincide, there cannot be any element $a <_A a_0$ such that $H(a) = F(a_0)$; thus, since $H[A] = A$, there must exist an element $a_1 >_A a_0$ such that $H(a_1) = F(a_0)$. This however, violates the assumption that H is order preserving.

5. This follows from the uniqueness of order isomorphisms between well orders and the fact that the identity map is an order isomorphism on any linear order.

6. As F is bijective by assumption, we have that $F[A] = B$. The ordering on B is obviously defined so that F becomes order preserving. The rest follows from part item 2. □

Definition 108. A linear order whose proper initial segments are all sets is called a *set-like* ordering.

Theorem 109 (Comparability theorem)**.** *If $(A, <_A)$ and $(B, <_B)$ are set-like well orders, then there is either an order isomorphism from A to an initial segment of B, or there is an order isomorphism from B to an initial segment of A.*

Proof. We use elementary comprehension to define a class F as follows. First, let X be the class

$$X = \{f \mid \exists a \in A\, \exists b \in B\, (f \text{ is an order isomorphism from } A_{<_A a} \text{ to } B_{<_B b}\}.$$

Let $f : A_{<_A a} \to B_{<_B b}$ and $g : A_{<_A a'} \to B_{<_B b'}$ be any elements of X. We prove that either $f \subset g$ or $g \subset f$ holds. We can assume without loss of generality that $a <_A a'$ holds. Since $g \restriction A_{<_A a}$ is an order isomorphism from $A_{<_A a}$ to $g[A_{<_A a}]$, we have as a consequence from the previous theorem that $g[A_{<_A a}] = B_{<_B b}$ and thus that $g \restriction A_{<a_A} = f$. Now, since X consists of pairwise compatible functions, the class $F = \cup X$ is a function. Clearly, F is an order isomorphism from some initial segment I_A of A to some initial segment I_B of B. If either $I_A = A$ or $I_B = B$ holds, then we are done. If both classes I_A and I_B are proper initial segments of A and B respectively, then, because A and B are set-like well orders, the class F is a $\subset$-maximal order isomorphism between proper initial segments of A and B respectively. Further, by Lemma 98, there are a, b in A and B such that $I_A = A_{<_A a}$ and $I_B = B_{<_B b}$. Since the function $\hat{F} = F \cup \{\langle a, b\rangle\}$ is an order isomorphism between $A_{\leq_A a}$ and $B_{\leq_B b}$, the $\subset$-maximality of F implies that either $A_{\leq_A a} = A$ or $B_{\leq_B b} = B$ holds. □

Corollary 110. *Every set-like well order is order isomorphic to a unique initial segment of $(On, \in)$. In particular, every set-like well order that is a set is order isomorphic to an ordinal, and every set-like well order that is a proper class is order isomorphic to $(On, \in)$.*

Corollary 111. *Every well order $(A, <_A)$ that is not set-like, contains an element $a \in A$ such that $(A_{<a}, <_A)$ is order isomorphic to $(On, \in)$.*

Proof. Let $(A, <_A)$ be as in the claim. Fix $a = \min_{<_A}\{x \in A \mid A_{<_A x} \text{ is not a set}\}$. Clearly, $(A_{<a}, <_A)$ is a set-like well order that is a proper class, thus the claim folows from Corollary 110. □

Remark 112. The previous corollary enables us to assign to each set-like well order $(A, <_A)$ an *ordertype* or *length* being the least ordinal α such that $(A, <_A)$ is order isomorphic to $(\alpha, \in)$ if there exists any such ordinal and On otherwise. Also note that every set-like well order is isomorphic to its order-type.

2 Trees and κ-well orders

2.1 κ-well orders

In this section, we introduce the concept of κ-well orders. Thereafter we will continue with investigating to which degree fundamental properties of well orders have to be weakened to suit the new concept. In particular, we will introduce appropriate induction and minimum principles. We will see that many of the characteristic properties of well orders can be adapted for κ-well orders by natural means, that are moreover compatible with our motif of interpreting small entities as sets and larger entities as proper classes.

Remark 113. It has been brought to our attention that κ-well orders have been studied before by Oikkonen in [Oik92]. However, we follow a different approach and our presentation is independent from Oikkonen's work.

Definition 114. A *cardinal* κ is an ordinal that satisfies $\forall\alpha\,(|\alpha| = |\kappa| \to \kappa \leq \alpha)$. We write $\mathsf{card}(x)$ to mean that x is a cardinal.

Lemma 115. *For every set x, there is a unique cardinal κ such that a bijective function $f : x \to \kappa$ exists.*

Proof. Recall that $|x| = |y|$ stands for the assumption that there is a bijective function from x to y. Let x be any set. Since x can be well ordered, there exists an ordinal α_x with $|\alpha_x| = |x|$. We get a cardinal with the desired property by stipulating $\kappa = \min\{\alpha \mid |\alpha| = |\alpha_x|\}$. The uniqueness of the cardinal κ is a consequence of the fact that two distinct cardinals can never be put into a one-to-one correspondence. $\square$

Definition 116. The *cardinality* of a set x is the unique cardinal κ such that $|x| = |\kappa|$. We write $|x|$ to mean the cardinality of x.

Definition 117. Let κ and λ be ordinals and let X be any class. We introduce the following abbreviations:

1. $[X]^{<\kappa} = \{x \mid x \subset X \wedge |x| < \kappa\}$.
2. $[X]_{>\kappa} = \{x \mid x \subset X \wedge |x| > \kappa\}$.
3. $[X]^{<\kappa}_{>\lambda} = [X]^{<\kappa} \cap [X]_{>\lambda}$.

Definition 118. Let $(A, <_A)$ be a linear order. A class $X \subset A$ is called *progressive* in $(A, <_A)$ if

$$\forall a \in A\, (A_{<_A a} \subset X \to a \in X)$$

holds. We will write $\mathsf{prog}(X, (A, <_A))$ to mean that X is progressive in $(A, <_A)$.

Given the notion of progressive subclasses, we can formulate transfinite induction along a linear order $(A, <_A)$ as the assumption that A contains no proper and progressive subclasses. In the following, we will generalize the notion of progressiveness and thus we will also obtain new induction principles.

Definition 119. Let $(A, <_A)$ be a linear order; a class $X \subset A$ is *κ-progressive* in $(A, <_A)$ if

$$\forall m \in [A]^{<\kappa}_{>0}\, (A_{<m} \subset X \to m \cap X \neq \varnothing).$$

We write $\mathsf{prog}_\kappa(X, (A, <_A))$ to mean that X is a κ-progressive subclass of $(A, <_A)$.

Remark 120. Note that our notion of k-progressive for any $1 < k < \omega$ coincides with the usual notion of progressive.

Proof. Let $1 < k < \omega$ be a natural number and let $(A, <)$ be any linear order. Let $X \subset A$ be progressive and let $m = \{x_0, \dots, x_n\}$ with $0 \leq n < k - 2$ be any subset of A. Since A is linearly ordered, we can assume that x_0 is the smallest element of m. Therefore, $A_{<m} = A_{<x_0}$ and thus it follows from the assumption that X is progressive that

$$A_{<m} \subset X \Rightarrow A_{<x_0} \subset X \Rightarrow x_0 \in X \Rightarrow m \cap X \neq \varnothing$$

and thus that X is k-progressive. The fact that every k-progressive set is also progressive is immediate from the observation that $[A]^{<k}_{>0}$ contains all singleton subsets of A if only $k > 1$. $\square$

Remark 121. Obviously, if $\alpha < \lambda$ are ordinals, then for any linear order $(A, <_A)$, every class X that is λ-progressive in $(A, <_A)$ is also α-progressive in $(A, <_A)$.

Lemma 122. *If κ is a nonzero ordinal, $(A, <_A)$ is a well order and $X \subset A$, then the following are equivalent.*

1. $\mathsf{prog}(X, (A, <_A))$
2. $\mathsf{prog}_\kappa(X, (A, <_A))$

Proof. In view of the preceding remark, we only have to show that *1.* above implies *2.* above. Let $m \subset A$ be a nonempty set of cardinality less than κ such that $A_{<_A m} \subset X$. If we fix $a_m = \min_{<_A}(m)$, then $A_{<_A a_m} \subset X$ and thus by $\mathsf{prog}(X, (A, <_A))$ we have that $a_m \in X$, hence $m \cap X \neq \varnothing$. $\square$

Definition 123. For a linear order $(A, <_A)$, we say that:

1. The *κ-minimum principle* holds in $(A, <_A)$ if for every class $X \subset A$, there exists a minimal subset $m \subset X$ of cardinality less than κ. If the κ-minimum principle holds in $(A, <_A)$, then we say that $(A, <_A)$ is a *κ-well order*. We write $\mathsf{wo}_\kappa(A, <_A)$ to mean that $(A, <_A)$ is a κ-well order.
2. The *κ-chain condition* holds in $(A, <_A)$, and write $\mathsf{cc}_\kappa(A, <_A)$ if there is no descending chain of length κ in $(A, <_A)$.
3. $(A, <_A)$ satisfies *κ-transfinite induction*, if no proper subclass of A that is κ-progressive exists. In that case, we write $\mathsf{ti}_\kappa(A, <_A)$.

Lemma 124. *Let $(A, <_A)$ be a linear order, κ a cardinal and X any subclass of A, then the following three assertions are pairwise equivalent.*

1. *X contains a minimal set of cardinality less than κ.*
2. *The class $\{a \in A \mid \exists x, y \in X \, (x \leq_A a \leq_A y)\}$ contains a minimal subset of cardinality less than κ.*
3. *The class $\{a \in A \mid \exists x \in X \, (x \leq_A a)\}$ contains a minimal subset of cardinality less than κ.*

Proof. Clearly, every minimal subset of X is also a minimal subset of

$$Y = \{a \in A \mid \exists x, y \in X\,(x \leq_A a \leq_A y)\}$$

and since Y is an initial segment of

$$Z = \{a \in A \mid \exists x \in X\,(x \leq_A a)\}$$

every minimal subset of Y is also a minimal subset of Z. Thus it is enough to prove that whenever Z has a minimal subset of a given cardinality, then so does X. Let $m = \{m_\alpha \mid \alpha < \lambda\}$ be a minimal subset of Z with $\lambda < \kappa$. If $m = \varnothing$, then this means that $Z = \varnothing$ and thus also that $X = \varnothing$ and we are done. If there exists an element $x \in m$ such that $X_{<_A x} = \varnothing$, then x is a minimal element of X and thus $\{x\}$ is a minimal subset of X of cardinality less that κ and we are done. In all the remaining cases, we know that for all $\alpha < \lambda$ it is the case that $X_{<_A m_\alpha} \neq \varnothing$. Hence we can define a function $f : \lambda \to A$ from

$$f(\alpha) = \min_{\lhd}(X_{<_A m_\alpha}).$$

In this case, $f[\lambda]$ is the desired minimal set of X with cardinality less than κ. $\square$

Proposition 125. *Let $(A, <_A)$ be any linear order. The following are equivalent:*

1. $\mathsf{ti}_\kappa(A, <_A)$
2. $\mathsf{wo}_\kappa(A, <_A)$

Proof. If $(A, <_A)$ is not a κ-well order, then there exists a subclass X of A whose minimal subsets all have at least cardinality κ. Let $Y = A \backslash X$. In order to prove that 1. $\Rightarrow$ 2. holds, it is enough to show that Y is κ-progressive. Let $m \in [A]^{<\kappa}_{>0}$ be arbitrary and consider

$$\begin{aligned} m \cap Y = \varnothing \Rightarrow m &\subset X \\ &\Rightarrow \exists x \in X\, \forall y \in m\,(x < y) \\ &\Rightarrow \exists x \in X\,(x \in A_{<m}) \\ &\Rightarrow A_{<m} \not\subset Y. \end{aligned}$$

For the converse direction, assume that X is a κ-progressive subclass of A. The class $Y = A\backslash X$ must, by the κ-minimum principle, possess a minimal subset $m \subset Y$ of cardinality less than κ. Because of $A_{<_A m} \subset X$ and X is κ-progressive, either $m = \varnothing$ or $m \cap X \neq \varnothing$. In the first case $\varnothing$ is minimal in Y and thus $X \supset A$. The latter case, however, cannot occur since any $x \in X \cap m \subset X \cap Y$ contradicts $X \cap Y = \varnothing$. □

Definition 126. Let α be an ordinal. The *cofinality* $cf(\alpha)$ of α is the least ordinal β such that there exists a sequence $f : \beta \to \alpha$ with the property that $f[\beta]$ is a cofinal subset of $(\alpha, \in)$. An ordinal α that satisfies $\alpha = cf(\alpha)$ is called a *regular cardinal.*

Remark 127. Note that for every ordinal α, the cofinality $cf(\alpha)$ of α is a cardinal.

Proposition 128. *Let $(A, <_A)$ be a linear order. The following are pairwise equivalent:*

1. $\forall (A, <_A) \left(\mathsf{wo}_\kappa(A, <_A) \leftrightarrow \mathsf{cc}_\kappa(A, <_A)\right)$
2. $\forall (A, <_A) \left(\mathsf{wo}_\kappa(A, <_A) \to \mathsf{cc}_\kappa(A, <_A)\right)$
3. *κ is a regular cardinal.*

Proof. 1⇒2. Clear.

2⇒3. Assume that for some ordinal κ the equivalence displayed in *1.* holds. We consider the linear ordering $(A, <_A) = \kappa^{-1} = (\kappa, >)$, i.e. κ together with its usual ordering inverted. Obviously, $(A, <_A)$ does not satisfy the κ-chain condition. Hence, by assertion *2.*, $(A, <_A)$ is not a κ-well order. Therefore, by Lemma 124, there is a coinitial (with respect to the inverted ordering $<_A$) subset $x \subset \kappa$ that has no minimal subset of size less than κ. In terms of the standard ordering of κ, this means that x is an initial segment of κ that has no cofinal subset of size less than κ. Therefore, there is an ordinal $\mu \leq \kappa$ with $cf(\mu) \geq \kappa$, hence $cf(\kappa) = \kappa$.

3⇒1. Now Let $(A, <_A)$ be any linear order and let κ be a regular cardinal. We have to prove that $\mathsf{cc}_\kappa(A, <_A) \leftrightarrow \mathsf{wo}_\kappa(A, <_A)$ holds. If a descending chain

$f : \kappa \to A$ exists, then the set $f[\kappa]$ has no minimal set of size less than $cf(\kappa) = \kappa$. On the other hand, if there is a subclass X of A that has no minimal subset of size less than κ, then we can define a function $f : \kappa \to A$ by stipulating

$$f(\alpha) = \min_{\triangleleft}(X_{<\{f(\mu)|\mu<\alpha\}}).$$

Such a function clearly violates the κ-chain condition. □

Corollary 129. *Let $(A, <_A)$ be a linear order. If κ is a regular cardinal, then the κ-chain condition, the κ-minimum principle and the κ-induction principle are pairwise equivalent.*

2.2 Trees

Definition 130. We call a function f whose domain is an ordinal a *sequence*. The class of all sequences is denoted by *Seq*. For any class X, we write $Seq(X)$ to mean the class of all sequences that range into X, i.e.

$$Seq(X) = \{f \in Seq \mid rng(f) \subset X\}.$$

If f, g are two sequences, then we write $f \perp g$ to mean that f and g are *incompatible* in the sense that they do not have a common extension, i.e. $f \cup g \notin Seq$.

Remark 131. Note that the subset relation on sequences corresponds to the notion of extending a sequence by some further elements.

Definition 132. A *tree* is a subclass T of the class *Seq* that is closed under *initial sequences*. Formally, this can be written as

$$\mathsf{tree}(T) \;\equiv\; T \subset Seq \wedge \forall f, \alpha\, (f \in T \to f \restriction \alpha \in T).$$

The tree $Bin = Seq(\{0, 1\})$ is called the *full binary tree*; its elements are referred to as *binary sequences*. Any tree $T \subset Bin$ is called a *binary tree*.

In the literature of set theory (cf. Definition 5.1 in [Kun80]), trees are usually defined as pairs $(T, <_T)$ where $<_T$ is a partial ordering of T such that for each

$x \in T$, the set $\{y \in T \mid y <_T x\}$ is well ordered. To see that our and the usual tree notions are interchangeable, we provide a uniform way to construct a tree $\hat{T}$ from a given tree $(T, <_T)$ and a function $F : T \to \hat{T} \subset Seq$ that satisfies

$$\forall x, y \in T\, (x <_T y \leftrightarrow F(x) \subsetneq F(y)).$$

Clearly, if $T \subset Seq$ is a tree, then the pair $(T, \subsetneq)$ is a tree in the usual sense. Vice versa, given any "usual tree" $(T, <_T)$, then we obtain a tree $\hat{T}$ if we enumerate the branches of the original tree, i.e. by stipulating

$$f \in \hat{T} \Leftrightarrow f \in Seq(T) \wedge \forall \alpha \in dom(f) \big(T_{<_T f(\alpha)} = f[\alpha]\big).$$

Note that for every $f \in \hat{T}$ and all ordinals $\alpha, \beta \in dom(f)$, we have that $\alpha < \beta$ if and only if $f(\alpha) <_T f(\beta)$. The function $F : T \to \hat{T}$ can be obtained from

$$\begin{aligned} x \mapsto\ & \text{The uniquely determined order isomorphism } f_x \text{ between} \\ & (\{y \in T \mid y <_T x\}, <_T) \text{ and its order-type.} \end{aligned}$$

Definition 133. 1. For any sequences $f, g \in Seq$ we define the *concatenation* $f \frown g$ of f and g as

$$(f \frown g)(\alpha) = \begin{cases} f(\alpha) & \text{if } \alpha < dom(f) \\ g(\beta) & \text{if } \alpha = dom(f) + \beta \,\wedge\, \beta < dom(g). \end{cases}$$

Note that $dom(f \frown g) = dom(f) + dom(g)$. We will also use the abbreviation $f \frown x$ to mean $f \frown \{\langle 0, x\rangle\}$.

2. Let x be any set. We write $(x)_\alpha = \{\langle \mu, x\rangle \mid \mu < \alpha\}$ to mean the *constant function* with domain α and range $\{x\}$.

3. Let T be a tree. We write $Lv(T, \alpha) = \{f \in T \mid dom(f) = \alpha\}$ to mean the α^{th} *level* of T. The *height* $H(T)$ of the tree T is the least ordinal λ such that $Lv(T, \lambda) = \varnothing$ if such an ordinal exists and the class On of all ordinals otherwise.

4. For any subclass X of Seq, we write $Tr(X) = \{f \in Seq \mid \exists h \in X\, (f \subset h)\}$

to mean the *least tree* that contains all of X.

5. An *(tree-)embedding* of a tree T into a tree H is a mapping $E : T \to H$ such that
$$\forall f, g \in T\, (f \subsetneq g \leftrightarrow E(f) \subsetneq E(g))$$
holds. An embedding that is onto we shall call an *isomorphism* of the respective trees.

Remark 134. Note that for any tree T and any subclass X of T, it is the case that $Tr(X)$ is the *least subtree* of T that contains X.

As in our setting we have a global well ordering, we can further simplify our investigations of trees by considering only ordinal trees, that is trees that are subclasses of $Seq(On)$.

Lemma 135. *If T is a tree and $F : V \to On$ is any bijection, then the class $\{F \circ f \mid f \in T\} \subset Seq(On)$ is a tree that is isomorphic to T by dint of the mapping $f \mapsto F \circ f$.*

Proof. The fact that the proposed mapping is indeed an isomorphism follows directly from the assumption that F is bijective. $\square$

Definition 136. Let κ be an cardinal. A tree T is *κ-branching* if for every $f \in T$, the cardinality of the set $\{x \mid f \frown x \in T\}$ is less than κ. Similarly, a tree is called *set branching* if for every $f \in T$, the class $\{x \mid f \frown x \in T\}$ is a set.

Proposition 137. *A tree T is set branching if and only if all its levels are sets.*

Proof. Let T be any tree. If all the levels of T are sets, then every class of the form $\{x \mid f \frown x \in T\}$ where $f \in T$ is a subclass of the set $Lv(T, dom(f) + 1)$ and hence itself is a set. For the converse direction, assume that T is set branching. We prove by induction on α that $Lv(T, \alpha)$ is a set. For the successor case, assume that $\alpha = \beta + 1$. We have that

$$\begin{aligned} Lv(T, \alpha) &= \{f \in T \mid dom(f) = \beta + 1\} = \{f \frown x \in T \mid dom(f) = \beta\} \\ &= \{f \frown x \mid f \in Lv(T, \beta) \land f \frown x \in T\} \\ &= \bigcup_{f \in Lv(T,\beta)} \{f \frown x \mid f \frown x \in T\}, \end{aligned}$$

which by induction hypothesis is a set. In case that α is a limit ordinal, note that a function $f : \alpha \to V$ can only be an element of T (and hence of $Lv(T, \alpha)$) if all the functions $\{f \restriction \beta \mid \beta < \alpha\}$ already belong to T. Hence, the mapping $f \mapsto \{f \restriction \beta \mid \beta < \alpha\}$ is clearly an injection of $Lv(T, \alpha)$ into the powerclass of $\bigcup_{\beta<\alpha} Lv(T, \beta)$. Since the latter is a set by induction hypothesis, the claim is proved. □

Definition 138. Let T be a tree and let λ be an ordinal. A *branch of length* λ, also called a λ-*branch*, of T is a function $f : \lambda \to V$ such that $\forall \alpha < \lambda\, (f \restriction \alpha \in T)$. A *path* through T is a function $F : On \to V$ such that $\forall \alpha\, (F \restriction \alpha \in T)$.

Naturally, all the trees under consideration in arithmetic are subsets of $Seq(\omega)$. It is a peculiarity of the arithmetic setting (assuming arithmetical comprehension) that for any tree T, the statement that every element of T can be properly extended is a sufficient condition that there is an infinite branch in T (note that this within arithmetic becomes the statement that T has a path). To see that, let T be a tree that satisfies the aforementioned condition. We can then define an infinite branch $f : \omega \to \omega$ by recursion on the natural numbers as given from

$$f(n) = \min\{k \in \omega \mid \exists s \in T\, (f \restriction n \subsetneq s) \wedge s(n) = k\}.$$

The following example is to illustrate that, in this respect, the situation is quite different in set theory. We explicitly give a tree in which every element can be arbitrarily extended within the tree, but there is no path through the tree nevertheless. In a wider sense, this example seems to give emphasis for our impression that in set theory it seems much harder to give *local* conditions that imply that a given tree has a path than in arithmetic.

Example 139. Let T be the tree that contains exactly the functions of the form

$$(\alpha_0)_{\lambda_0} \frown (\alpha_1)_{\lambda_1} \frown \cdots \frown (\alpha_n)_{\lambda_n}$$

with $n \in \omega$ and $\alpha_1 < \alpha_2 < \cdots < \alpha_n$ and for all $i \leq n$ the constants α_i are greater than λ_i. Figuratively, the elements of T resemble stairs whose treads are at most as long as their height over ground. Note that given any element f of T and any ordinal γ, there is always a function $h \in T$ that has length (at least) γ and that extends f. For example, if f is

$$(\alpha_0)_{\lambda_0} \frown (\alpha_1)_{\lambda_1} \frown \cdots \frown (\alpha_n)_{\lambda_n},$$

then pick h to be the function

$$h = (\alpha_0)_{\lambda_0} \frown (\alpha_1)_{\lambda_1} \frown \cdots \frown (\alpha_n)_{\lambda_n} \frown (\alpha_n + \gamma + 1)_{\lambda_n + \gamma}.$$

Note that despite the fact that every element of T has arbitrary long extensions, we can show that T has no path. By way of contradiction, assume that T has a path $F : On \to On$. If $rng(F)$ is finite, then there must exist a maximal element α in $rng(F)$, since all elements of T are monotonously increasing sequences, this implies that F is of the form $g \frown (\alpha)_{On}$ for some suitable initial sequence g, contradicting $\forall \gamma\,(F \restriction \gamma \in T)$. If otherwise the range of F is infinite, then we can define an infinite increasing sequence of ordinals $\{\gamma_i \mid i \in \omega\}$ by

$$\gamma_i = \min\{\alpha \mid |rng(F \restriction \alpha)| > i\}.$$

Stipulating $\gamma = \sup\{\gamma_i \mid i \in \omega\}$ it follows that $F \restriction \gamma$ has a range of infinite cardinality, in contradiction to $F \restriction \gamma \in T$.

Comparing to the situation as it is in the setting of arithmetic, i.e. when speaking of trees with height less or equal than ω, where it is a sufficient condition for a tree to have a branch of length ω that every element of the tree can be extended by one further step. The previous example tells us that in the set theoretic environment, the situation is more complicated; there are trees that do not have paths but where but still every element of the tree can be extended arbitrarily far.

Now we turn our attention to the so called *Kleene Brouwer ordering*, a linear ordering of the class Seq that allows us to connects trees and linear orders in a beneficial way. The Kleene Brouwer ordering will be an important tool in our discussion of the Π^1_1 completeness of weak well orders.

Definition 140. The *Kleene Brouwer ordering* $<_{KB}$ is the linear ordering of

Seq given from

$$f <_{KB} g \Leftrightarrow g \subsetneq f \vee \exists \alpha\, (\alpha \in dom(f) \cap dom(g) \wedge f \restriction \alpha = g \restriction \alpha \wedge f(\alpha) \lhd g(\alpha)).$$

Given a tree T, we say that $<_{KB} \cap (T \times T)$ is the *Kleene Brouwer ordering of* T. We will write $<_T$ or KB_T, whichever makes the respective text component more readable, to mean the Kleene Brouwer ordering of T.

Remark 141. For any tree T, the pair $(T, <_T)$ is a linear order.

Proposition 142. *Let κ be any cardinal. For any tree T, the following implications hold:*

$$\mathsf{cc}_\kappa(T, <_T) \Rightarrow \ T \textit{ has no } \kappa\textit{-branch} \Rightarrow \mathsf{wo}_\kappa(T, <_T).$$

Proof. The first implication follows from the fact that if $b : \kappa \to T$ is a branch of length κ, then the mapping $\alpha \mapsto b \restriction \alpha$ induces a descending chain of length κ in $(T, <_T)$, violating the κ-chain condition. For the second implication, assume that $(T, <_T)$ is not a κ-well order. By Lemma 124, we can assume that there exists a coinitial subclass (and thus a subtree) X of T that has no minimal subset of cardinality less than κ. Let $\varphi(f, x)$ be a formula that signifies

$$x = \min_{\lhd}\{y \mid f \frown y \in X\}.$$

From corollary 74 it follows that there exists a function $F : On \to V$ with the property that $\forall \alpha\, (\exists x\, \varphi(F \restriction \alpha, x) \to \varphi(F \restriction \alpha, F(\alpha)))$. If we can prove for all ordinals $\alpha < \kappa$ that $F \restriction \alpha \in X$ holds, then $F \restriction \kappa$ is a branch of length κ in T and we are done. By way of contradiction, assume that μ is the least ordinal less than κ such that $F \restriction \mu \notin X$. It follows that $y = \{F \restriction \nu \mid \nu < \mu\} \subset X$. If there is a sequence f such that $f \in X_{<y}$, then f cannot be a common and proper extension of all the elements of y, as otherwise f extends $F \restriction \mu$, contradicting the fact that X is closed under initial segments. Thus, there must be an ordinal $\theta < \mu$ such that $F \restriction \theta = f \restriction \theta$ and $f(\theta) \lhd F(\theta)$; this cannot be the case since $f \in X$ and the way we build F. Therefore, $X_{<y}$ is empty and thus y is a minimal subset of X of cardinality less than κ, contradicting our choice of X. $\square$

Corollary 143. *If κ is a regular cardinal and T is any tree, then the following are equivalent:*

1. *There is no branch of length κ in T.*

2. *The Kleene Brouwer ordering $<_T$ is a κ-well ordering of T.*

Proof. This is a direct consequence of the κ-chain condition and κ-weak well orderedness being equivalent for regular cardinals κ. $\square$

König's Lemma and its relatives in class theory

The standard formulation of König's Lemma is the assertion that every finitely branching infinite tree posseses an infinite branch. We call this principle König's Lemma for ω (cf. Proposition 152). The principle which plays the role that König's Lemma for ω plays in the arithmetical setting, is what we call König's Lemma for On (cf. Definition 153); the assertion that every set branching tree that is a proper class has a path. Further, in order to obtain the weak versions of König's Lemma for ω and On, as usual, the respective principles are restricted to binary trees. While König's Lemma in the arithmetical setting is provable in the theory ACA_0, König's Lemma for On is not provable in our base theory NBG (cf. Remark 155). While in general it is not possible to reduce the question whether or not a given tree has a path to the same question about some specific binary tree, we present a construction that transforms an arbitrary tree into a binary tree (cf. Definition 144), such that the original tree has a path if and only if the binary tree has a particular kind of path (cf. Proposition 150). Moreover, if the construction is initialized with a set branching tree, then we obtain that the original tree has a path if and only if the binary tree has the path (cf. Proposition 151). As a result, we can prove that the König's Lemma for On and weak König's Lemma for On are equivalent (cf. Theorem 154). Finally, we will introduce the theory $s\Pi^1_1$-Ref an show that this theory is strong enough to prove König's Lemma for On (cf. Theorem 164).

Definition 144. We define a *canonical embedding*

$$\mathcal{B} : Seq(On) \to Bin$$

of the tree $Seq(On)$ into[1] the tree Bin by recursion on the levels of $Seq(On)$ as follows:

$$\mathcal{B}(f) = \begin{cases} \varnothing & \text{if } dom(f) = 0 \\ \mathcal{B}(f \upharpoonright \beta) \frown (1)_{f(\beta)} \frown (0)_1 & \text{if } dom(f) = \beta + 1 \\ \bigcup_{\alpha < dom(f)} \mathcal{B}(f \upharpoonright \alpha) & \text{if } dom(f) \text{ is a limit ordinal.} \end{cases}$$

Remark 145. Informally, for any sequence $f \in Seq$, the sequence $\mathcal{B}(f)$ takes the form

$$\mathcal{B}(f) = \langle \underbrace{1, \cdots, 1}_{f(0)}, 0, \underbrace{1, \cdots, 1}_{f(1)}, 0, \cdots, 0, \underbrace{1, \cdots, 1}_{f(\alpha)}, 0, \cdots \rangle.$$

Lemma 146. *For any tree $T \subset Seq(On)$, the function $\mathcal{B} \upharpoonright T$ is an embedding of T into Bin.*

Proof. Let T be a tree as in the claim. First, we prove that for any two elements $f, g \in T$, we have that

$$f \subset g \to \mathcal{B}(f) \subset \mathcal{B}(g). \tag{1}$$

Proceed by induction on the domain of g. Let λ be the domain of g. Without loss of generality, assume that $f = g \upharpoonright \beta$ for some $\beta < \lambda$. If $\lambda = \alpha + 1$ is a successor ordinal, then it follows from the induction hypothesis that

$$\begin{aligned} \mathcal{B}(f) = \mathcal{B}(g \upharpoonright \beta) \subset \mathcal{B}(g \upharpoonright \alpha) \\ \subset \mathcal{B}(g \upharpoonright \alpha) \frown (1)_{g(\alpha)} \frown (0)_1 = \mathcal{B}(g \upharpoonright \alpha + 1) = \mathcal{B}(g). \end{aligned}$$

For the case where $\lambda = dom(g)$ is a limit ordinal and $f = g \upharpoonright \beta$ with $\beta < \lambda$, we get that

$$\mathcal{B}(f) = \mathcal{B}(g \upharpoonright \beta) \subset \bigcup_{\mu < \lambda} \mathcal{B}(g \upharpoonright \mu) = \mathcal{B}(g).$$

[1] To see that $\mathcal{B}$ maps, as indicated above, sequences to sequences, note that at the successor stages clearly sequences are mapped to sequences and for the limit stages consider the first part in the proof of the next lemma, where it is shown that $\mathcal{B}$ preserves the subset relation.

Next, we prove that for any $f, g \in T$, we have that

$$f \perp g \to \mathcal{B}(f) \perp \mathcal{B}(g). \tag{2}$$

Let $f, g \in T$ be incompatible and let α be the least ordinal with $f(\alpha) \neq g(\alpha)$. Without loss of generality, we can assume that $f(\alpha) < g(\alpha)$. We have that

$$\mathcal{B}(f \restriction \alpha + 1) = \mathcal{B}(f \restriction \alpha) \frown (1)_{f(\alpha)} \frown 0 = \mathcal{B}(g \restriction \alpha) \frown (1)_{f(\alpha)} \frown 0$$

and

$$\mathcal{B}(g \restriction \alpha + 1) = \mathcal{B}(g \restriction \alpha) \frown (1)_{g(\alpha)} \frown 0.$$

Hence, it follows that

$$\mathcal{B}(f)(dom(\mathcal{B}(f \restriction \alpha) + f(\alpha))) = 0 \neq 1 = \mathcal{B}(g)(dom(\mathcal{B}(f \restriction \alpha) + f(\alpha))).$$

Therefore, $\mathcal{B}(f) \perp \mathcal{B}(g)$. Now we are ready to show that

$$\mathcal{B}(f) \subsetneq \mathcal{B}(g) \to f \subsetneq g \tag{3}$$

holds for all sequences $f, g \in T$. We prove the contrapositive. We assume that $\neg(f \subsetneq g)$ holds, then either $g \subset f$ or $f \perp g$. Applying (1) and (2) for the respective cases, we get that $\mathcal{B}(g) \subset \mathcal{B}(f)$ or $\mathcal{B}(g) \perp \mathcal{B}(f)$, therefore, it is the case that $\neg \mathcal{B}(f) \subsetneq \mathcal{B}(g)$. Regarding (1) and (3), we only have to verify that the properness of inclusions is preserved by $\mathcal{B}$ to finish the proof. Let $f \in T$ and $\alpha < dom(f)$. To see this, consider $\mathcal{B}(f \restriction \alpha) \subsetneq \mathcal{B}(f \restriction \alpha + 1) \subset \mathcal{B}(f)$. $\square$

Definition 147. A sequence $f \in Seq$ is called *good* if it is not terminating in a sequence of 1's, that is if

$$\forall \alpha \in dom(f) \, \exists \beta \in dom(f) \, (\alpha \leq \beta \wedge f(\beta) \neq 1)$$

holds. Similarly, a function $F : On \to V$ is good if

$$\forall \alpha \, \exists \beta \geq \alpha \, (F(\beta) \neq 1).$$

Lemma 148. *Let T be any tree and let f be an element of $Tr(\mathcal{B}[T])$. We have the following equivalence:*

$$f \in \mathcal{B}[T] \Leftrightarrow f \textit{ is good.}$$

Proof. Clearly, any element of the image of $\mathcal{B}$ is good. We have to prove that whenever $f \in Tr(\mathcal{B}[T])$ is good, then there is a function $g \in T$ such that $f = \mathcal{B}(g)$. Since $f \in Tr(\mathcal{B}[T])$, there must exist an element $h \in T$ such that $f \subset h' = \mathcal{B}(h)$. Without loss of generality, we assume that h is of minimal domain with the aforementioned property. If $dom(h) = \delta + 1$ is a successor ordinal, it follows from the minimality of the domain of h that $f \not\subset \mathcal{B}(h \restriction \delta)$. Since f and $\mathcal{B}(h \restriction \delta)$ have a common extension, this implies that they are compatible and thus that $\mathcal{B}(h \restriction \delta) \subset f$. In particular, we get that

$$\mathcal{B}(h \restriction \delta) \subset f \subset \mathcal{B}(h) = \mathcal{B}(h \restriction \delta) \frown (1)_{h(\delta)} \frown 0.$$

Together with the fact that f is good, this implies that $\mathcal{B}(h) = f$. Now assume that $dom(h) = \lambda$ is a limit ordinal. Because $\mathcal{B}$ preserves the subset relation, we know that for every ordinal $\alpha < \lambda$ the function $\mathcal{B}(h)$ is a common extension of $\mathcal{B}(h \restriction \alpha)$ and f. Thus, we have that

$$\forall \alpha < \lambda \, (\mathcal{B}(h \restriction \alpha) \subset f \, \vee \, f \subset \mathcal{B}(h \restriction \alpha)).$$

This implies that for all $\alpha < \lambda$ the assumption that $\mathcal{B}(h \restriction \alpha) \not\subset f$ contradicts the minimality of $dom(h)$, we conclude that $\forall \alpha < \lambda \, (\mathcal{B}(h \restriction \alpha) \subset f)$ and therefore that

$$\mathcal{B}(h) = \bigcup_{\alpha < \lambda} \mathcal{B}(h \restriction \alpha) \subset f$$

holds. Thus we have $\mathcal{B}(h) = f$ as desired. □

Lemma 149. *If $T \subset Seq(On)$ is a set branching tree, then $Tr(\mathcal{B}[T])$ has only, if any, good paths.*

Proof. Let F be a path through $Tr(\mathcal{B}[T])$. We have to show that F is good. Suppose that F is not good and let α be the least ordinal with $\forall \beta > \alpha \, (F(\beta) = 1)$.

From the minimality of α, it follows that $F \restriction \alpha$ is a good element of $Tr(\mathcal{B}[T])$. Applying the previous lemma, we get that there exists a sequence $f \in T$ with $\mathcal{B}(f) = F \restriction \alpha$. Since T is set branching, there exists an ordinal κ such that $\kappa > \sup\{\mu \mid f \frown \mu \in T\}$. Since F is a path through $Tr(\mathcal{B}[T])$, it is the case that $F \restriction (\alpha + \kappa) \in Tr(\mathcal{B}[T])$, hence there must exist an element $g \in T$ that satisfies $F \restriction (\alpha + \kappa) \subset \mathcal{B}(g)$. Summarizing, we have that

$$\mathcal{B}(f) = F \restriction \alpha \subsetneq F \restriction \alpha \frown (1)_\kappa = F \restriction (\alpha + \kappa) \subset \mathcal{B}(g)$$

and in regard of the fact that $\mathcal{B}$ is an embedding, that f is a proper initial segment of g. Therefore, we can conclude that

$$\begin{aligned}\mathcal{B}(f) \frown (1)_{g(dom(f))} \frown 0 &= \mathcal{B}(g \restriction dom(f)) \frown (1)_{g(dom(f))} \frown 0 \\ &= \mathcal{B}(g \restriction (dom(f)+1)) \subset \mathcal{B}(g).\end{aligned}$$

As the functions

$$\underbrace{F \restriction \alpha}_{\mathcal{B}(f)} \frown (1)_\kappa \qquad \text{and} \qquad \mathcal{B}(f) \frown (1)_{g(dom(f))} \frown 0$$

have the common extension $\mathcal{B}(g)$, they must be compatible. Since

$$\mathcal{B}(f) \frown (1)_{g(dom(f))} \frown 0 \not\subset \mathcal{B}(f) \frown (1)_\kappa,$$

this implies that

$$\mathcal{B}(f) \frown (1)_\kappa \subset \mathcal{B}(f) \frown (1)_{g(dom(f))} \frown 0.$$

From this, it follows that $\kappa \leq g(dom(f))$, which contradicts our choice of κ since $f \frown g(dom(f)) \in T$ and $\kappa > \sup\{\mu \mid f \frown \mu \in T\}$. $\square$

Proposition 150. *A tree $T \subset Seq(On)$ has a path if and only if the tree $Tr(\mathcal{B}[T])$ has a good path.*

Proof. Let G be a good path through $Tr(\mathcal{B}[T])$. Since G is good, the class $X = \{\alpha \mid G \restriction \alpha \text{ is good in } Tr(\mathcal{B}[T])\}$ is unbounded in On. Hence, the preimage

Y of $\{G \upharpoonright \alpha \mid \alpha \in X\}$ under $\mathcal{B}$ is a proper class of pairwise compatible functions in T. Therefore, $\cup Y$ is a path through T. For the reverse implication, let F be a path through T. The subtree $S = \{F \upharpoonright \alpha \mid \alpha \in On\}$ of T has a path, hence also $Tr(\mathcal{B}[S])$ has a path. Since clearly S is a set branching tree, we can apply the previous lemma to get that the path through $T(\mathcal{B}[S])$ is a good path. $\square$

The following theorem reduces the question whether or not any set branching tree T of ordinal sequences (and thus any set branching tree) has a path to the question whether or not the binary tree $Tr(\mathcal{B}(T))$ has a path. This observation will be useful when we reduce König's Lemma to the later introduced principle of Π^1_1 reflection.

Theorem 151. *For any set branching tree $T \subset Seq(On)$, we have the following equivalence:*

$$T \textit{ has a path } \Leftrightarrow Tr(\mathcal{B}[T]) \textit{ has a path.}$$

Proof. Since

$$\begin{aligned} T \text{ has a path } &\Leftrightarrow Tr(\mathcal{B}[T]) \text{ has a good path} \\ &\Leftrightarrow Tr(\mathcal{B}[T]) \text{ has a path,} \end{aligned}$$

the claim follows immediately from Lemma 149 and Proposition 150. $\square$

Proposition 152 (König's Lemma for ω)**.** *Every ω-branching tree is either finite or has an infinite branch.*

Proof. Let T be ω-branching and infinite. We define a sequence $f \in Seq(\omega)$ recursively as follows:

$$f(n) = \min\{k \in \omega \mid \omega \leq |\{h \in T \mid (f \upharpoonright n \frown k) \subset h\}|\}.$$

This function is well defined since there is no finite partition of an infinite set into finite parts. Clearly, f is an infinite branch of T. $\square$

The generalization of König's Lemma from ω to arbitrary ordinals is false. There are *Aronszajin trees*, that is trees of height $\omega_1 = \min\{\alpha \mid |\alpha| > \omega\}$ whose levels are all at most countable that have no uncountable branches. For a construction of an Aronszajin tree, see Theorem 9.16 in [Jec03]. An ordinal to which König's Lemma generalizes is called *weakly compact*. For a detailed exposition on weakly compact cardinals and other large cardinals see [Kan09] or II.17 in [Jec03].

Definition 153. The *König's Lemma for On* is the assumption that every set-branching tree is either a set or has a path. We will denote this assumption by KL^+. The principle WKL^+, *weak König's Lemma for On*, is the statement that every binary tree is either a set or has a path.

Theorem 154. *The principles* WKL^+ *and* KL^+ *are equivalent over* NBG.

Proof. Let T be any set branching tree. Without loss of generality, let T be a subclass of $Seq(On)$. Assuming WKL^+, we can conclude:

$$\begin{aligned} T \text{ has a path } &\Leftrightarrow Tr(\mathcal{B}[T]) \text{ has a path} \\ &\Leftrightarrow Tr(\mathcal{B}[T]) \text{ is a proper class} \\ &\Leftrightarrow T \text{ is a proper class.} \end{aligned}$$

□

Remark 155. A weakly compact cardinal is a cardinal κ such that $(V_\kappa, V_{\kappa+1})$ is a model of $\mathsf{NBG} + \mathsf{KL}^+$. Weakly compact cardinals are well known and have been thoroughly studied (e.g. [Jec03] Chapters 9 and 17 and [Kan09]). Since not every inaccessible cardinal is weakly compact (cf. [Jec03] Chapter 17), but for every inaccessible cardinal λ it is the case that $(V_\lambda, V_{\lambda+1})$ is a model of MK, the principle WKL^+ is not provable from MK set theory.

After we have seen that proving $(\mathsf{W})\mathsf{KL}^+$ is far beyond the capabilities of NBG, we briefly mention two principles which both imply KL^+ over NBG. The first is called *weak compactness principle* (cf. Definition 158) and the second is the *(strict)*Π^1_1 *reflection principle* (cf. Definition 162). Apart from our preparation in establishing the equivalence between KL^+ and WKL^+, the proofs (cf. Theorems 159 and 164) that show that these principles entail KL^+ are a straightforward adaption of standard set theory proofs (cf. Theorems II.2.17.13 and II.17.18 in [Jec03]).

Definition 156. The infinitary language $\mathcal{L}_{On,\omega}$ is given from the usual language $\mathcal{L}$ of logical symbols together with the following parts:

1. For each ordinal α a variable x_α.

2. Relation, function and constant symbols.

3. Infinitary connectives $\bigvee_{\alpha<\lambda}$ and $\bigwedge_{\alpha<\lambda}$ for ordinals α and λ.

Formulas and terms of the language $\mathcal{L}_{On,\omega}$ are built up as in first order logic with additional formulas of the type $\bigvee_{\alpha<\kappa}\sigma_\alpha$ and $\bigwedge_{\alpha<\kappa}\sigma_\alpha$ for any set $\{\sigma_\alpha \mid \alpha<\kappa\}$ of formulas.

Definition 157. Let L be a subclass of the language $\mathcal{L}_{On,\omega}$. A *model* for L is a pair $\mathfrak{M}=(A,I)$ where

1. A is a class

2. I is a function that maps the variables, constants, relations symbols and functions of L to elements, relations and functions on A.

For a model $\mathfrak{M}=(A,I)$ of L and a symbol R of L we also write $R^{\mathfrak{M}}$ for $I(R)$. The interpretation of terms of the language L and the satisfaction relation $\models$ is defined as in the case of finitary logic with the additional stipulation that satisfaction of infinitary formulas is defined such that formulas of the form $\bigwedge_{\alpha<\kappa}\sigma_\alpha$ are satisfied if for all $\alpha<\kappa$ the formula σ_α is satisfied, while the formula $\bigvee_{\alpha<\kappa}\sigma_\alpha$ is satisfied if there exists an ordinal $\alpha<\kappa$ such that σ_α is satisfied.

Definition 158. The *weak compactness principle for On* is the assumption that every class of $\mathcal{L}_{On,\omega}$ sentences whose subsets all have models, has itself a model.

Theorem 159. *König's Lemma on On,* KL^+*, is provable from* NBG *together with the weak compactness principle.*

Proof. We assume the weak compactness principle. Let T be a proper class and a set branching tree. We consider the $\mathcal{L}_{On,\omega}$ language with a constant symbol for every element of T and a unary predicate symbol P. We consider the set $\Sigma=\Sigma_1\cup\Sigma_2$ where

$$\Sigma_1=\{\neg(P(f)\wedge P(g)) \mid f,g\in Seq\wedge(f\perp g)\}$$

and

$$\Sigma_2 = \{ \bigvee_{f \in Lv(T,\alpha)} P(f) \mid \alpha \in On\}.$$

Whenever x is a subset of Σ, we can find an ordinal λ such that whenever the sentence $\bigvee_{f \in Lv(T,\alpha)} P(f)$ is an element of x, then $\alpha < \lambda$. Now fix a subset x of Σ and an ordinal λ as described. Since in a set branching tree all the levels are sets, and since T is a proper class, the height of T is On. Thus, we can find a sequence $f \in T$ with $dom(f) = \lambda$. We claim that if we fix $I(P) = \{f \restriction \alpha \mid \alpha < \lambda\}$ and $I(f) = f$ for all elements of T, then (T, I) is a model of x. Since all elements of $I(P)$ are pairwise compatible, we know that every element of Σ_1 is satisfied in (T, I). On the other hand, if an element of x is not in Σ_1, then it is of the form $\bigvee_{f \in Lv(T,\alpha)} P(f)$ for some $\alpha < \lambda$, which is satisfied because $f \restriction \alpha \in I(P)$ and $f \restriction \alpha \in Lv(T, \alpha)$. Applying the weak compactness principle, we conclude that the class Σ has itself a model $\mathfrak{M}$. We find a path through T by stipulating $F = \cup\{f \in T \mid P^{\mathfrak{M}}(f^{\mathfrak{M}})\}$. Since $\mathfrak{M}$ is a model of Σ_1, we know that F is a function, thus, to see that F is a path through T, we have to verify that $dom(F) = On$. For every ordinal α, we know that $P^{\mathfrak{M}}(f^{\mathfrak{M}})$ holds for some $f \in Lv(T, \alpha)$ and therefore that $f \subset F$, i.e. $dom(F) \geq \alpha$. $\square$

Definition 160. We define the relativization φ^U of a formula φ (not containing U as a variable) to any class U by induction on the complexity of φ.

1. Quantifier free formulas remain unchanged when relativized to U.
2. Formulas of the form $\varphi_1 \wedge \varphi_2$ relativize to $\varphi_1^U \wedge \varphi_2^U$.
3. Formulas of the form $\varphi_1 \vee \varphi_2$ relativize to $\varphi_1^U \vee \varphi_2^U$.
4. Formulas of the form $\varphi_1 \rightarrow \varphi_2$ relativize to $\varphi_1^U \rightarrow \varphi_2^U$.
5. Formulas of the form $\neg\varphi$ relativize to $\neg(\varphi^U)$.
6. Formulas of the form $\forall x\, \varphi$ and $\exists x\, \varphi$ relativize to $\forall x\, (x \in U \rightarrow \varphi^U)$ and $\exists x\, (x \in U \wedge \varphi^U)$ respectively.
7. Formulas of the form $\forall X\, \varphi$ and $\exists X\, \varphi$ relativize to $\forall X\, (X \subset U \rightarrow \varphi^U)$ and $\exists X\, (X \subset U \wedge \varphi^U)$ respectively.

Definition 161. Formulas of the form $\forall X\,\sigma(X)$, where $\sigma(X)$ is a Σ^0_1 formula are called *strict* Π^1_1 *formulas*. We will use the term $s\Pi^1_1$ to denote the collection of strict Π^1_1 formulas.

Definition 162. Let $\mathcal{F}$ be a collection of formulas. The $\mathcal{F}$ reflection principle $\mathcal{F}$-Ref is the assumption that for every formula $\psi(X_1,\ldots,X_n,x_1,\ldots,x_k)$ in $\mathcal{F}$ and all sets $\vec{x}=x_1,\ldots,x_k$ and all classes $X_1,\ldots,X_n$, it is the case that

$$\psi(X_1,\ldots,X_n,\vec{x})\rightarrow\exists\alpha\,(\mathsf{lim}(\alpha)\;\wedge\;\vec{x}\in V_\alpha\wedge\psi^{V_\alpha}(X_1\cap V_\alpha,\ldots,X_n\cap V_\alpha,\vec{x})).$$

For a more detailed analysis of the theories Π^1_1-Ref and $s\Pi^1_1$-Ref and their comparison respectively, the reader is referred to 3.4 and 3.5 in [Sal05] and to [Glo76].

Proposition 163. *In the theory* $\mathsf{NBG}+s\Pi^1_1$-Ref*, it is provable that for any binary tree T we have that*

$$T\textit{ has a path }\Leftrightarrow\forall\alpha\,(Lv(T,\alpha)\neq\varnothing).$$

Proof. We work in the theory $\mathsf{NBG}+s\Pi^1_1$-Ref and prove that if T is a binary tree with no path, then there must exist an ordinal α such that $Lv(T,\alpha)=\varnothing$. Let $\psi(T)$ be the statement that T has no path, i.e.

$$\psi(T)\;\equiv\;\forall F\,\exists\alpha\,(F\restriction\alpha\notin T).$$

Since the clauses $\mathsf{fun}(f)$, $\mathsf{ord}(x)$ and $x=dom(y)$ can be expressed by Δ^0_0 formulas, the statement $\psi(T)$ can be reformulated as a $s\Pi^1_1$ formula as follows:

$$\forall F\,\exists g,\alpha\,(g\subset F\wedge dom(g)=\alpha\wedge g\notin T).$$

If T is a binary tree with no path, then we can apply the reflection principle to obtain the existence of a nonzero limit ordinal λ such that $\psi^{V_\lambda}(T\cap V_\lambda)$. Considering that T is a binary tree, we obtain that for any element f of T that

$$dom(f)<\lambda\Leftrightarrow f\in V_\lambda$$

and thus that

$$T \cap V_\lambda = \bigcup_{\alpha<\lambda} Lv(T, \alpha).$$

Hence, we conclude that

$$\begin{aligned}
\psi^{V_\lambda}(T \cap V_\lambda) &\Leftrightarrow \big(\forall F\, \exists \alpha\, (F \restriction \alpha \notin T \cap V_\lambda)\big)^{V_\lambda} \\
&\Leftrightarrow \forall f \in V_{\lambda+1}\, \exists \alpha < \lambda\, (f \restriction \alpha \notin \bigcup_{\mu<\lambda} Lv(T, \mu)) \\
&\Leftrightarrow \forall f \in {}^{\lambda}2\, \exists \alpha < \lambda\, (f \restriction \alpha \notin T)
\end{aligned}$$

and thus that $Lv(T, \lambda) = \varnothing$. For the converse direction, note that if F is a path though a tree T, then $F \restriction \alpha \in Lv(T, \alpha) \neq \varnothing$ for all ordinals α. □

Theorem 164. KL^+ *is provable in the theory* $\mathsf{NBG} + s\Pi^1_1\text{-}\mathsf{Ref}$.

Proof. In view of Theorem 154, it is enough to prove WKL^+ within the theory $\mathsf{NBG} + s\Pi^1_1\text{-}\mathsf{Ref}$. Let $T \subset Bin$ be a binary tree. If T is a proper class, then all its levels must be nonempty. As we have seen in Proposition 163, from this it is provable in $\mathsf{NBG} + s\Pi^1_1\text{-}\mathsf{Ref}$ that T has a path. □

Summary

We introduced the notion of κ well orders together with the principle of κ-induction and the κ-minimum principle. In Proposition 128 we obtained a characterization of regular cardinals in terms of κ-well orders and the κ-chain condition. In the second part of the past section, we fixed the notion of trees in set theory. In Example 139, we illustrated that in set theory, as opposed to arithmetic, it is much more complicated distinguish trees that have a path from those that do not. In line with this finding, we observed in Remark 155 that König's Lemma (for On) is a much stronger assertion in set theory than it is in arithmetic. Subsequently, we reduced König's Lemma to its weak form over NBG (cf. Theorem 154). This preparation enabled us to adopt the standard arguments (e.g. Theorem II.17.18 and Theorem II.17.13 in [Jec03]) in order to give upper bounds for the strength of KL^+ in terms of strict Π^1_1 reflection on limit ordinals (cf. Theorem 164) and the weak compactness principle (cf. Theorem 159).

3 Weak well orders

Weak well orders are introduced to be the class theoretic counterpart of countable well orders in arithmetic. While weak well orders are a natural translation of well orders with respect to our doctrine of translating numbers as sets and (countable) sets as proper classes, they also exhibit similar technical features as well orders in arithmetic. The most striking resemblance is a set theoretic analog of Kleene's Σ^1_1 normal form lemma (cf. Theorem 172). Following our exposition of the normal form lemma, it is investigated to what extent this result enables us to preserve some of the "landscape" of arithmetic to our setting of sets and classes.

Following our framework of the first chapter, a weak well order is an "On-well order". In the same fashion as before, we introduce weak well orders by providing three pairwise equivalent conditions: an induction principle, a descending chain condition and a minimum principle respectively. Also, the notion of a superprogressive subclass of linear orders is based on the respective notion fixed for κ-well orders before.

Before we turn our attention to investigate weak well orders, let us briefly recall the situation as it presents itself in arithmetic. The notion of a progressive subset X of a (countable) linear ordering $(A, <_A)$ is given from the condition

$$\forall a \in A \, (A_{<_A a} \subset X \to a \in X).$$

This condition, however, is equivalent to saying that for every nonempty finite subset m of A, the statement

$$A_{<_A m} \subset X \to X \cap m \neq \varnothing$$

is satisfied. This is the formulation that we will translate to our setting in order to obtain the notion of a *superprogressive* class and consequently the concept of a weak well order.

Definition 165. If $(A, <_A)$ is a linear order and X is a subclass of A, then we call X a *superprogressive* subclass of A if, for all nonempty subsets m of A,

$$A_{<_A m} \subset X \to X \cap m \neq \varnothing$$

holds. We write $\mathsf{prog}^+(X, (A, <_A))$ to mean that X is a superprogressive subclass of the linear order $(A, <_A)$.

Remark 166. A subclass $X \subset A$ is exactly then superprogressive in $(A, <_A)$ if X is κ-progressive for all ordinals κ.

Lemma 167. *Let $(A, <_A)$ be a linear order. The following properties are all equivalent for $(A, <_A)$.*

1. *(Weak induction principle) No proper and superprogressive subclass of A exists.*
2. *(Weak chain condition) No strictly descending function $F : On \to A$ exists.*
3. *(Weak minimum principle) For every subclass $X \subset A$, there exists a set $m \subset X$ that is minimal in X.*

Proof. Let $(A, <_A)$ be a linear order.

1⇒2: By contrapositive, assume that a strictly decreasing function $F : On \to A$ exists. Use elementary comprehension to define $X = A \setminus rng(F)$. Let $m \subset A$ be a nonempty set such that $A_{<_A m} \subset X$. Note that F is one-to-one; hence, the inverse image $F^{-1}(m) = \{\alpha \mid F(\alpha) \in m\}$ of m must be bounded in On by an ordinal λ. If $X \cap m$ was empty, then $m \subset rng(F)$; hence, $F[F^{-1}(m)] = m$. Therefore, if we pick $\beta > \lambda$, it follows that $F(\beta) <_A m$ in contradiction to the choice of m. This proves that $X \cap m$ cannot be empty and thus that X is superprogressive, i.e. we have found a superprogressive subclass of A that is clearly a proper subclass.

2⇒3: By contrapositive, let $X \subset A$ be such that there exists no minimal subset in X. Note that for every subset $m \subset X$, this implies that $X_{<_A m} \neq \varnothing$. Define a strictly descending function $F : On \to A$ by recursion on the ordinals as follows:

$$F(\lambda) = \min_{\lhd}(X_{<_A \{F(\mu) \mid \mu < \lambda\}}).$$

3⇒1: By contrapositive, assume that $X \subsetneq A$ is superprogressive. If $A \setminus X$ has a minimal subset m, then $A_{<_A m} \subset X$, and because $X \neq A$, also $m \neq \varnothing$. From the assumption that X is superprogressive it follows that $X \cap m \neq \varnothing$ contradicting the choice of m. Hence, $A \setminus X$ has no minimal subset and therefore the third condition is violated. □

Definition 168. A linear order that satisfies any, and thus all, of the properties presented in Lemma 167, is called a *weak well order*. We write $\mathsf{wwo}(A, <_A)$ to mean that $(A, <_A)$ is a weak well order.

Remark 169. The definition of progressive subsets cannot be altered from

$$A_{<_A m} \subset X \to m \cap X \neq \varnothing$$

for all nonempty finite sets m, to

$$A_{<_A m} \subset X \to m \subset X$$

for the same range of sets m. Otherwise, any linear ordering with some least element, for example $\omega + 1$ with the natural ordering inverted, would satisfy transfinite induction, i.e. the nonexistence of proper yet progressive subsets. Similarly, the weak induction principle is not (and should not be) equivalent to

$$\forall X \left(\forall m \subset A \left(A_{<_A m} \subset X \to m \subset X\right) \to A \subset X\right), \tag{3.1}$$

since every linear ordering that contains a minimal subset (of the whole field) satisfies already equation (3.1).

Proof. Let $(A, <_A)$ be any linear order with minimal subset m_0. Let X be a subclass of A that satisfies the requirement

$$\forall m \subset A \left(A_{<_A m} \subset X \to m \subset X\right).$$

Now pick any $a \in A$ and consider the set $m_1 = m_0 \cup \{a\}$. Since m_0 is minimal in A, so is m_1 and we have that $A_{<_A m_1} = \varnothing \subset X$. Therefore, it follows from our assumption on X that $m_1 \subset X$ and hence $a \in X$. □

Lemma 170. *If $<_A$ is a linear ordering of A such that every coinitial subclass X of A contains a minimal subset, then $(A, <_A)$ is already a weak well order.*

Proof. The proof is very similar as in the case of κ-well orders. Let $(A, <_A)$ be a linear order as in the statement and let X be any proper and nonempty subclass of A. Since $A_{<_A X}$ is initial in A, by assumption, there must exist a set m that is minimal in $Y = A \setminus A_{<_A X} = \{a \in A \mid \exists x \in X\,(x \leq_A a)\}$. Assume that $m = \{m_\alpha \mid \alpha < \lambda\}$. We pick any $a_0 \in X$ and define for each ordinal $\alpha < \lambda$ the set x_α as

$$x_\alpha = \begin{cases} \min_\lhd(X_{\leq m_\alpha}) & \text{if } X_{\leq m_\alpha} \neq \varnothing \\ a_0 & \text{else} \end{cases}$$

and set $x = \{x_\alpha \mid \alpha < \lambda\} \subset X$. To see that x is a minimal subset of X, let $a \in X$ be arbitrary. Since a is an element of Y, there exists an ordinal $\alpha < \lambda$ such that $m_\alpha \leq a$. Since $m_\alpha \in Y_{\leq m_\alpha}$, it follows from the minimality of X in Y that there exists a $b \in Y_{\leq m_\alpha} \cap X = X_{\leq m_\alpha}$. Hence, $x_\alpha \leq m_\alpha \leq a$. □

3.1 The logical complexity of weak well orders

It is one of the fundamental results in arithmetic that in the subsystem ACA_0 of second order arithmetic the predicate $\mathsf{wo}(X)$, which says that X is a well ordering of the integers, is provably a Π^1_1 complete predicate (e.g. Theorem II.1.8 in [Pro05] or Lemma V.1.8 in [Sim98]). In particular, this implies that it is not expressible by any Σ^1_1 formula that a given linear ordering is well founded (cf. Theorem 7 and Theorem II.1.9 in [Pro05] and Theorem V.1.9 in [Sim98]). This fact is of pivotal importance for many of the results that have been obtained from the comparison of meta predicative subsystems of second order arithmetic. In particular, it is the rudiment of any argument that involves pseudo hierarchies.

The proof in arithmetic that $\mathsf{wo}(X)$ is a Π^1_1 complete predicate is based on the fact that a tree is well founded if and only if it is well ordered by its Kleene Brouwer ordering. In order to show this, one can[1] use König's Lemma (e.g. Lemma V.1.3 in [Sim98]) which is at disposal in ACA_0. In set theory, as we have

[1] However, there is no urge to use König's Lemma since the proof presented here works also in arithmetic.

observed before, the situation regarding König's Lemma is quite different and it is far from being at our disposal in the base theory. However, we can prove the equivalence of a tree not having a path and of it being weakly well ordered by its Kleene Brouwer ordering by applying our weak minimum principle. We can therefore avoid the use of König's Lemma and, as a consequence, we will be able to prove the Π^1_1 completeness of $\mathsf{wwo}(X)$ in the theory NBG.

Proposition 171. *It is provable in* NBG *that for every tree T, the equivalence*

$$T \textit{ has a path } \Leftrightarrow KB_T \textit{ is not a weak well ordering of } T$$

holds.

Proof. The proof is very similar to the proof of Proposition 142, however, since being a weak well order does not imply being a κ-well order for some κ, we cannot directly apply the proposition. First assume that T has no path. Let X be any coinitial subclass and thus a subtree of T. We prove that X contains a minimal subset, which in view of Lemma 170 is sufficient to prove that the Kleene Brouwer ordering of T is a weak well order. We apply Corollary 74 to the formula

$$\varphi(f,x) \equiv x = \min_{\lhd}\{y \mid f \cup \langle dom(f), y\rangle \in X\}$$

to obtain a function $F : On \to V$ that maps any α to the $\lhd$-least set so that $F \restriction \alpha + 1 \in X$ holds, whenever α is such that there exists any set x with the property that $F \restriction \alpha \frown x \in X$. Since T has no path, neither does X and we can define

$$\lambda = \min\{\alpha \mid F \restriction \alpha + 1 \notin X\}.$$

In case that $F \restriction \lambda \in X$ it is clear that $F \restriction \lambda$ is the least element of X with respect to the Kleene Brouwer ordering of T and we are done. If $F \restriction \lambda \notin X$, then λ must clearly be a limit ordinal. We prove that $m = \{F \restriction \alpha \mid \alpha < \lambda\}$ is minimal in X. Since X is a subtree of T, it is clear that $m \subset X$. Let $h \in X$ be arbitrarily chosen. Since X is a tree and $\cup m = F \restriction \lambda \notin X$, there must be an ordinal $\alpha < \lambda$ such that $F \restriction \alpha \not\subset h$. If $h \subset F \restriction \alpha$, then $F \restriction \alpha \leq_T h$ and we are done. Otherwise let β be the least ordinal such that $F \restriction \alpha(\beta) \neq h(\beta)$; from the definition of F, it follows that $F \restriction \alpha(\beta) \lhd h(\beta)$ and we are done. The converse direction is straight forward as every path induces a descending chain in (T, KB_T). □

Theorem 172 (Σ^1_1 normal form lemma)**.** *Fix a bijective function $\Omega : V \to On$ and define for all ordinals α the set $\Omega_\alpha = \{x \mid \Omega(x) < \alpha\}$, further write X_α to mean $X \cap \Omega_\alpha$. If $\varphi(\vec{X}, \vec{z})$ is a Σ^1_1 formula (with all free variables shown), then we can find a $\Delta^1_0(\Omega)$ formula, that is an elementary formula that contains the additional constant Ω, $\varphi^*(\vec{X}, f, \vec{z})$ such that*

$$\forall \vec{X}, \vec{z} \left(\varphi(\vec{X}, \vec{z}) \leftrightarrow \exists F \, \forall \alpha \, \varphi^*(\vec{X}_\alpha, F \restriction \alpha, \vec{z})\right)$$

is provable in NBG.

Proof. We assume for the first step of our proof that φ is an elementary formula (in prenex normal form with possibly a "dummy" universal quantifier), that is φ takes the form

$$\varphi(\vec{X}, \vec{z}) \equiv \forall x_1 \exists y_1 \ldots \forall x_k \exists y_k \, \varphi_0(\vec{X}, x_1, y_1, \ldots, x_k, y_k, \vec{z})$$

where φ_0 is quantifier free. Since there are no quantifiers involved in the formula $\varphi_0(\vec{X}, x_1, y_1, \ldots, x_k, y_k, \vec{z})$, we know that

$$\begin{aligned} &\forall \alpha \, \forall x_1, y_1, \ldots, x_k, y_k, \vec{z} \in \Omega_\alpha \\ &\quad \left(\varphi_0(\vec{X}, x_1, y_1, \ldots, x_k, y_k, \vec{z}) \leftrightarrow \varphi_0(\vec{X}_\alpha, x_1, y_1, \ldots, x_k, y_k, \vec{z})\right) \end{aligned} \tag{3.2}$$

holds for all classes $\vec{X}$ and all sets $\vec{z}$. Because φ is elementary, we can use global choice obtain the following equivalences for all classes $\vec{X}$ and all sets $\vec{z}$:

$$\begin{aligned} &\forall x_1 \exists y_1 \ldots \forall x_k \exists y_k \, \varphi_0\big(\vec{X}, x_1, y_1, \ldots, x_k, y_k, \vec{z}\big) \\ &\Leftrightarrow \exists F \, \forall x_1, \ldots, x_k \, \varphi_0\big(\vec{X}, x_1, F(\langle x_1 \rangle), \ldots, x_k, F(\langle x_1, \ldots, x_k \rangle), \vec{z}\big) \\ &\Leftrightarrow \exists F \, \forall x_1, \ldots, x_k \, \varphi_0\big(\vec{X}, x_1, F(\Omega(\langle x_1 \rangle)), \ldots, x_k, F(\Omega(\langle x_1, \ldots, x_k \rangle)), \vec{z}\big). \end{aligned} \tag{3.3}$$

If we write $\Lambda(\alpha, x_1, \ldots, x_k)$ to mean

$$\begin{aligned} &x_1, \ldots, x_k, \langle x_1 \rangle, \ldots, \langle x_1, \ldots, x_k \rangle \in \Omega_\alpha \\ &\wedge \; F(\Omega(\langle x_1 \rangle)), \ldots, F(\Omega(\langle x_1, \ldots, x_k \rangle)) \in \Omega_\alpha \end{aligned}$$

we can use equation (3.2) to continue from (3.3) as follows:

$$\exists F\,\forall\alpha\,\forall x_1,\ldots,x_k\,\big(\Lambda(\alpha,x_1,\ldots,x_k)\\ \to \varphi_0(\vec{X},x_1,F(\Omega(\langle x_1\rangle)),\ldots,x_k,F(\Omega(\langle x_1,\ldots,x_k\rangle)),\vec{z})\big)$$

which is equivalent to

$$\exists F\,\forall\alpha\,\forall x_1,\ldots,x_k\,\big(\Lambda(\alpha,x_1,\ldots,x_k)\\ \to \varphi_0(\vec{X}_\alpha,x_1,F\restriction\alpha(\Omega(\langle x_1\rangle)),\ldots,x_k,F\restriction\alpha(\Omega(\langle x_1,\ldots,x_k\rangle)),\vec{z})\big).$$

Thus, the formula

$$\Theta(\vec{X},F,\vec{z},\alpha)\;\equiv\;\forall x_1,\ldots,x_k\,\big(\Lambda(\alpha,x_1,\ldots,x_k)\;\to\;\Delta(\vec{X},F,\vec{z},\alpha,x_1\ldots,x_k)\big)$$

where $\Delta(\vec{X},F,\vec{z},\alpha,x_1\ldots,x_k)$ stands for

$$\varphi_0(\vec{X},x_1,F(\Omega(\langle x_1\rangle)),\ldots,x_k,F(\Omega(\langle x_1,\ldots,x_k\rangle)),\vec{z}),$$

satisfies

$$\varphi(\vec{X},\vec{z})\;\leftrightarrow\;\exists F\,\forall\alpha\,\Theta(\vec{X}_\alpha,F\restriction\alpha,\vec{z},\alpha)$$

for all classes $\vec{X}$ and all sets $\vec{z}$. Therefore, we can stipulate

$$\varphi^*(\vec{X},f,\vec{z})\;\equiv\;\Theta(\vec{X},f,\vec{z},rk(dom(f)))$$

to obtain the desired formula. We now turn to the case where $\varphi(\vec{X},\vec{z})$ is of the form

$$\varphi(\vec{X},\vec{z})\equiv\exists Y\,\psi(\vec{X},Y,\vec{z})$$

with $\psi(\vec{X},Y,\vec{z})$ elementary. From the first case we get an elementary formula $\psi^*(\vec{X},Y,f,\vec{z})$ such that

$$\forall\vec{X},Y,\vec{z}\,\big(\psi(\vec{X},Y,\vec{z})\leftrightarrow\exists F\,\forall\alpha\,\psi^*(\vec{X}_\alpha,Y_\alpha,F\restriction\alpha,\vec{z})\big) \tag{3.4}$$

holds. We obtain $\varphi^*(\vec{X},g,\vec{z})$ from $\psi^*(\vec{X},Y,f,\vec{z})$ by replacing all subformulas that are of the form $z\in Y$ (z being any set variable) by the fragment

$\exists v\,(\langle\Omega(z),\langle 1,v\rangle\rangle \in g)$ (with a new variable v), and all subformulas of the form $z \in f$ by $\exists a, b\,(z = \langle a,b\rangle \wedge (\langle a,\langle 0,b\rangle\rangle \in g \vee \langle a,\langle 1,b\rangle\rangle \in g))$ (with new variables a and b). We will now show that

$$\forall \vec{X}, \vec{z}\,\big(\varphi(\vec{X},\vec{z}) \leftrightarrow \exists G\,\forall\alpha\,\varphi^*(\vec{X}_\alpha, G \restriction \alpha, \vec{z})\big) \tag{3.5}$$

holds, and thus complete the proof of the theorem. Let $F : On \to V$ be any function and Y some arbitrary class. We use elementary comprehension to obtain a function $G_{Y,F} : On \to V$ with

$$G_{Y,F}(\alpha) = \begin{cases} \langle 0, F(\alpha)\rangle & \text{if } \Omega^{-1}(\alpha) \notin Y \\ \langle 1, F(\alpha)\rangle & \text{if } \Omega^{-1}(\alpha) \in Y. \end{cases}$$

We get for all ordinals α that

$$Y_\alpha = \{x \mid \exists v\,(G_{Y,F}(\Omega(x)) = \langle 1,v\rangle)\}_\alpha = \{x \mid \exists v\,(\langle\Omega(x),\langle 1,v\rangle\rangle \in G_{Y,F} \restriction \alpha)\},$$

and for all sets x and y that

$$\langle x,y\rangle \in F \restriction \alpha \Leftrightarrow \exists w\,(\langle x,\langle w,y\rangle\rangle \in G_{Y,F} \restriction \alpha).$$

Therefore, we have that

$$\begin{aligned} &\forall\alpha\,\psi^*(\vec{X}_\alpha, Y_\alpha, F \restriction \alpha, \vec{z}) \\ &\Leftrightarrow \forall\alpha\,\psi^*(\vec{X}_\alpha, \{x \mid \exists v\,(\langle\Omega(x),\langle 1,v\rangle\rangle \in G_{Y,F} \restriction \alpha)\}, F \restriction \alpha, \vec{z}) \\ &\Leftrightarrow \forall\alpha\,\varphi^*(\vec{X}_\alpha, G_{Y,F} \restriction \alpha, \vec{z}). \end{aligned} \tag{3.6}$$

Thus, in view of equations (3.4) and (3.6), we have that

$$\forall \vec{X}, \vec{z}\,\big(\varphi(\vec{X},\vec{z}) \leftrightarrow \exists Y, F\,\forall\alpha\,\varphi^*(\vec{X}_\alpha, G_{Y,F} \restriction \alpha, \vec{z})\big).$$

Therefore, what remains to be shown, is that the equivalence

$$\exists Y, F\,\forall\alpha\,\varphi^*(\vec{X}_\alpha, G_{Y,F} \restriction \alpha, \vec{z}) \leftrightarrow \exists G\,\forall\alpha\,\varphi^*(\vec{X}, G \restriction \alpha, \vec{z}) \tag{3.7}$$

holds for all classes $\vec{X}$ and all sets $\vec{z}$. The direction from left to right is obviously true. For the converse direction, we can simply invert the construction of $G_{Y,F}$ presented before, i.e. given a function G, we can use elementary comprehension to define the function

$$F_G = \{\langle x, y\rangle \mid \langle x, \langle 1, y\rangle\rangle \in G \vee \langle x, \langle 0, y\rangle\rangle \in G\}$$

and the class

$$Y_G = \{x \mid \exists y\, (\langle \Omega(x), \langle 1, y\rangle\rangle \in G)\}.$$

Since for all ordinals α and all sets y, we have that

$$G_{Y_G,F_G}(\alpha) = \langle 1, y\rangle \Leftrightarrow G(\alpha) = \langle 1, y\rangle$$

and

$$G_{Y_G,F_G}(\alpha) = \langle 0, y\rangle \Leftrightarrow G(\alpha) = \langle 0, y\rangle,$$

the desired equivalence (3.7) follows from the definition of $\varphi^*(\vec{X}, g, \vec{z})$. $\square$

Lemma 173 (Representation lemma I)**.** *If $\varphi(\vec{X}, \vec{x})$ is a Π^1_1 formula, then there is a formula $\mathsf{T}_\varphi(\vec{X}, f, \vec{x}) \in \Delta^1_0(\Omega)$ such that it is provable in* NBG *that for all classes $\vec{X}$ and all sets $\vec{x}$ the class $T_\varphi(\vec{X}, \vec{x})$, defined from*

$$f \in T_\varphi(\vec{X}, \vec{x}) \leftrightarrow \mathsf{T}_\varphi(\vec{X}, f, \vec{x}),$$

has the following properties:

1. *The class $T_\varphi(\vec{X}, \vec{x})$ is a tree for all classes $\vec{X}$ and all sets $\vec{x}$.*
2. *For all classes $\vec{X}$ and all sets $\vec{x}$, we have that*
$$\varphi(\vec{X}, \vec{x}) \leftrightarrow T_\varphi(\vec{X}, \vec{x}) \text{ has no path.}$$
3. *For all classes $\vec{X}$ and all sets $\vec{x}$, we have that*
$$\varphi(\vec{X}, \vec{x}) \leftrightarrow \big(T_\varphi(\vec{X}, \vec{x}), KB_{T_\varphi(\vec{X}, \vec{x})}\big) \text{ is a weak well order}$$

Proof. Let $\varphi(\vec{X}, \vec{x})$ be any Π^1_1 formula and further let $\psi(\vec{X}, \vec{x})$ be the Σ^1_1 formula obtained from $\neg\varphi(\vec{X}, \vec{x})$ in negation normal form. Further, $\psi^*(\vec{X}, f, \vec{x})$ be the $\Delta^1_0(\Omega)$ formula that satisfies

$$\forall \vec{X}, \vec{x}\big(\psi(\vec{X}, \vec{x}) \leftrightarrow \exists F\, \forall\alpha\, \psi^*(\vec{X}_\alpha, F \restriction \alpha, \vec{x})\big)$$

as presented in the normal form lemma. We stipulate

$$\mathsf{T}_\varphi(\vec{X}, f, \vec{x}) \;\equiv\; f \in \mathit{Seq} \wedge \forall\alpha \leq \mathit{dom}(f)\, \psi^*(\vec{X}_\alpha, f \restriction \alpha, \vec{x}).$$

We now turn to check the first and the second claim of the theorem.

1. This is obvious from the definition of $\mathsf{T}_\varphi(\vec{X}, \vec{x})$ and the fact that if $f \in \mathit{Seq}$ and $\alpha \leq \mathit{dom}(f)$ then $\{(f \restriction \alpha) \restriction \mu \mid \mu \in \mathit{On}\} \subset \{f \restriction \mu \mid \mu \in \mathit{On}\}$.

2. Since

$$\begin{aligned}
T_\varphi(\vec{X}, \vec{x}) \text{ has no path } &\Leftrightarrow \forall F\, \exists\alpha\, (F \restriction \alpha \notin T_\varphi(\vec{X}, \vec{x}))\\
&\Leftrightarrow \forall F\, \exists\alpha\, \neg \mathsf{T}_\varphi(\vec{X}, F \restriction \alpha, \vec{x})\\
&\Leftrightarrow \forall F\, \exists\alpha\, \neg\forall\beta \leq \alpha\, \psi^*(\vec{X}_\beta, F \restriction \beta, \vec{x})\\
&\Leftrightarrow \forall F\, \exists\alpha \neg\psi^*(\vec{X}_\alpha, F \restriction \alpha, \vec{x})\\
&\Leftrightarrow \neg\exists F\, \forall\alpha\, \psi^*(\vec{X}_\alpha, F \restriction \alpha, \vec{x})\\
&\Leftrightarrow \neg\psi(\vec{X}, \vec{x}),
\end{aligned}$$

   the claim follows from the equivalence $\neg\psi(\vec{X}, \vec{x}) \Leftrightarrow \varphi(\vec{X}, \vec{x})$.

Because the third claim is a direct consequence of the second claim and Proposition 171, the proof is complete. $\square$

Definition 174. Let T be a tree, X any class and F either an element of Seq or a function whose domain is the class On.

1. We write $[X, F]$ to mean the mapping $\mathit{dom}(F) \ni \alpha \mapsto \langle X_\alpha, F(\alpha)\rangle \in V \times V$.

2. Given a tree S with $\bigcup_{f \in S} \mathit{rng}(f) \subset V \times V$, we use elementary comprehen-

sion to form a new tree

$$T_S(X) = \{f \in Seq \mid [X, f] \in S\}.$$

Further, we write $X \epsilon S$ to mean that a function $F : On \to V$ so that $[X, F]$ is a path through S exists.

Remark 175. Let T, S, X be as before, then $T_S(X)$ has a path if and only if $X \epsilon S$.

Lemma 176 (Representation lemma II)**.** *Let $\varphi(X)$ be a Σ^1_1 formula. It is provable in* NBG *that there is a tree S with $\bigcup_{f \in S} rng(f) \subset V \times V$ such that*

$$\forall X \, (\varphi(X) \leftrightarrow X \epsilon S).$$

Proof. Let $\varphi(X)$ be a Σ^1_1 formula (with no further parameters). Pick $\varphi^*(X, f)$ as in Theorem 172. Hence,

$$\forall X \left(\varphi(X) \leftrightarrow \exists F \, \forall \alpha \, (\varphi^*(X_\alpha, F \restriction \alpha))\right).$$

We use elementary comprehension to form the tree

$$S = \{f \otimes g \mid f, g \in Seq \land \forall \alpha < dom(f \otimes g) \, \varphi^*(f(\alpha), g \restriction \alpha)\}$$

where for $f, g \in Seq$, the term $f \otimes g$ stands for the mapping $\alpha \mapsto \langle f(\alpha), g(\alpha) \rangle$ and $dom(f \otimes g) = \min\{dom(f), dom(g)\}$. Obviously, S is a tree with $\bigcup_{f \in S} rng(f) \subset V \times V$. It remains to show that

$$\forall X \left(\varphi(X) \leftrightarrow T_S(X) \text{ has a path}\right)$$

holds. Fix any class X with $\varphi(X)$ and a function F such that

$$\forall \alpha \, \varphi^*(X_\alpha, F \restriction \alpha).$$

Hence, the sequence $[X, F] \restriction \alpha$ belongs to S for every ordinal α and thus $T_S(X)$

has a path. For the converse direction, consider

$$\begin{aligned}
T_S(X) \text{ has a path} &\Leftrightarrow \exists F\, \forall\alpha\, (F \restriction \alpha \in T_S(X)) \\
&\Rightarrow \exists F\, \forall\alpha\, ([X, F] \restriction \alpha \in S) \\
&\Rightarrow \exists F\, \forall\alpha\, ((\mu \mapsto X_\mu) \restriction \alpha \otimes F \restriction \alpha \in S) \\
&\Rightarrow \exists F\, \forall\alpha\, \forall\mu < \alpha\, (\varphi^*(X_\mu, F \restriction \mu)) \\
&\Rightarrow \exists F\, \forall\alpha\, \varphi^*(X_\alpha, F \restriction \alpha) \\
&\Rightarrow \varphi(X).
\end{aligned}$$

This proves the claim. □

Proposition 177. *If $\varphi(X)$ is a Σ^1_1 formula, then it is provable in* NBG *that there is a tree T such that*

$$\neg\, (\varphi(T) \leftrightarrow T \textit{ has no path})$$

holds.

Proof. By way of contradiction, assume that φ is a Σ^1_1 formula such that

$$\forall X\, \big(\mathsf{tree}(X) \to (\varphi(X) \leftrightarrow X \text{ has no path})\big)$$

holds. Consider a negated Σ^1_1 formula $\psi(X)$ that expresses

$$\mathsf{tree}(X) \wedge \bigcup_{f \in X} rng(f) \subset V \times V \wedge \neg\varphi(T_X(X)).$$

From the second representation lemma, we get a tree S such that

$$\bigcup_{f \in S} rng(f) \subset V \times V$$

and

$$\forall X\, \big(\psi(X) \leftrightarrow T_S(X) \text{ has no path}).$$

From

$$\psi(S) \Leftrightarrow \neg\varphi(T_S(S)),$$

we get

$$\psi(S) \Leftrightarrow T_S(S) \text{ has a path},$$

but we already know that also

$$\psi(S) \Leftrightarrow T_S(S) \text{ has no path}$$

holds, which is a contradiction. □

Theorem 178. *If $\varphi(A, <_A)$ is a Σ^1_1 formula, then it is provable in* NBG *that*

$$\neg\forall(A, <_A)\left(\varphi(A, <_A) \leftrightarrow \mathsf{wwo}(A, <)\right)$$

holds.

Proof. By way of contradiction, assume that a formula $\varphi(A, <_A)$ exists such that

$$\forall(A, <_A)\,(\varphi(A, <_A) \;\leftrightarrow\; \mathsf{wwo}(A, <_A))$$

holds. Hence, we obtain from Proposition 171 that for every tree T

$$\varphi(T, KB_T) \;\leftrightarrow\; T \text{ has no path},$$

in contradiction to Proposition 177. □

Summary

We used the minimum principle for weak well orders to prove that the Kleene Brouwer ordering of a tree T is exactly then weakly well founded if T has no path (cf. Proposition 171). This fact enabled us to prove the Σ^1_1 normal form lemma (cf. Theorem 172). From the Σ^1_1 normal for Lemma we then deduced that the predicate $\mathsf{wwo}(A, <_A)$ is Π^1_1 complete in the sense that every Π^1_1 assertion about a class can be stated in the form that a specific linear order is a weak well order (cf. Lemma 173). As a consequence, we proved that the predicate $\mathsf{wwo}(A, <_A)$ cannot be stated as a Σ^1_1 formula (cf. Theorem 178).

3.2 Elementary transfinite recursion

Recursion principles have undergone a long history of research, and they play a key role in many academic disciplines. Similar principles are used and investigated in mathematics, logic, formal languages, linguistics and computer science – to name but a few.

The starting point of the investigation of the theory ATR_0 is the seminal work [Fri74] of H. Friedman. In [Fri74], second order arithmetic is enriched by a principle of transfinitely iterating arithmetical comprehension along any given well ordering. Owing to the important role of the resulting theory, ATR_0, in reverse mathematics, the book [Sim98] lists[2] the theory as one of the "five basic systems" of second order arithmetic (cf. [Sim98] I.12).

In this section, we discuss what theories of sets and classes are suitable representatives of set theoretic counterparts of the theory ATR_0. Here, "suitable" is to be understood in a twofold manner. The underlying meaning of the original theory should be preserved, i.e. the theory should be recognized as being based on the same principle of iterating elementary comprehension along some (carefully chosen) relations. Second, the new theory should relate to the set theoretic interpretation of other subsystems of second order arithmetic in a similar way as the theories originally relate to each other. Since the theory ATR_0 can be quoted verbatim in set theory, the first requirement is not difficult to meet (cf. Definition 188). However, since the theory ATR_0 ows much of its capabilities to the fact that in arithmetic the well foundedness predicate is Π^1_1 complete, the latter requirement is not fulfilled by employing a one-to-one translation of ATR_0. Consequently, we try to find a way to take advantage of the Σ^1_1 normal form lemma and thus suggest a formulation that involves weak well orders (cf. Definition 196). However, simply allowing iterations of elementary comprehension along weak well orders, or along any non well founded structures for that matter, immediately results in inconsistent theories (cf. the follow-up discussion and in particular Theorem 185, or our related work presented in [FS13]). Under these circumstances, we evidently have to be very careful in combining recursion and weak well orders. What we eventually gain by reformulating elementary transfi-

[2] Among the theories $\mathsf{RCA}_0, \mathsf{WKL}_0, \mathsf{ACA}_0$ and $\Pi^1_1\text{-}\mathsf{CA}_0$.

nite recursion with weak well orders as opposed to well orders, is preserving the logical structure of the axiom in being formed as a Σ^1_1 property of some linear order $(A, <_A)$ being entailed by a Π^1_1 complete predicate of $(A, <_A)$. We will thus be able to further extend the theory of elementary transfinite recursion by a principle of "uniform recursion" (cf. Definition 219) that when translated to the situation in arithmetic trivially holds, but in our setting allows for "pseudo hierarchy like arguments" to be applied. We will demonstrate this ability of our theory by proving that fixed points of positive inductive definitions exist, something that cannot be shown with the usual principle of elementary transfinite recursion.

Definition 179. Let R be a binary relation on a class A. A *progressive subclass* X of (A, R) is a class $X \subset A$ such that

$$\forall b \in A \ (\{a \in A \mid aRb\} \subset X \ \rightarrow \ b \in X)$$

holds. We will write $\mathsf{prog}(X, (A, R))$ to mean that X is a progressive subclass of (A, R).

Remark 180. Note that in case of linear orders, Definition 179 and Definition 118 coincide.

Definition 181. A binary relation R on a class A is said to be *well founded* (on A) if one of the following three, pairwise equivalent, statements hold

1. There is no function $f : \omega \to A$ such that $\forall n \in \omega \, (f(n+1)Rf(n))$.
2. Every nonempty subclass of A contains R-minimal elements.
3. $\forall X \ (\mathsf{prog}(X, (A, R)) \ \rightarrow \ A \subset X)$.

We write $\mathsf{wf}(A, R)$ to express that R is well founded and transitive on A. Similarly, we write $\mathsf{wf}(R)$ to mean $\mathsf{wf}(Fld(R), R)$.

Remark 182. Clearly, $\mathsf{wo}(A, R)$ if $\mathsf{wf}(A, R)$ and R is a linear ordering of A.

Definition 183. The transitive closure of a relation R on a class A is the pair (A, R^*) where xR^*y holds if there is a natural number $0 < n$ and a function

$f \in \mathit{Seq}$ with $\mathit{dom}(f) = n$ such that

$$f(0) = x \wedge f(\mathit{dom}(f) - 1) = y \wedge \forall i < \mathit{dom}(f)\,(i = 0 \vee f(i-1) R f(i)).$$

Remark 184. A relation R is well founded if and only if the transitive closure of R is well founded.

Theorem 185. *There is an elementary formula $\varphi(X, x, y)$ such that it is provable in* NBG *that for every transitive relation $(A, <_A)$,*

$$(\forall x\, \exists H\, \forall a \in A\, ((H)_a \;=\; \{y \mid \varphi(H \restriction A_{<_A a}, x, y)\})) \;\rightarrow\; \mathsf{wf}(A, <_A).$$

Proof. Let $\varphi(X, x, y)$ be the elementary formula

$$y = \varnothing \;\wedge\; \forall a \in x\, ((H)_a = \varnothing).$$

In particular, note that

$$\{y \mid \varphi(H, x, y)\} = \begin{cases} 1 & \text{if } \forall a \in x\, ((H)_a = \varnothing) \\ \varnothing & \text{otherwise} \end{cases} \tag{3.8}$$

holds of any class H. Now assume that $<_A$ is a transitive relation on A such that for any set x, there is a class H such that for all $a \in A$

$$(H)_a = \{y \mid \varphi(H \restriction A_{<_A a}, x, y)\} \tag{3.9}$$

holds. Pick any nonempty subset x of A, some $a_0 \in x$ and a class H that satisfies the above equation. If $H \restriction x = \varnothing$, then we have that $\forall a \in x\, ((H)_a = \varnothing)$ and thus, by equation (3.8) that

$$(H)_{a_0} = \{y \mid \varphi(H \restriction A_{<_A a_0}, x, y)\} = 1,$$

contradicting $(H)_{a_0} = \varnothing$. If $H \restriction x \neq \varnothing$, then there must exist an $a \in x$ such that $(H)_a \neq \varnothing$, therefore we can conclude with (3.8) that

$$1 = (H)_a = \{y \mid \varphi(H \restriction A_{<_A a}, x, y)\},$$

and thus that

$$\forall c \in x\,((H \restriction A_{<_A a})_c = \varnothing). \tag{3.10}$$

We now prove that a is minimal in x. If $b <_A a$ we can apply the assumption on the transitivity of $<_A$ to conclude that $A_{<_A b} \subset A_{<_A a}$ and consequently that $(H \restriction A_{<_A b})_c \subset (H \restriction A_{<_A a})_c$ for all c. Therefore, it follows from equation (3.10) that

$$\forall c \in x\,((H \restriction A_{<_A b})_c = \varnothing).$$

From equations (3.9) and (3.8) it therefore follows that

$$(H \restriction A_{<_A a})_b = (H)_b = 1.$$

Thus (3.10) implies that b cannot be an element of x and we are done. $\square$

Remark 186. In the previous theorem, the transitivity condition can be dropped for the necessity of a nontrivial application of the axiom of choice, which in the proof presented above is not necessary (a short proof is presented in the discussion part of [FS13]).

Corollary 187. *If $\Psi(A, <_A)$ is a formula with the following properties*

1. $\forall (A, <_A)\ (\Psi(A, <_A) \rightarrow \mathsf{tran}(A, <_A))$
2. $\exists (A, <_A)\ (\neg \mathsf{wf}(A, <_A) \wedge \Psi(A, <_A))$,

then the schema

$$\Psi(A, <_A) \rightarrow \exists H\, \forall a \in A\, \big((H)_a = \{x \mid \varphi(H \restriction A_{<_A a}, x)\}\big),$$

where $\varphi(X, x)$ ranges over all elementary formulas is inconsistent with classical first order logic (in fact also with much weaker systems).

Now we present three formulations of usual transfinite recursion and show that they are pairwise equivalent.

Definition 188. The schemata $\Delta^1_0\text{-}\mathsf{TR}_{\mathsf{wo}}$ and $\Delta^1_0\text{-}\mathsf{TR}_{\mathsf{wf}}$ are as follows:

$$\mathsf{wo}(A, <_A) \to \exists F\, \forall a \in A\, \big((F)_a = \Gamma_\varphi(F \restriction A_{<_A a}, a)\big) \qquad (\Delta^1_0\text{-}\mathsf{TR}_{\mathsf{wo}})$$
$$\mathsf{wf}(A, R) \to \exists F\, \forall a \in A\, \big((F)_a = \Gamma_\varphi(F \restriction \{b \in A \mid bRa\}, a)\big) \qquad (\Delta^1_0\text{-}\mathsf{TR}_{\mathsf{wf}})$$

where $\varphi(X, x, a)$ ranges over elementary formulas (possibly with further parameters) and $\Gamma_\varphi(X, a) = \{x \mid \varphi(X, x, a)\}$.

Remark 189. It is provable in NBG that the schema

$$\mathsf{wo}(A, <_A) \to \exists H\, \forall a \in A\, \big((H)_a = \Gamma_\varphi(H \restriction A_{<_A a})\big) \tag{3.11}$$

where $\varphi(X, x)$ ranges over elementary formulas, proves the schema $\Delta^1_0\text{-}\mathsf{TR}$. To see this, let

$$X \sqcup Y = \{\langle u, v\rangle \mid (u \in X \wedge v = 0) \vee (u \in Y \wedge v = 1)\}$$

denote the disjoint union of any classes X and Y. Given some elementary formula $\varphi(X, x, a)$, we stipulate $\psi(X, x)$ such that

$$\Gamma_\psi(X \sqcup Y) = \{\langle u, v\rangle \mid (\varphi(X, u, a_Y) \wedge v = 0) \vee (v = 1 \wedge u = a_Y)\}$$

where a_Y stands for

$$\min_{<_A}\{b \in A \mid b \notin Y\}.$$

Thus, applying (3.11) provides a class H, such that $(H)_a = \Gamma_\psi(H \restriction A_{<_A a})$. Stipulating F such that for all $a \in A$ $(F)_a = \{u \mid \langle u, 0\rangle \in (H)_a\}$ provides the desired class.

Definition 190. We call a function $R : A \to \mathcal{P}(A)$ a recursor for $(A, <_A)$ and write $\mathsf{rec}(R, (A, <_A))$ if the conjunction of the following clauses is satisfied:

1. $\forall a \in A\, (a \in R(a))$
2. $\forall a, b \in A\, (a <_A b \to A_{<_A R(a)} \subset A_{<_A R(b)})$
3. $\forall X \prec A\, \big(X \neq A \to \exists a \in \overline{X}\, (A_{<_A R(a)} \subset X)\big)$

Remark 191. Let $R : A \to \mathcal{P}(A)$ be some function and let $(A, <_A)$ be a linear order. If we stipulate

$$a <_R b \Leftrightarrow a <_A R(b),$$

then the third clause in Definition 190 is equivalent to the statement that $(A, <_R)$, is a well founded relation. This is crucial for Definition 196 where essentially elementary transfinite recursion along the relation $<_R$ is introduced.

Proof of the equivalence claim. Let $R : A \to \mathcal{P}(A)$ be some function and let $(A, <_A)$ be a linear order. First assume that $<_R$ is well founded on A and let X be a proper initial segment of $(A, <_A)$. Since $Y = A \setminus X$ is a nonempty subclass of A, we can use the well foundedness of $<_R$ on $(A, <_R)$, to pick a $<_R$-least element $a \in Y$. Because $A_{<_R a} = A_{<_A R(a)}$ it follows that $A_{<_A R(a)}$ is disjoint from Y and thus that $A_{<_A R(a)} \subset X$ as desired. For the converse direction, assume that R satisfies the third clause of Definition 190. We have to show that $<_R$ is well founded on A. Let Y be any nonempty subclass of A. First, assume that Y is coinitial in $(A, <_R)$. Thus, $X = A \setminus Y$ is a proper initial segment of $(A, <_R)$ and hence we can apply the third clause of Definition 190 to obtain an element $a \in Y$ such that $A_{<_R a} \subset X$, clearly a is $<_R$-minimal in Y. If Y is not coinitial, we can use the previous case to get a $<_R$-minimal element of the class $\{z \in A \mid \exists a \in Y\, (a \leq_R z)\}$, clearly such an element is also $<_R$-minimal in Y. $\square$

Remark 192. In cases where $(A, <_A)$ is a weak well order, all proper initial segments of $(A, <_A)$ are of the form $A_{<_A m}$ for some nonempty subset m of A; thus, the last clause of Definition 190 can be expressed by the following elementary formula

$$\forall m \subset A \left(m = \varnothing \vee \exists a \in A \,(a \notin A_{<_A m} \wedge A_{<_A R(a)} \subset A_{<_A m}\right). \tag{3.12}$$

Remark 193. If R is already monotone, i.e. R satisfies the second clause of Definition 190, then the last item is equivalent to

$$\forall X \prec A \left(\{a \in A\backslash X \mid A_{<_A R(a)} \subset X\} \text{ is nonempty and initial in } A\backslash X \right).$$

Example 194. 1. A linear order $(A, <_A)$ is a well order if and only if the mapping $R : x \to \{x\}$ is a recursor for $(A, <_A)$.

2. If $(A, <_A)$ is a linear order and $m \subset A$ is some minimal subset of A, then the mapping $R : x \to m \cup \{x\}$ is a (admittedly uninteresting) recursor for $(A, <_A)$.

Proof. For the first claim, note that for the function R, as defined in the claim, we have that

$$a <_R b \Leftrightarrow a <_A \{b\} \Leftrightarrow a <_A b$$

for all elements $a, b \in A$. Thus, it follows from Remark 191 that $(A, <_A)$ is a well order if and only if R, as defined above, is a recursor for $(A, <_A)$. For the second claim, if m, R and $(A, <_A)$ are as stated in the claim, then clearly $\forall a \in A\, (A_{<_A R(a)}) = \varnothing$. Thus the second and the third clause of Definition 190 are trivially satisfied. However, since $x \in m \cup \{x\}$ also the first clause is fulfilled. □

Remark 195. Let $\mathcal{C}$ be a collection of linear orders, if $\varphi(R, (A, <_A))$ is some formula such that for all linear orders exactly one function $R : A \to \mathcal{P}(A)$ with $\varphi(R, (A, <_A))$ exists, then we say that $\varphi(R, (A, <_A))$ *characterizes* $\mathcal{C}$ *with recursors* if for all linear orders $(A, <_A)$

$$(A, <_A) \in \mathcal{C} \Leftrightarrow \mathsf{rec}(R_\varphi(A, <_A), (A, <_A)),$$

where $R_\varphi(A, <_A)$ is the unique function that satisfies $\varphi(R, (A, <_A))$, holds. In that sense, we have seen in the first part of Example 194 that

$$\varphi(R, (A, <_A)) \equiv R = \{\langle x, \{x\}\rangle \mid x \in A\}$$

characterizes the collection of all well orders with recursors. However, there is no simple (elementary) formula $\varphi(R, (A, <_A))$ that can characterize weak well orders in general. Thus, it is difficult to give interesting examples of actual recursors for weak well orders at this point. We will see in Lemma 214, that a particular collection of weak well orders can be characterized with recursors by an elementary formula and in section 3.2.1, we discuss a theory that in some sense axiomatically ads a characterization with recursors for the collection of all weak well orders.

Definition 196. The schema Δ^1_0-$\mathsf{TR}_{\mathsf{wwo}}$ is

$$\mathsf{wwo}(A,<_A) \wedge \mathsf{rec}(R,(A,<_A)) \to \exists F\,\forall a \in A\,\big((F)_a = \Gamma_\varphi(F \restriction A_{<_A R(a)}, a)\big)$$

for elementary formulas φ.

The classes whose existence is guaranteed from the different schemata of elementary recursion Δ^1_0-$\mathsf{TR}_{\mathsf{wwo}}$, Δ^1_0-$\mathsf{TR}_{\mathsf{wo}}$ and Δ^1_0-$\mathsf{TR}_{\mathsf{wf}}$ respectively, are all, as shown below, unique.

Lemma 197. *Let $<_A$ be a binary relation on a class A and let $\varphi(X,x)$ be an elementary formula.*

1. *If $(A,<_A)$ is well founded, then there exists at most one class $H \subset A \times V$ such that*
$$\forall a \in A\,((H)_a = \Gamma_\varphi(H \restriction A_{<_A a}, a)).$$

2. *If $(A,<_A)$ is a weak well order and R is a recursor for $(A,<_A)$, then there is at most one class $H \subset A \times V$ that satisfies*
$$\forall a \in A\,((H)_a = \Gamma_\varphi(H \restriction A_{<_A R(a)}, a)).$$

Proof. The proofs of the first and the second part of the lemma are essentially the same, with the only difference that in the first case, we will use elementary transfinite induction, whereas in the second case, we will employ weak elementary induction.

1. Assume that H and G are two classes that both satisfy the formula presented in the lemma. We use transfinite induction to prove that the class $X = \{a \in A \mid (H)_a = (G)_a\} = A$ holds. Assume that $A_{<_A a} \subset X$ for some $a \in A$. It follows that for every $b <_A a$ that $(H)_b = (G)_b$ and thus that $H \restriction A_{<_A a} = G \restriction A_{<_A a}$. This immediately leads to $(H)_a = (G)_a$ as desired.

2. Assume that H and G are two classes that both satisfy the formula presented in the second part of the lemma. Let X be defined as before. We

have to show that for all nonempty subsets $m \subset A$, we have that

$$A_{<_A m} \subset X \to X \cap m \neq \varnothing$$

holds. Let m be a nonempty subset of A such that $A_{<_A m} \subset X$ holds. Since $A_{<_A m}$ clearly is a proper initial segment of $(A, <_A)$, it follows from the fact that R is a recursor of $(A, <_A)$ that there exists an element $a \in A \backslash A_{<_A m}$ such that $A_{<_A R(a)} \subset A_{<_A m}$ holds. Since we have that $a \not<_A m$, there must exist a $b \in m$ with $b \leq a$. It follows from the fact that R is a recursor for $(A, <_A)$ that $A_{<_A R(b)} \subset A_{<_A R(b)} \subset X$ and thus that

$$(H)_b = \Gamma_\varphi(H \restriction A_{<_A R(b)}, b) = \Gamma_\varphi(G \restriction A_{<_A R(b)}, b) = (G)_b.$$

Thus, $b \in X \cap m$ as desired. □

Theorem 198. *The schemata* Δ^1_0-$\mathsf{TR}_{\mathsf{wwo}}$*,* Δ^1_0-$\mathsf{TR}_{\mathsf{wo}}$ *and* Δ^1_0-$\mathsf{TR}_{\mathsf{wf}}$ *are all equivalent over* NBG*.*

Proof. Δ^1_0-$\mathsf{TR}_{\mathsf{wf}} \Rightarrow \Delta^1_0$-$\mathsf{TR}_{\mathsf{wwo}}$: Let $\varphi(X, x)$ be any elementary formula and let $(A, <_A)$ be a weak well order with a recursor R. As we have observed in Remark 191, the relation $(A, <_R)$ is well founded. Since $<_R$ clearly is transitive, the principle Δ^1_0-$\mathsf{TR}_{\mathsf{wf}}$ provides a class F such that for any $a \in A$,

$$(F)_a = \Gamma_\varphi(F \restriction \{b \in A \mid b <_R a\}, a) = \Gamma_\varphi(F \restriction A_{<_A R(a)}, a)$$

holds.
Δ^1_0-$\mathsf{TR}_{\mathsf{wwo}} \Rightarrow \Delta^1_0$-$\mathsf{TR}_{\mathsf{wo}}$: This follows from the fact that for every well order $(A, <_A)$, the function $A \ni a \mapsto \{a\} \in \mathcal{P}(A)$ is a recursor for $(A, <_A)$.
Δ^1_0-$\mathsf{TR}_{\mathsf{wo}} \Rightarrow \Delta^1_0$-$\mathsf{TR}_{\mathsf{wf}}$: Let $(A, \prec)$ be any transitive, well founded relation and let Γ_φ be any operator with $\varphi(X, x, a)$ elementary. Let T be the tree of all $(A, \prec)$ descending chains. The linear order $(T, <_T)$ where $<_T$ means the Kleene Brouwer ordering of T is a well order because $(A, \prec)$ is well founded. We write $t(s) = s(dom(s) - 1)$ to mean the tail of any $s \in T \backslash \{\varnothing\}$. For any $a \in A$, let $s_a = \min_{<_T}\{s \in T \mid t(s) = a\}$ be the least sequence in T that has a as its tail element. Given a class F, let $\hat{F}$ be the class $\{(a, x) \mid (s_a, x) \in F\}$. In particular,

we have that $(\hat{F})_a = (F)_{s_a}$. Define the formula $\psi(X, x, s, a)$ as follows

$$
\begin{aligned}
& s \neq s_{s(t)} \wedge (s, x) \in (X)_{s_{t(s)}} \\
& \vee \\
& s = s_{t(s)} \wedge \varphi(\hat{X} \restriction \{a \in A \mid s \subsetneq s_a\}, x, a)
\end{aligned}
$$

Now we use elementary transfinite recursion along $(T, <_T)$ to get a class H that satisfies for any $s \in T$ that

$$(H)_s = \Gamma_\psi(H \restriction T_{<_T s}, s, a),$$

and therefore also for any $a \in A$ that

$$
\begin{aligned}
(\hat{H})_a = (H)_{s_a} &= \Gamma_\psi(H \restriction T_{<_T s_a}, s_a, a) = \{x \mid \psi(H \restriction T_{<_T s_a}, x, s_a, a)\} \\
&= \{x \mid \varphi(\widehat{H \restriction T_{<_T s_a}} \restriction \{b \in A \mid s_a \subsetneq s_b\}, x, a)\}. \qquad (3.13)
\end{aligned}
$$

We observe that all the sequences s_x with $x \in A$ satisfy for all $k \in dom(s_x)$ the equations

$$s_x(k) = \min_{\lhd} \big(A_{\prec x} \backslash \{s_x(n) \mid n < k\}\big).$$

Hence an easy induction argument reveals that for any two distinct such sequences one is a proper extension of the other. Together with the observation that for any $a, b \in A$ we have that

$$b \prec a \Rightarrow (s_a \frown b) \in T \Rightarrow s_b \leq_T (s_a \frown b) <_T s_a,$$

we conclude that

$$b \prec a \Leftrightarrow s_a \subsetneq s_b.$$

Therefore,

$$
\begin{aligned}
& \widehat{H \restriction T_{<_T s_a}} \restriction \{b \in A \mid s_a \subsetneq s_b\} \\
& = \{(b, x) \mid (s_b, x) \in H \wedge s_b <_T s_a\} \restriction \{b \in A \mid s_a \subsetneq s_b\} \\
& = \{(b, x) \mid (s_b, x) \in H \wedge s_b <_T s_a \wedge s_a \subsetneq s_b\}
\end{aligned}
$$

$$
\begin{aligned}
&= \{(b,x) \mid (s_b,x) \in H \wedge s_a \subsetneq s_b\} \\
&= \{(b,x) \mid (s_b,x) \in H \wedge b \prec a\} \\
&= \{(b,x) \mid (b,x) \in \hat{H} \wedge b \prec a\} = \hat{H} \restriction A_{\prec a}.
\end{aligned}
$$

Returning to the equation (3.13), we therefore get that

$$
\begin{aligned}
(\hat{H})_a &= \{x \mid \varphi(\widehat{H \restriction T_{<_T s_a}} \restriction \{b \in A \mid s_a \subsetneq s_b\}, x, a)\} \\
&= \{x \mid \varphi(\hat{H} \restriction A_{\prec a}, x, a)\} = \Gamma_\varphi(\hat{H} \restriction A_{\prec a}, a),
\end{aligned}
$$

as desired. □

Definition 199. The theory Δ^1_0-TR of *elementary transfinite recursion* stands for any (and thus every) of the theories $\mathsf{NBG} + \Delta^1_0$-$\mathsf{TR}_{\mathsf{wwo}}$, $\mathsf{NBG} + \Delta^1_0$-$\mathsf{TR}_{\mathsf{wo}}$ or $\mathsf{NBG} + \Delta^1_0$-$\mathsf{TR}_{\mathsf{wf}}$.

Our first step in analyzing the theory of elementary transfinite recursion is to prove that it properly extends NBG. We do this by proving that the theory is strong enough to provide for a truth set relative to NBG.

Definition 200 (Gödel numbers and truth sets)**.** We extend the language of set theory by a constant symbol c_x for each set x. The collection of all $\mathcal{L}_1^{con}$ formulas is obtained by allowing free variables in $\mathcal{L}_1$ formulas to be replaced by any of the new constant symbols. We define for each $\mathcal{L}_1^{con}$ formula φ a finite set $\ulcorner\varphi\urcorner$ recursively on the build up of φ:

1. If φ is of the form $x \in y$ or $x = y$ where x and y are either constants or variables, then $\ulcorner\varphi\urcorner$ is the sequence $\langle 0, \tilde{x}, \tilde{y}\rangle$ or $\langle 1, \tilde{x}, \tilde{y}\rangle$, where $\tilde{x}$ and $\tilde{y}$ are given from

$$
\tilde{z} = \begin{cases} 2^k + 5 & \text{if } z = v_k \\ \langle 5, u\rangle & \text{if } z = c_u \end{cases}
$$

   where z stands for either x or y.

2. If φ is of the form $\sigma \wedge \delta$, then $\ulcorner\varphi\urcorner$ is the sequence $\langle 2, \ulcorner\sigma\urcorner, \ulcorner\delta\urcorner\rangle$.

3. If φ is of the form $\neg\psi$, then $\ulcorner\varphi\urcorner$ is the sequence $\langle 3, \ulcorner\psi\urcorner\rangle$.

4. If φ is of the form $\sigma \vee \delta$, then $\ulcorner\varphi\urcorner$ is $\ulcorner\neg(\neg\sigma \wedge \neg\delta)\urcorner$.
5. If φ is of the form $\sigma \rightarrow \delta$, then $\ulcorner\varphi\urcorner$ is $\ulcorner\neg\sigma \vee \delta\urcorner$.
6. If φ is of the form $\exists v_i\,(\psi)$, then $\ulcorner\varphi\urcorner$ is the sequence $\langle 4, 2^i + 5, \ulcorner\psi\urcorner\rangle$.
7. If φ is of the form $\forall v_i\,(\psi)$, then $\ulcorner\varphi\urcorner$ is $\ulcorner\neg(\exists v_i\,(\neg\psi))\urcorner$.

We will write Fml to mean the class of all codes of $\mathcal{L}_1^{con}$ formulas.

Definition 201. A *truth set* is a set t of codes of closed $\mathcal{L}_1$ formulas that satisfies

$$\sigma \leftrightarrow \ulcorner\sigma\urcorner \in t$$

for all closed $\mathcal{L}_1$ formulas σ.

Lemma 202 (Tarski)**.** *If t is a truth set, then t is not definable by a first order formula, i.e. there is no first order formula $\varphi(x)$ such that $t = \{x \mid \varphi(x)\}$ holds.*

Proof. A proof of this standard result can be found in most textbooks on logic and set theory, see for example [Jec03] Theorem 12.7. □

Theorem 203. *It is provable in* NBG *together with the principle of elementary transfinite recursion restricted to the well order $(\omega, <)$, that a truth set exists.*

Proof. We define a class T of codes of $\mathcal{L}_1^{con}$ formulas by elementary transfinite recursion as follows: let $B \subset Fml$ stand for the true atomic formulas, i.e. B is the least class that contains all formulas that are of the form $\ulcorner x = x\urcorner$ where x is either a constant symbol or a variable and all formulas that are of the form $\ulcorner c_x \in c_y\urcorner$ where $x \in y$. $\bar{B}$ on the other hand is the class of all false atomic formulas, i.e. $\bar{B}$ is the least class that contains all formulas of the form $\ulcorner x = y\urcorner$ where x, y are different constant symbols or variables, and all formulas that are of the form $\ulcorner c_x \in c_y\urcorner$ where $x \notin y$. Further, if x is the code of some formula $\ulcorner\exists v_i\,\varphi\urcorner = \langle 4, 2^i + 5, y\rangle$, then let x_u be the code of the formula $\varphi(\frac{c_u}{v_i})$, i.e. the unique sequence that satisfies $dom(x_u) = dom(y)$ and for all $z \in dom(y)$

$$x_u(z) = \begin{cases} y(z) & \text{if } y(z) \neq 2^i + 5 \\ \langle 5, u\rangle & \text{otherwise.} \end{cases}$$

We use elementary transfinite recursion along ω in order to simultaneously define two subclasses T and $\bar{T}$ of Fml such that for all $n \in \omega$ we have that

$$\begin{aligned} x \in (T)_n \leftrightarrow x \in B \vee \exists y, \exists z\, (x = \langle 2, y, z\rangle \wedge y, z \in \bigcup_{k<n} (T)_k) \\ \vee \exists y\, (x = \ulcorner 3, y \urcorner \wedge y \in \bigcup_{k<n} (\bar{T})_k) \\ \vee \exists u, y\, \exists i \in \omega\, (x = \langle 4, 2^i + 5, y\rangle \wedge x_u \in \bigcup_{k<n} (T)_k) \end{aligned}$$

and

$$\begin{aligned} x \in (\bar{T})_n \leftrightarrow x \in \bar{B} \vee \exists y, \exists z\, (x = \langle 2, y, z\rangle \wedge y \in \bigcup_{k<n} (\bar{T})_k \vee z \in \bigcup_{k<n} (\bar{T})_k) \\ \vee \exists y\, (x = \ulcorner 3, y \urcorner \wedge y \in \bigcup_{k<n} (T)_k) \\ \vee \exists y\, \exists i \in \omega \forall u\, (x = \langle 4, 2^i + 5, y\rangle \wedge x_u \in \bigcup_{k<n} (\bar{T})_k). \end{aligned}$$

Let σ be any formula. We prove by induction on the build up of σ that

$$\sigma \to \ulcorner \sigma \urcorner \in \bigcup_{n\in\omega} (T)_n \tag{1}$$

$$\neg\sigma \to \ulcorner \sigma \urcorner \in \bigcup_{n\in\omega} (\bar{T})_n \tag{2}$$

holds. In case that σ is an atomic formula the claim is clearly true. Assume that $\sigma = \varphi \wedge \psi$, if σ holds, then by induction hypothesis both codes $\ulcorner\varphi\urcorner$ and $\ulcorner\psi\urcorner$ are elements of $\bigcup_{n\in\omega} T_n$ and thus so is $\ulcorner\sigma\urcorner$. If otherwise, $\ulcorner\sigma\urcorner$ is false, then by induction hypothesis at least one of the codes $\ulcorner\varphi\urcorner$ or $\ulcorner\psi\urcorner$ must be an element of $\bigcup_{n\in\omega}(\bar{T})_n$ and thus so is $\ulcorner\sigma\urcorner$. If $\sigma = \neg\psi$, then $\ulcorner\sigma\urcorner$ is an element of $\bigcup_{n\in\omega}(T)_n$ iff $\ulcorner\psi\urcorner$ is an element of $\bigcup_{n\in\omega}(\bar{T})_n$, and vice versa. Hence, the claim follows from the induction hypothesis. For the case $\sigma = \exists v_i \psi$, first assume that σ holds, thus there exists a set u such that $\psi(\frac{u}{v_i})$ holds. Therefore, by induction hypothesis, we have that $\ulcorner\psi(\frac{c_u}{v_i})\urcorner \in \bigcup_{n\in\omega}(T)_n$ and thus that $\ulcorner\sigma\urcorner \in \bigcup_{n\in\omega}(T)_n$. If $\exists v_i\, \psi$ is false, then that means that for all sets u it is the case that $\neg\psi(\frac{u}{v_i})$ holds and thus, by induction hypothesis, that all codes of the form $\ulcorner\psi(\frac{u}{v_i})\urcorner$ are elements of $\bigcup_{n\in\omega}(\bar{T})_n$ and hence that $\ulcorner\sigma\urcorner \in \bigcup_{n\in\omega}(\bar{T})_n$. The cases where $\sigma = \varphi \vee \psi$ or $\sigma = \varphi \to \psi$ are reduced to the previous cases. Now let t be the class of all sets x such that x is the code of a closed $\mathcal{L}_1$ formula that is an element of some stage $(T)_n$ of T, then clearly for any closed formula σ of $\mathcal{L}_1$ we have that

$$\sigma \leftrightarrow \ulcorner\sigma\urcorner \in t.$$

□

Corollary 204. *Elementary transfinite recursion is not provable from* NBG.

Definition 205. The theory FP is NBG together with the assertions that positive elementary operators have a fixed point. That is

$$\exists Z\, \forall z\, \big(z \in Z \leftrightarrow \varphi(Z, z)\big)$$

where $\varphi(Z^+, z)$ ranges over elementary formulas and is positive in Z (cf. Definition 12).

Remark 206. Note that Lemma 13 also applies to the set theoretic case.

Remark 207. By the same argument that works in the arithmetic setting, the theory FP can be shown to prove elementary transfinite recursion. The proof of Theorem 3.1 in [Avi96] can be quoted verbatim to fit our framework.

The following two theorems are a remarkable example on how different the situation in arithmetic and set theory is for theories that substantially rely on the notion of well foundedness.

Theorem 208 (Sato)**.** *The theory* FP *proves the consistency of the theory* Δ^1_0-TR.

Proof. See [Sat12] Corollary 34. □

An even more striking result presented in [Sat13] is the fact that the theory FP is strong enough to prove the existence of least fixed points for any elementary monotone operator.

Theorem 209 (Sato)**.** *For every X-positive elementary formula $\varphi(X^+, y)$ the following is provable in the theory* FP*.*

$$\exists Z\, \forall U\, \big(\forall z\, (\varphi(Z, z) \to z \in Z) \,\wedge\, \forall z\, (\varphi(U, z) \to z \in U) \;\to\; Z \subset U\big)$$

Proof. See [Sat13] Corollary 5.4.i. □

In arithmetic, the theory that corresponds to FP is called FP_0, it is, as proved by Avigad (cf. [Avi96] Theorem 3.1), equivalent to the theory of arithmetical transfinite recursion ATR_0. As proved by Kleene in [Kle55], assuming least fixed points for positive arithmetical operators is powerful enough to prove the Π^1_1 comprehension axiom. Thus, in arithmetic, claiming the existence of least fixed points of positive arithmetical operators is beyond what is provable in the theory FP_0. Hence, while the results of Sato are interesting on their own, their real importance stems from the fact that they reveal a deep asymmetry between the situations in set theory as opposed to arithmetic. In the following, we will first see how an adaption of Avigad's proof of $\mathsf{FP}_0 = \mathsf{ATR}_0$ can be used to show that the logical complexity of the predicate wwo is reduced when specific changes are made (cf. Theorem 215). Thereafter, we will introduce the axiom of uniform recursion with the objective of pinning down a little more accurately which peculiarities of set theory and arithmetic respectively lead to the aforementioned differences.

Lemma 210. *Let $\varphi(X, a, x)$ be an elementary formula and let $(A, <_A)$ be a weak well order with a recursor R. It is provable in the theory Δ^1_0-TR that there is a class H such that*

$$\forall a, b \in A \left((H)_a = \bigcup_{b <_A R(a)} \{x \mid \varphi((H)_b, b, x)\} \right),$$

and in particular also

$$\forall a, b \in A\; (a <_A b \to (H)_a \subset (H)_b)\,.$$

Proof. Let $\varphi(X, a, x)$ be an elementary formula and let R be any recursor. We stipulate

$$\psi(X, a, x) = \exists b <_A R(a)\, (\varphi((X)_b, b, x)).$$

By an application of elementary transfinite recursion, we obtain a class H that satisfies

$$(H)_a = \{x \mid \psi(H \upharpoonright A_{<_A R(a)}, a, x)\},$$

and thus that

$$\begin{aligned} x \in (H)_a &\Leftrightarrow \psi(H \upharpoonright A_{<_A R(a)}, a, x) \\ &\Leftrightarrow \exists b <_A R(a)\, \varphi((H \upharpoonright A_{<_A R(a)})_b, b, x) \\ &\Leftrightarrow \exists b <_A R(a)\, \varphi((H)_b, b, x) \Leftrightarrow x \in \bigcup_{b <_A R(a)} \{z \mid \varphi((H)_b, b, z)\} \end{aligned}$$

as desired. The second part of the claim follows from the first part together with the property $a <_A b \to A_{<_A R(a)} \subset A_{<_A R(b)}$ of the recursor R of $(A, <_A)$. $\square$

Corollary 211. *It is provable in the theory Δ^1_0-TR that for every well order $(A, <_A)$ and every elementary formula $\varphi(X, a, x)$, there is a class H such that for all elements $a, b \in A$ we have that $(H)_a = \bigcup_{b <_A a} \{x \mid \varphi((H)_b, b, x)\}$ and in particular that $a <_A b \to (H)_a \subset (H)_b$*

Proof. This follows from Lemma 210 when considering that every well order is also a weak well order and the mapping $x \mapsto \{x\}$ is a recursor on any well order. $\square$

Definition 212. Let $(A, <_A)$ be a linear order. We write $\mathsf{wwo}^*(A, <_A)$ to mean that every subclass X of A contains a minimal subset that is also an initial segment of X (or equivalently a nonempty initial subset). More formally,

$$\forall X \subset A\, \exists m \subset X \big(\forall x \in X\, \exists y \in m\, (y \leq x) \wedge \forall x \in m\, (X_{<_A x} \subset m)\big).$$

From the arithmetic perspective, the notion of a wwo^* ordering as defined above seems equally adequate as weak well orders to translate the concept of a countable well ordering from arithmetic to set theory. Especially since the typical minimal sets of well orders, singletons, are also initial segments. However, as we will prove next, the predicate wwo^* can be expressed by a Σ^1_1 formula and is thus not useful in our context.

Definition 213. If $(A, <_A)$ is a linear order, then $R^*(A, <_A) : A \to \mathcal{P}(A)$ is the function defined from

$$R^*(A, <_A)(a) = \{b \in A \mid b \leq_A a \wedge [b, a] \text{ is a set}\},$$

where $[b, a]$ stands for the class $\{c \in A \mid b \leq_A c \wedge c \leq_A a\}$.

Lemma 214. *It is provable in* NBG *that for every linear order* $(A, <_A)$*,*

$$\mathsf{wwo}^*(A, <_A) \leftrightarrow \mathsf{rec}(R^*(A, <_A), (A, <_A))$$

holds.

Proof. We write R^* for $R^*(A, <_A)$. For the direction from right to left, let $Y \subset A$ be any nonempty subclass of A. First assume that Y is coinitial in $(A, <_A)$. Let $X = A \setminus Y$ and thus $X \prec A \wedge X \neq A$. Since R^* is a recursor for $(A, <_A)$, there is an element $a \in Y$ such that $A_{<_A R^*(a)} \subset X$, thus $R^*(a) \cap Y$ is a nonempty initial subset of Y, as desired. In case that Y is not coinitial in $(A, <_A)$, we can apply the previous case to the coinitial class $Y' = \{a \in A \mid \exists b \in Y\, (b \leq a)\}$ in order to obtain a minimal and initial subset m of Y'. Clearly $m \cap Y$ is minimal and initial in Y. We now turn to the converse direction of the claim. We first have to make sure that for all $a \in A$ the class $R^*(a)$ is a set. Let $a \in A$ be arbitrarily chosen. Since by assumption we have that $\mathsf{wwo}^*(A, <_A)$ holds, we can assume that there is an initial and minimal subset m of $R^*(a)$; further, since $R^*(a) \neq \varnothing$, we can find a $b \in m$, and since $b \in R^*(a)$, we can write $R^*(a) = [b, a] \cup m$ as the union of two sets. In order to verify that R^* is a recursor for $(A, <_A)$, we still have to check

1. $\forall a \in A\, (a \in R^*(a))$
2. $\forall a, b \in A\, (a <_A b \to A_{<_A R^*(a)} \subset A_{<_A R^*(b)})$
3. $\forall X \prec A\, (X \neq A \to \exists b \in A \setminus X\, (A_{<_A R^*(b)} \subset X))$.

Items 1. and 2. are clearly satisfied. To verify 3., let X be any proper initial segment of $(A, <_A)$. Since $\mathsf{wwo}^*(A, <_A)$ holds, there exists a minimal and initial subset m of $Y = A \setminus X$. Because $Y \neq \varnothing$, it follows from the minimality of m

that also $m \neq \varnothing$. Pick any $a \in m$. We want to prove that $A_{<_A R^*(a)} \subset X$. Because of $A_{<_A m} = X$, it is enough to show that $A_{<_A R^*(a)} \subset A_{<_A m}$, i.e. that $\forall b \in A\,(b <_A R^*(a) \rightarrow b <_A m)$. To see this, consider that for all b with $b <_A R^* a$ we have that $[b,a]$ is not a set and thus that $[b,a] \not\subset m$ which implies that $b <_A m$ because m is an initial segment of a coinitial subclass of $(A, <_A)$ and $a \in m$. $\square$

Theorem 215. *There is a Σ^1_1 formula $\varphi(A, <_A)$ such that*

$$\forall (A, <_A)\ (\varphi(A, <_A) \leftrightarrow \mathsf{wwo}^*(A, <_A))$$

is consistent with Δ^1_0-TR.

Proof. By way of contradiction. We assume that there is no Σ^1_1 formula σ such that $\sigma(A, <_A) \leftrightarrow \mathsf{wwo}^*(A, <_A)$ holds for all $(A, <_A)$, and then we will use this assumption to prove the schema FP within Δ^1_0-TR and thus contradict Theorem 208. We work in the theory Δ^1_0-TR. Let $\varphi(X, x)$ be any elementary and X-positive formula. If $\mathsf{wwo}^*(A, <_A)$ holds of some linear order $(A, <_A)$, then, as seen in Lemma 214, the function $R^* = R^*(A, <_A)$ is a recursor for $(A, <_A)$. Because further we have for all $b \in A$ that $[b,a] \notin V \Leftrightarrow b <_A R^*(a)$, we can apply elementary transfinite recursion and Lemma 210 to obtain a class H such that for all elements $a, b \in A$

1. $a <_A b \rightarrow (H)_a \subset (H)_b$

2. $(H)_a = \bigcup_{b <_A R^*(a)} \{x \mid \varphi((H)_b, x)\} = \bigcup_{[b,a] \notin V} \Gamma_\varphi((H)_b)$

holds because $[b,a] \notin V \Leftrightarrow b < a \wedge [b,a] \notin V \Leftrightarrow b <_A R^*(a)$. Further, note that since $\mathsf{wwo}^*(A, <_A)$, for all sets x there are minimal and initial subsets of the class $\{a \in A \mid x \in (H)_a\}$. Therefore, and in view of Lemma 214, the following is provable from elementary transfinite recursion:

$$
\begin{aligned}
\mathsf{wwo}^*(A,<_A) \to \exists H \big(&\mathsf{lo}(A,<_A)\\
&\wedge\ \forall a,b \in A\,(a <_A b \to (H)_a \subset (H)_b)\\
&\wedge\ \forall a \in A\,((H)_a = \bigcup_{[b,a]\notin V} \Gamma_\varphi((H)_b)\\
&\wedge\ \forall x\,\exists m\,(m \text{ is minimal and initial in } \{a \mid x \in (H)_a\})\big)\\
&\wedge\ \forall a \in A\,(\{b \in A \mid \varnothing \neq [b,a] \text{ is a set}\} \text{ is a set}).
\end{aligned}
\tag{3.14}
$$

As a consequence of our assumption that $\mathsf{wwo}^*(A,<_A)$ cannot be expressed by any Σ^1_1 formula and the fact that the right hand side of the above expression is indeed Σ^1_1, it follows that there must exist a linear order $(A,<_A)$ such that $\neg\mathsf{wwo}^*(A,<_A)$, but the right hand side of (3.14) holds of $(A,<_A)$. Fix any monotone elementary formula $\varphi(X,x)$ and let $(A,<_A)$ be any such linear order and let H be a class that satisfies the right side of (3.14). Since $(A,<_A)$ is not a wwo^* order, there must exist an initial subclass X of A such that for $Y = A\backslash X$ there is no minimal and initial subset. We stipulate

$$F = \bigcup_{a\in X} (H)_a.$$

If we can prove that F is a fixed point of the operator Γ_φ, then we are done. First, consider that

$$
\begin{aligned}
F = \bigcup_{a\in X} (H)_a &= \bigcup_{a\in X} \big(\bigcup_{[b,a]\notin V} \Gamma_\varphi((H)_b)\big) = \bigcup_{\substack{a,b\in X\\ [b,a]\notin V}} \Gamma_\varphi((H)_b)\\
&\subset \bigcup_{b\in X} \Gamma_\varphi((H)_b) \subset \Gamma_\varphi\big(\bigcup_{b\in X}(H)_b\big) = \Gamma_\varphi(F)
\end{aligned}
$$

where monotonicity of Γ_φ is applied in the second line. In order to see also $\Gamma_\varphi(F) \subset F$, we prove the identity

$$F = \bigcap_{a\in Y} (H)_a, \tag{3.15}$$

and then apply monotonicity of Γ_φ to obtain

$$\Gamma_\varphi(F) = \Gamma_\varphi(\bigcap_{a\in Y}(H)_a) \subset \bigcap_{a\in Y}(\Gamma_\varphi(H)_a).$$

Thereafter, in order to verify that $\Gamma_\varphi(F) \subset F$, it is enough to prove that for every $b \in Y$ there exists a $a \in Y$ such that $\Gamma_\varphi((H)_a) \subset (H)_b$. Let a be any element of Y. By (3.14), we know that if $R^*(a)$ is minimal in Y, then it is a minimal and initial subset, which cannot be the case by our choice of Y, hence there must be an element $b \in Y$ such that $[b, a] \notin V$. Therefore,

$$\Gamma_\varphi((H)_b) \subset \bigcup_{[c,a]\notin V} \Gamma_\varphi((H)_c) = (H)_a.$$

To finish the proof, we still need to verify the identity (3.15). The inclusion $F \subset \bigcap_{a\in Y}(H)_a$ follows immediately from the second clause of (3.14). For the reverse inclusion, assume that some x is not in $F = \bigcup_{a\in X}(H)_a$. Hence, $Z = \{a \mid x \in (H)_a\}$ is a subclass of Y. Because of the second to last clause of (3.14), the class Z has a minimal and initial subset. Therefore, there must exist an $a \in Y_{<_A Z}$; hence, we have an $a \in Y$ such that $x \notin (H)_a$ and thus that $x \notin \bigcap_{a\in Y}(H)_a$. $\square$

3.2.1 The uniform recursion axiom

After we found that, regardless of whether or not we formulate elementary transfinite recursion with well orders or weak well orders, the resulting theory will still be much weaker than the theory FP of fixed points, we will now try to pin down the differences between set theory and arithmetic that cause this dissimilarity of the two settings. When analyzing Avigad's original proof (cf. [Avi96], Theorem 3.1) of the equivalence $\mathsf{FP}_0 = \mathsf{ATR}_0$, an important difference between arithmetic and set theory seems to be the fact that in arithmetic, it is possible to arithmetically define a function $R : \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\mathbb{N})$ such that a linear ordering $\prec$ of $\mathbb{N}$ is a well ordering if and only if R is a recursor for $(\mathbb{N}, \prec)$. In fact, a linear order $\prec$ on $\mathbb{N}$ is exactly then a well order if the mapping $x \mapsto \{x\}$ is a recursor for $(\mathbb{N}, \prec)$. Moreover, we have seen that the situation is similar in the case of wwo^* orderings (cf Lemma 214). We now turn our attention towards the question of what happens if for every linear ordering $<$ of V, a function $R(<) : \mathcal{P}(V) \to \mathcal{P}(V)$ exists with the property that for any linear ordering $<$ the class $R(<)$ is a recursor for $(V, <)$ if and only if $<$ weakly well orders V. As it turns out, with this assumption we can prove FP from Δ^1_0-TR in a similar fashion as we did in proving Theorem 215. In order to process this vague observation into an actual theorem, we will need to add a class with similar properties such as the aforementioned function R to our base theory, and then we have to prove FP within this new theory.

First, we note that a class $R(<)$ as described above must depend on $<$, unlike in the case of well orders but similar as in the case of wwo^* orderings. In other words, as shown in the next proposition, it is inconsistent to merely add a function $R : V \to V$ and claim that for any linear ordering $<$ of the universe one has $\mathsf{wwo}(<) \leftrightarrow \mathsf{rec}(R, <)$.

Lemma 216. *For every function $R : V \to V$, there exists a weak well ordering $<$ of V such that R is not a recursor for $<$.*

Proof. Let $R : V \to V$ be any function. We define a sequence a_i for $i \in \omega$ by

$$a_n = \min_{\lhd}(V \setminus \bigcup_{i<n} R(a_i)).$$

In the following, we define a linear ordering $<^*$ on $X = \bigcup_{i \in \omega} R(a_i)$. First, we define

$$R^*(a_n) = R(a_n) \setminus \bigcup_{i<n} R^*(a_i)$$

which are a partition of X into pairwise nonempty sets. Each of the sets $R^*(a_n)$ shall be ordered by $<^*$ such that a_n is the top element and $<^*$ equals $\lhd$ on $R^*(a_n) \setminus \{a_n\}$. Further, for sets $x \in R^*(a_n)$ and $y \in R^*(a_k)$ with n and k distinct, we let $x <^* y$ if $n > k$. To see that R is not a recursor for $<^*$, consider that if $a \in X$, then there is some natural number k such that $a \in R^*(a_k)$, and thus that $a_{k+1} \in X_{<^* R(a)}$. Therefore, there exists no set a in X such that $X_{<^* R(a)} = \varnothing$; hence, R is not a recursor for $(X, <^*)$. Since clearly $(X, <^*)$ is a weak well order, this proves the claim. $\square$

Lemma 216 shows that if we want to have some kind of uniform recursor, that is, a mapping such as $< \mapsto R(<)$ as described above, we first have to clarify how we can code a function that maps classes to classes. Then, as the next step we will state the *uniform recursion axiom* as the assertion that there is a (coded) function $\mathcal{F}$ that satisfies the following condition:

$$\mathcal{F}(X) \text{ is a recursor for } (V, X) \Leftrightarrow (V, X) \text{ is a weak well order.}$$

Definition 217. For any class X, we write $Ch(X)$ to mean the function $Ch(X) : On \to 2$ which maps an ordinal α to 1 if the αth (with respect to $\lhd$) set is an element of X. Formally, $Ch(X)(\alpha)$ is defined by transfinite recursion as follows:

$$x_\alpha = \min_{\lhd}(V \setminus \{x_\mu \mid \mu < \alpha\})$$

$$Ch(X)(\alpha) = \begin{cases} 1 & \text{if } x_\alpha \in X \\ 0 & \text{otherwise.} \end{cases}$$

We will refer to $Ch(X)$ as the *characteristic function* of X.

Definition 218. A class $F \subset Seq(2 \times 2)$ is a *coded function* if

$$\forall X, Y, Z \big(\forall \alpha \, (Ch(X) \otimes Ch(Y)) \restriction \alpha \in F \wedge \forall \alpha \, ((Ch(X) \otimes Ch(Z)) \restriction \alpha \in F) \\ \to Y = Z\big).$$

We will write $\mathsf{FUN}(F)$ to mean that F is a coded function. Further, if F is a coded function and $\forall \alpha \, (Ch(X) \otimes Ch(Y)) \restriction \alpha \in F$ holds, then we will write $F(X) = Y$.

Definition 219. The axiom of uniform recursion is the assertion that there exists a coded function U such that

$$\forall X \, \big(\mathsf{wwo}(V, X) \leftrightarrow \mathsf{rec}(U(X), (V, X))\big)$$

holds.

Now, after having stated our axiom of uniform recursion, we have to make sure that the new principle is consistent relative to some established theory.

Theorem 220. *The uniform recursion axiom is provable in* $\mathsf{NBG} + \Pi^1_1\text{-}\mathsf{CA}$.

Proof. We work in $\mathsf{NBG} + \Pi^1_1\text{-}\mathsf{CA}$ to define a class U such that $f \in U$ if and only if

$$
\begin{aligned}
&\exists F, G, X \big(F \otimes G \upharpoonright dom(f) = f \\
&\wedge F = Ch(X) \\
&\wedge \big(\neg \mathsf{wwo}(V, X) \wedge G = Ch(\{\langle x, \{x\}\rangle \mid x \in V\}) \\
&\quad \vee G = Ch(\{\langle x, m_0 \cup \{x\}\rangle \mid m_0 = \min_{\lhd}\{m \mid m \text{ is a minimal set of } X\}\})\big)\big).
\end{aligned}
$$

Identifying classes with their characteristic functions, basically, a sequence $f : \alpha \to 2 \times 2$ is an element of U if there are classes X, G such that f is an initial segment of $X \otimes G$ and if X is not a weak well ordering of V, then G is the function $x \mapsto \{x\}$; otherwise, there is a (unique) least (with respect to $\lhd$) minimal subset m of V (with respect to the ordering X) and G is the mapping $x \mapsto m \cup \{x\}$. Hence, U is a coded function such that

$$
U(X) \doteq \begin{cases} x \mapsto \{x\} & \text{if } \neg \mathsf{wwo}(V, X) \\ x \mapsto \{x\} \cup \text{“the least minimal set w.r.o } X\text{”} & \text{otherwise} \end{cases}
$$

where $\doteq$ reads as “is the characteristic function of”. Clearly, $U(X)$ is the characteristic function of a recursor for (V, X) exactly if X weakly well orders V. $\square$

Definition 221. The theory $\Delta^1_0\text{-}\mathsf{TR}^+$ consists of $\Delta^1_0\text{-}\mathsf{TR}$ and the uniform recursion axiom.

Theorem 222. *The principle* FP *is provable within the theory* $\Delta^1_0\text{-}\mathsf{TR}^+$.

Proof. The proof is very similar to that of Theorem 215. We work in the theory $\Delta^1_0\text{-}\mathsf{TR}^+$. Fix a class U as stated in the uniform recursion axiom. Applying elementary transfinite recursion and Lemma 210, we get that for every elementary formula $\varphi(X, x)$, the following holds.

$$\begin{aligned}
\mathsf{wwo}(V,<) \to \exists H \big(&\mathsf{lo}(V,<) && (3.16)\\
&\wedge\ \forall x, y\,(x < y \to (H)_x \subset (H)_y) && (3.17)\\
&\wedge\ \forall x\,((H)_x = \bigcup_{b<(U(<))(b)} \Gamma_\varphi((H)_b)) && (3.18)\\
&\wedge\ \forall x\,(x \in (U(<))(x)) && (3.19)\\
&\wedge\ \forall x, y\,(x < y \to \{z \mid z < U(<)(x)\} \subset \{z \mid z < U(<)(y)\}) && (3.20)\\
&\wedge\ \forall x\,\exists m\,(m \text{ is minimal in } \{z \mid x \in (H)_z\}) && (3.21)\\
&\wedge\ \forall m\,\exists x\,(\neg\,(x < m)\ \wedge \forall y\,(y < U(<)(x) \to y < m))\big). && (3.22)
\end{aligned}$$

We justify the last two clauses as follows. While (3.21) is just a way to express that some specific subclasses of V have minimal subsets with respect to $(V,<)$ and thus is granted by the assumption that $<$ weakly well orders V, the last part is a slightly harder to decipher. Remember that any recursor R of $(V,<)$, by definition, has the property

$$\forall X\,(X \text{ is an initial subclass of } (V,<) \to \exists y\,(y \notin X \wedge \{x \mid x < R(y)\} \subset X)). \tag{3.23}$$

Since this property is Π^1_1, we cannot directly include it into our formula. However, we can restrict the above statement to only initial subclasses of the form $V_{<m}$ for all sets m, that is initial subclasses whose complements have minimal subsets, and then apply the formula to the recursor $U(<)$ to get (3.22). Now fix a monotone elementary formula $\varphi(X, x)$. Since the right side of (3.16)$\wedge\cdots\wedge$(3.22) is a Σ^1_1 formula, it follows from Theorem 178 that there exists a linear ordering $<$ of V that is not a weak well order, but still satisfies the right side of (3.16)$\wedge\cdots\wedge$(3.22). Fix a class H such as stated in (3.16)$\wedge\cdots\wedge$(3.22). Since $<$ does not weakly well order V, the class $R = U(<)$ is not a recursor for $(V,<)$. Because of (3.19) and (3.20), the statement made in (3.23) must be violated, i.e. there must be an initial subclass X of $(V,<)$ such that there is no set $x \notin X$ with the property that $\{y \mid y < R(x)\} \subset X$. First, note that because of (3.22), the class $V\backslash X$

cannot have a minimal subset. We claim that

$$\bigcup_{x\in X}(H)_x = \bigcap_{x\notin X}(H)_x$$

holds. The $\subset$ part of the equality is immediate from the second clause of the right side of (3.16)$\wedge\cdots\wedge$(3.22) and the fact that X is initial relative to $(V,<)$. To see the reverse inclusion, suppose that $y \notin \bigcup_{x\in X}(H)_x$. Therefore, y is such that the class $Y = \{x \mid y \in (H)_x\}$ is a subclass of $V\backslash X$. From (3.21), it follows that Y has a minimal subset and thus that Y is not minimal in $V\backslash X$, since otherwise any minimal subset of Y would also be a minimal subset of $V\backslash X$. Therefore, there must be a set $z \in V\backslash X$ that is below Y, thus

$$y \notin (H)_z \supset \bigcap_{x\notin X}(H)_x.$$

Now, we prove that $F = \bigcup_{x\in X}(H)_x = \bigcap_{x\notin X}(H)_x$ is a fixed point of Γ_φ, i.e. $\Gamma_\varphi(F) = F$. For the inclusion from left to right, we use monotonicity of Γ_φ together with (3.17) and (3.18) to obtain

$$\begin{aligned} F = \bigcup_{x\in X}(H)_x &= \bigcup_{x\in X}\big(\bigcup_{y<R(x)}\Gamma_\varphi((H)_y)\big) \\ &\subset \bigcup_{x\in X}\Gamma_\varphi((H)_x) \\ &\subset \Gamma_\varphi\big(\bigcup_{x\in X}(H)_x\big) = \Gamma_\varphi(F). \end{aligned}$$

For the inclusion of $\Gamma_\varphi(F)$ in F, since by monotonicity of Γ_φ we already know that

$$\Gamma_\varphi(F) = \Gamma_\varphi\big(\bigcap_{x\notin X}(H)_x\big) \subset \bigcap_{x\notin X}\Gamma_\varphi((H)_x)$$

holds, it is enough to show that for all sets y, we have

$$y \notin \bigcap_{x\notin X}(H)_x \to \exists z\,(z \notin X \wedge y \notin \Gamma_\varphi((H)_z)).$$

Suppose that there is a set x such that $x \notin X$ and $y \notin (H)_x$. It follows from how we chose X that $Z = \{z \mid z < R(x)\} \setminus X$ is a nonempty class. Choosing any z from Z yields

$$y \notin (H)_x = \bigcup_{i<R(x)} \Gamma_\varphi((H)_i) \supset \Gamma_\varphi((H)_z)$$

as desired. $\square$

Summary

We saw in Theorem 185 that generalizing the principle of elementary transfinite recursion cannot be done by only relaxing the well foundedness condition on the underlying relation. Nevertheless, we managed to give a translation of elementary transfinite recursion that is based on weak well orders (cf. Definition 196). In Theorem 198 we showed that our new principle is equivalent to the usual formulation of elementary transfinite recursion. For completeness, we gave a proof in Theorem 203 that elementary transfinite recursion is not provable from NBG. We suppose that the argument given is folklore. We continued by using our new interpretation of elementary transfinite recursion in order to use a result of Sato (cf. Theorem 208) to show that wwo^* orderings are not complete Π^1_1 (cf. Theorem 215). We observed that the existence of uniform recursors for wwo^* orderings allowed us to transform Avigad's proof of $\mathsf{FP}_0 \equiv \mathsf{ATR}_0$ into the proof for Theorem 215. We then formalized this line of thought and asserted the existence of uniform recursors for weak well orders in the uniform recursor axiom (cf. Definition 219). We proved in Theorem 220 that the uniform recursor axiom is provable from Π^1_1 comprehension and continued with proving the existence of fixed points of elementary monotone operators from elementary transfinite recursion together with uniform recursors for weak well orders (cf. Theorem 222).

3.3 Comparing weak well orders

As we have seen in Theorem 109, for any two well orderings $(A, <_A)$ and $(B, <_B)$, if at least one of $(A, <_A)$ or $(B, <_B)$ respectively is a set like well order, then it is provable in NBG that either $(A, <_A)$ can be embedded into $(B, <_B)$ or vice versa. It is a natural question to ask whether in Theorem 109 the condition of being set like can be dropped and the claim can still be proved from NBG, or otherwise what the consequences are if such a principle is added to the base theory. Unfortunately, we where not successful in providing conclusive answers to these questions. In particular, it remains open whether or not NBG proves the comparability of arbitrary well orders. What we do know, and what we assume to be folklore, is that the comparability of well orders follows from elementary transfinite recursion (cf. Theorem 225). In regard of the fact that in arithmetic the theory CWO is of the same strength as ATR_0 and FP_0 (cf. Theorem 15), the question arises whether or not it is possible to implement a comparability principle for weak well orders that can prove FP or at least can be added to Δ^1_0-TR to obtain a theory within which FP is provable. However, it became evident in the course of our investigations that comparing weak well orders is a very meticulous task. In particular, since there are weak well orders that clearly do not compare in the usual sense, it is clear that the usual approach which employs order isomorphisms has to be altered. In literature, the system CWO is often quoted as the weakest extension ACA_0 that allows for a reasonable treatement of ordinals (e.g. [Sim98] V.6.). In the second part of this section we attempt to translate this particular characteristic of the theories CWO and ATR_0 to the class theoretic setting. Informally, a decent framework of ordinals should at least provide enough ordinals to have order-types for all well orders. Our framework provides order-types for potentially all weak well orders and additionally incorporates the following basic features (cf. Proposition 253):

- The order-type of every weak well order is a well order
- Every well order is order isomorphic to its order-type
- Order isomorphic orders have order isomorphic order-types.

Based on our framework of order-types, we introduce a theory that guarantees order-types for all weak well orders and discuss dome basic properties of that theory. However, most questions about the abilities of this theory remain open.

3.3.1 Comparability of well orders

First, we will prove that the comparability of well orders is a consequence of the theory Δ^1_0-TR.

Definition 223 (Comparability of well orders)**.** The theory[3] CWO is NBG together with the statement that from any two well orders at least one embeds into the other (cf. Definition 103).

Lemma 224. *If Y is any class and $\varphi(X,x)$ is an elementary formula such that $\forall X\,(\Gamma_\varphi(X) \in Y)$, then it is provable in the theory Δ^1_0-TR that for every well order $(A,<)$ there is a function $F : A \to Y$ such that $\forall a \in A\,(F(a) = \Gamma_\varphi(F \restriction A_{<a}))$.*

Proof. For any well order $(A,<)$, a function F with the aforementioned properties can be obtained as follows. Let $\varphi(X,x)$ be an elementary formula. Stipulate

$$\psi(X,x) \equiv \forall z\,(z \in x \leftrightarrow \varphi(X,z)),$$

and let H be a class that satisfies the usual recursion schema for the operator Γ_ψ. Now define F by elementary comprehension

$$F = \{\langle a,y\rangle \in A \times Y \mid y = \min_{\lhd}(H)_a\}.$$

Let $a \in A$ and consider

$$\begin{aligned} F(a) &= \min_{\lhd}((H)_a) = \min_{\lhd}(\Gamma_\psi(H \restriction A_{<a})) \\ &= \min_{\lhd}(\{\Gamma_\varphi(H \restriction A_{<a})\}) = \Gamma_\varphi(H \restriction A_{<a}). \end{aligned}$$

□

[3] Due to the verbatim translation, we use the same acronym to name the comparability of well orders in arithmetic and in set theory respectively. It will, however, always be clear from the context which theory we are referring to.

Theorem 225. *In the theory* Δ^1_0-TR*, it is provable that* CWO *holds.*

Proof. Let $(A, <_A)$ and $(B, <_B)$ be any two well orders. Let a_0 and b_0 be the smallest elements of $(A, <_A)$ and $(B, <_B)$ respectively.

$$(\forall b \in B\, \exists a \in A\, (X(a) = b) \wedge x = b_0) \vee (x = \min_{<_B}\{b \in B \mid \neg\exists a \in A\, (X(a) = b)\})$$

Applying Lemma 224, we obtain a function $F : A \to B$ such that for all $a \in A$

$$F(a) = \Gamma_\varphi(F \upharpoonright A_{<_A a})$$

holds. In particular, we get a function F such that for all a in A

$$F(a) = \begin{cases} \min_{<_B}(B \backslash F[A_{<_A a}]) & \text{if possible} \\ b_0 & \text{otherwise.} \end{cases}$$

holds. We distinguish two cases. First, assume that for all $a \in A$ we have that $F(a) = b_0 \leftrightarrow a = a_0$. Obviously, F satisfies for all $a, a\prime \in A$ that $a < a\prime$ holds iff $F(a) < F(a\prime)$ holds. Therefore, if we show that $F[A]$ is an initial segment of $(B, <_B)$ we have shown that F embeds $(A, <_A)$ into $(B, <_B)$. Assume that $b < F(a)$ holds of some $b \in B$ and $a \in A$, then it follows that

$$b < F(a) = \min_{<_B}(B \backslash F[A_{<_A a}]) \Rightarrow b \in F[A_{<_A a}] \subset F[A].$$

In the case there is an $a \in A \backslash \{a_0\}$ with the property that $F(a) = b_0$, let a_1 be the least such element. Since $F \upharpoonright A_{<_A a_1}$ is a bijective mapping, we can fix $G = (F \upharpoonright A_{<a_1})^{-1}$. Clearly, G embeds $(B, <_B)$ into $(A_{<a_1}, <_A)$. $\square$

Corollary 226. *As a consequence of Theorem 225 and Theorem 208, we have that, unlike in the arithmetic case, comparability of well orders is not enough to prove the existence of fixed points of elementary positive operators.*

3.3.2 Order-types and comparability of weak well orders

As mentioned before, it is an immediate observation that simply extending the scope of the principle CWO from well orders to weak well orders yields an inconsistent theory.

Remark 227. For two given weak well orders, it is not necessarily the case that one embeds into the other.

Proof. Consider the order $\omega^* = (\omega, >)$, the set of natural numbers with their natural order inverted, and the order $(\omega, <)$, the natural numbers ordered as usual. Obviously both orders are weak well orders, but neither embeds into the other. $\square$

In view of this observation, we will turn our attention to the question on how to relax the notion of embedability, so that a corresponding comparability principle for weak well orders is meaningful. We will examine the strengths and shortcomings of the principles obtained that way. As we will see below (cf. Corollary 230), order embeddings are exactly those one-to-one functions between linear orders that preserve initial segments in both directions, i.e. every image of an initial segment is an initial segment in the target order and every preimage of an initial segment is an initial segment of the domain.

The following terminology is motivated by the fact that the collection of all initial subclasses of a linear order forms a topological space on the field of the relation.

Definition 228. A function $F : A \to B$ is a *continuous* map between the linear orders $(A, <_A)$ and $(B, <_B)$ if preimages of initial segments of $(B, <_B)$ are initial segments of $(A, <_A)$, i.e. if the following is satisfied.

$$\forall X\,(X \prec B \to F''X \prec A).$$

The function F is *open* if initial segments are mapped onto initial segments, i.e.

$$\forall X\,(X \prec A \to F[X] \prec B).$$

Proposition 229. *If $(A, <_A)$ and $(B, <_B)$ are linear orders, then the continuous functions from A to B are exactly the (not necessarily strict) order preserving functions from A to B, i.e. functions $F : A \to B$ that suffice the condition*

$a \leq_A a' \to F(a) \leq_B F(a')$ for all $a, a' \in A$[4]*.*

Proof. We assume that $F : A \to B$ satisfies $\forall a, a' \, (a \leq_A a' \to F(a) \leq_B F(a'))$ and we fix some initial segment Y of $(B, <_B)$. If $X = F''Y$ is empty, then it is initial in $(A, <_A)$ and we are done. Otherwise let $a \in X$ and $a' < a$ then, since F is order preserving and Y is initial in $(B, <_B)$, we know that $F(a') \in Y$ and thus that $a' \in X$. For the converse implication, assume that F is continuous and that $a \leq a'$ for some elements of A. Since $Y = B_{\leq_B F(a')}$ is an initial segment of $(B, <_B)$ so is $F''Y$. Because $a' \in F''Y$ and $a \leq a'$, we conclude that $a \in Y$ and thus that $F(a) \leq F(a')$. $\square$

Corollary 230. *Let $(A, <_A)$ and $(B, <_B)$ be linear orders and let $F : A \to B$. The following are equivalent.*

1. *The function F is an embedding of $(A, <_A)$ into $(B, <_B)$.*
2. *The function F is one-to-one, open and continuous.*

In view of the above proposition, continuous functions between any given linear orders $(A, <_A)$ and $(B, <_B)$ exist, and trivial examples are constant mappings. On the other hand, continuous one-to-one functions are strictly order preserving and thus, as we have seen before, too restrictive as comparison maps for anything beyond well orders. Moreover, even if we relax the one-to-one condition to the point where all functions $F : A \to B$ with the property that there is no element $b \in B$ such that $F'b$ is a proper class are considered, the resulting comparability principle is still inconsistent. A counter example is presented in the following remark.

Remark 231. Let A be the class $\omega \times On$ and fix

$$(n, \alpha) <_1 (m, \beta) \Leftrightarrow n < m \vee (n = m \wedge \alpha < \beta)$$

and

$$(n, \alpha) <_2 (m, \beta) \Leftrightarrow n > m \vee (n = m \wedge \alpha < \beta).$$

[4]Note that the defining implication for non strict order preservation is required only from left to right but not necessarily from right to left. A function $F : A \to B$ that has the property $F(a) \leq_B F(a') \to a \leq_A a'$ for all a, a' in A, is in fact already strictly order preserving and thus one-to-one, which is not necessarily the case for non strict order preserving functions.

Both $(A, <_1)$ and $(A, <_2)$ are weak well orders, but every continuous function from one order into the other will map some proper subclass of A to a single point.

Proof. Let $F : A \to A$ be such that $\forall a \in A\,(F'a \in \mathcal{P}(A))$. First, assume that F is continuous from $(A, <_1)$ to $(A, <_2)$. First, note that for all ordinals α, β and all natural numbers n, k with $n < k$ we have that $k' < n'$ if α', β', n' and k' are such that $F(n, \alpha) = (n', \alpha')$ and $F(k, \beta) = (k', \beta')$. Now, since $(0, 0)$ is the least element of $(A, <_1)$, a natural number n such that $F[A] \subset \{(k, \alpha) \mid k \leq n \wedge \alpha \in On\}$ exists. This contradicts our previous observation. Now assume that F is continuous from $(A, <_2)$ to $(A, <_1)$. Similarly as in the previous case, the elements of $X = \{(n, 0) \mid n \in \omega\}$ form an infinite descending chain in $(A, <_2)$. Since for any two distinct elements $a <_2 b$ of X we have that $\{x \in A \mid a <_2 x \wedge x <_2 b\}$ is a proper class, we know that $F \restriction X$ is one-to-one, contradicting the fact that $(A, <_1)$ is a well order. $\square$

Now that we have seen that continuous (i.e. non strict order preserving) functions are no viable candidates for comparison functions between weak well orders, we now turn to investigate to which degree open functions can be tuned into a more suitable alternative. First, a definition:

Definition 232. Let $(A, <_A)$ and $(B, <_B)$ be linear orders. We call a function $F : A \to B$ is *weakly order preserving* if

$$a <_A F'b \to F(a) <_B b$$

for all $a \in A$ and all $b \in B$.

We will continue with proving a few basic facts in order to display which key properties of (strict) order preserving functions are conserved in the process of weakening the notion in the sense of Definition 232. We will subsequently elaborate on some general facts of these functions and then discuss how they could be applied to obtain comparability principles of weak well orders.

First, we note that in the one-to-one case, weakly order preserving functions are already (strict) order preserving.

Lemma 233. *Let $(A, <_A)$ and $(B, <_B)$ be linear orders. Every weakly order preserving one-to-one function $F : A \to B$ is already strictly order preserving, i.e. $\forall a, b \in A\, (a <_A b \leftrightarrow F(a) <_B F(b))$.*

Proof. Let $(A, <_A)$, $(B, <_B)$ and $F : A \to B$ be as in the claim. It is enough to show that for all $a, b \in A$, we have $a <_A b \to F(a) <_B F(b)$. Since F is one-to-one, we can write $\{b\}$ as $F'F(b)$ for every element $b \in A$. Thus we get that

$$a <_A b \Leftrightarrow a <_A F'F(b) \Rightarrow F(a) <_B F(b)$$

for all $a \in A$, as desired. □

Lemma 234. *If $F : A \to B$ is some weakly order preserving function between linear orders $(A, <_A)$ and $(B, <_B)$, then*

$$a <_A F''Y \;\to\; F(a) <_B Y$$

holds for all $a \in A$ and all classes $Y \subset B$.

Proof. Let F be a weakly order preserving function from $(A, <_A)$ to $(B, <_B)$. Assume that $a <_A F''Y$ for some class $Y \subset B$. Consider

$$\begin{aligned} a <_A F''Y \;&\Rightarrow\; \forall y \in Y\, (a <_A F'y) \\ &\Rightarrow\; \forall y \in Y\, (F(a) <_B y) \\ &\Rightarrow\; F(a) <_B Y. \end{aligned}$$ □

As a consequence, we can prove that under weakly order preserving functions, preimages of minimal subclasses are minimal. More precisely:

Corollary 235. *Let F be a weakly order preserving function from $(A, <_A)$ to $(B, <_B)$. For all subclasses $M, Y \subset B$ we have that $F''M$ is minimal in $F''Y$ whenever M is minimal in Y.*

Proof. Let $M \subset Y$ be some subclasses of B. If $F''M \subset F''Y$ is not minimal, then there is an element $a \in F''Y$ below $F''M$, i.e. some set a satisfies $F(a) \in Y$ and $a < F''M$. Applying that F is weakly order preserving, we obtain that $F(a) \in Y$ and $F(a) < M$, thus M cannot be a minimal subclass of Y. □

The purpose of the following series of lemmas is to gather some more general facts about weakly order preserving functions and thus establish a suitable intuition for our further studies.

Lemma 236. *If* $(A, <_A)$ *and* $(B, <_B)$ *are linear orders and* $F : A \to B$ *is weakly order preserving, then every restriction* $F \upharpoonright X$ *of* F *to any initial segment* X *of* $(A, <_A)$ *is weakly order preserving.*

Proof. Without loss of generality, we assume that $X \neq \varnothing$. Let $a <_A (F \upharpoonright X)'b$ with $a \in X$ and $b \in B$. We apply $(F \upharpoonright X)'b = X \cap F'b$, and we distinguish the two cases where $F'b \cap X$ is and is not empty respectively. If $F'b \cap X = \varnothing$, then because X is an initial segment of $(A, <_A)$, we know that $a <_A F'b$ and thus that $F(a) <_B b$. In case that $F'b \cap X$ is not empty, it is minimal in $F'b$ and thus, since $a <_A F'b \cap X$, we know that $a <_A F'b$ so $F(a) <_B b$ as desired. □

As a result of this lemma, we obtain a proof that that weakly order preserving functions correspond exactly to open functions.

Lemma 237. *If* $(A, <_A)$ *and* $(B, <_B)$ *are linear orders, then every weakly order preserving function* $F : A \to B$ *is also open.*

Proof. First, we note that the image of any weakly order preserving function is an initial segment of the target order. To see this, assume that $F : A \to B$ is weakly order preserving. If $a \in A$ is arbitrarily chosen and $b \in B$ is not an element of $F[A]$, then $F'b = \varnothing$ and thus $a <_A F'b$, therefore $F(a) <_B b$. Hence, $F[A] = \{x \in B \mid \forall b \in B \backslash F[A]\, (x <_B b)\}$ is an initial segment of $(B, <_B)$. Now, if X is any initial segment of $(A, <_A)$, then we know from Lemma 236 that $F \upharpoonright X$ is a weakly order preserving and thus that $F[X] = (F \upharpoonright X)[X]$ is an initial segment of B. □

Lemma 238. *Let* $(A, <_A)$ *and* $(B, <_B)$ *be linear orders. Every open function* $F : A \to B$ *is weakly order preserving.*

Proof. Let $a \in A$ and $b \in B$ be such that $a <_A F'b$ and assume that $F : A \to B$ is open. We have to show that $F(a) <_B b$. If $X = A_{<_A F'b}$, then, because F is open, $Y = F[X]$ is an initial segment of $(B, <_B)$ with $F(a) \in Y$ and $b \notin Y$, thus $F(a) <_B b$. □

Corollary 239. *Let $(A, <_A)$ and $(B, <_B)$ be linear orders and let $F : A \to B$, then F is weakly order preserving if and only if F is open.*

From the fact that weakly order preserving functions are open, it is an immediate consequence that the composition of weakly order preserving functions are weakly order preserving.

Corollary 240. *If $(A, <_A)$, $(B, <_B)$ and $(C, <_C)$ are linear orders and if further $F : A \to B$ and $G : B \to C$ are weakly order preserving, then $G \circ F : A \to C$ is weakly order preserving between the orders $(A, <_A)$ and $(C, <_C)$.*

Combining corollary 239 and Lemma 233, we obtain that weakly order preserving functions that are one-to-one are already embeddings.

Corollary 241. *Let $(A, <_A)$ and $(B, <_B)$ be linear orders and let $F : A \to B$. The following statements are equivalent:*

1. *F is an embedding of $(A, <_A)$ into $(B, <_B)$*

2. *F is weakly order preserving and one-to-one*

In view of the previous lemma, we can think of approaching the notion of order preserving functions with weakly order preserving functions by tuning the degree to which they are allowed to map different elements to the same image. A very permissive approach is presented in the next definition, when we fix the notion of a *set to one* function to be a function F whose fibers, i.e. classes of the form $F'b$ for some b, are all sets.

Definition 242. A function $F : X \to Y$ is *set to one* if $\forall y \in Y\ (F'y \in \mathcal{P}(X))$. A *weak embedding* of some linear order $(A, <_A)$ into some other linear order $(B, <_B)$ is a function $F : A \to B$ that is both weakly order preserving and set to one.

Lemma 243. *The restriction of any weak embedding $F : A \to B$ of some linear order $(A, <_A)$ onto some other linear order $(B, <_B)$ to some initial segment X of $(A, <_A)$ is a weak embedding $F \restriction X : X \to B$ of $(X, <_X)$ into $(B, <_B)$.*

Proof. This follows directly from Lemma 236. $\square$

Lemma 244. *For all linear orders* $(A, <_A)$, $(B, <_B)$ *and* $(C, <_C)$ *and all weak embeddings* $F : A \to B$ *and* $H : B \to C$*, the composite* $H \circ F : A \to C$ *is a weak embedding of* $(A, <_A)$ *into* $(C, <_C)$.

Proof. Let $F : A \to B$ and $H : B \to C$ be weak embeddings as indicated in the claim. First note that for every element $c \in C$ the class

$$(H \circ F)'c = \{a \in A \mid H(F(a)) = c\}$$

is the union of set many sets and thus is itself as a set. To see that $H \circ F$ is weakly order preserving, it is enough to verify that the function is open. Let X be any initial segment of $(A, <_A)$. Since F is open, $F[X]$ is an initial segment of $(B, <_B)$ and because H is also open we can conclude that $H[F[X]] = (H \circ F)[X]$ is an initial segment of $(C, <_C)$. □

Now that we have established some basic properties of weak embeddings, the stage is set in order to see how far we can get with weak embeddings in terms of comparing weak well orders.

Theorem 245. *For every cardinal* κ*, if* $(V_\kappa, V_{\kappa+1})$ *is a model of* NBG*, then* $(V_\kappa, V_{\kappa+1})$ *is also a model of*

$$\mathsf{wwo}(A, <_A) \rightarrow \exists (W, \prec) \exists F \left(\mathsf{wo}(W, \prec) \wedge F : (A, <_A) \xrightarrow{w.e.} (W, \prec) \right),$$

where $F : (A, <_A) \xrightarrow{w.e.} (W, \prec)$ *is an abbreviation for the statement that* F *is a weak embedding of* $(A, <_A)$ *into* $(W, \prec)$.

Proof. Let $\mathfrak{M} = (V_\kappa, V_{\kappa+1})$ be a model of NBG. Clearly, a weak well order $(A, <_A)$ in $\mathfrak{M}$ [5] is a κ-well order of cardinality at most κ. We define a sequence $\{m_\alpha \mid \alpha \in On\}$ by transfinite recursion as follows:

$$m_\alpha = \begin{cases} \min_\lhd \{m \subset A \mid m \text{ is minimal in } A \setminus \bigcup_{\mu<\alpha} m_\mu\} & \text{if } \{\dots\} \neq \varnothing \\ \varnothing & \text{otherwise.} \end{cases}$$

[5] In spite of the fact that A and F are sets, but for the benefit of readability, we use upper case letters here.

Since we can assume without loss of generality that $|A| = \kappa$ and because the elements of $\{m_\alpha \mid \alpha \in On\}$ are pairwise disjoint and $A = \bigcup_\alpha m_\alpha$ holds, there is an ordinal λ with $\kappa \leq \lambda = \{\alpha \in On \mid m_\alpha \neq \varnothing\} < \kappa^+$, where κ^+ denotes the least cardinal greater than κ. Therefore, there is a bijective function $h : V_\kappa \to \lambda$ based on which we introduce the well order $(V_\kappa, <_h)$ where $x <_h y \leftrightarrow h(x) < h(y)$. Since, in $\mathfrak{M}$, the function $F : A \to V_\kappa$ as given from

$$F(a) = x \leftrightarrow \exists\alpha\,(h(x) = \alpha \wedge a \in m_\alpha).$$

is a weak embedding of $(A, <_A)$ into $(V, <_h)$, this proves the claim. $\square$

As a consequence of Theorem 245, we get that the consistency of NBG together with the formula

$$\mathsf{wwo}(A, <_A) \to \exists(W, \prec)\,\exists F\,\big(\mathsf{wo}(W, \prec) \wedge F : (A, <_A) \xrightarrow{w.e.} (W, \prec)\big), \tag{3.24}$$

can be proved within ZFC+ *"inaccessible cardinals exist"*. Thus, if anything else, a theory based on (3.24) can be assumed consistent. Moreover, as we will see later (cf. Definition 252), assuming a formula such as presented in (3.24) can be seen as an intermediate step towards a machinery to match weak well orders with well orders. Of course, (3.24) only guarantees that for every weak well order $(A, <_A)$ a nonempty collection of well orders into which $(A, <_A)$ can be weakly embedded exists, it has to be further clarified how these well orderings can be used to propose a well order to associate with $(A, <_A)$.

As we have seen in Theorem 185, if the theory Δ^1_0-TR is extended through weakening of the well foundedness precondition in the schema Δ^1_0-$\mathsf{TR}_{\mathsf{wo}}$, then the resulting theory is inconsistent. However, utilizing the formula (3.24) as an additional axiom might strengthen the theory Δ^1_0-TR by way of extending the range of *well orders* to which Δ^1_0-$\mathsf{TR}_{\mathsf{wo}}$ could be applied. More specifically, working in (3.24)+Δ^1_0-TR and given any weak well order $(A, <_A)$, one can find a class H that is the result of iterating a given elementary formula along a well order into which $(A, <_A)$ can be weakly embedded. More formally, if $\varphi(X^+, x)$[6]

[6]The superscript + in $\varphi(X^+, x)$ indicates that the parameter X might only occur positively in φ.

and $\psi(H, F, (W, \prec))$ are elementary formulas and $\mathsf{hier}_\varphi(H, (W, \prec))$ stands for the formula

$$\forall x, y \in W \big(x \prec y \to (H)_x \subset (H)_y \wedge (H)_x = \bigcup_{z \prec x} \Gamma_\varphi\big((H)_z\big)\big),$$

then one can prove from $\mathsf{wwo}(A, <_A)$ (with an application of Lemma 210) that

$$\exists (W, \prec), F, H \big(F : (A, <_A) \xrightarrow{w.e.} (W, \prec) \wedge \mathsf{hier}_\varphi(H, (W, \prec)) \wedge \psi(H, F, (W, \prec))\big) \tag{3.25}$$

holds. Since (3.25) is a Σ^1_1 formula, we conclude that there must also be a linear order $(A, <_A)$ that is *not* a weak well order but still satisfies (3.25). At first glance, it seems as if a particularly clever combination of the facts that $F : (A, <_A) \xrightarrow{w.e.} (W, \prec)$ and that $(A, \prec)$ is not weakly well founded might reveal the existence of a pseudo hierarchy[7] and along with it a fixed point for the operator Γ_φ. However, this is not the case as it is always possible to choose the formula ψ such that $\psi(H, F, (W, \prec)) \to \mathsf{wo}(W, \prec)$. In fact, we obtain the following proposition as a byproduct of the discussion above.

Proposition 246. *It is not provable in* NBG *that for all* $F, (A, <_A)$ *and* $(W, \prec)$ *it is the case that*

$$\big(\mathsf{wo}(W, \prec) \wedge F : (A, <_A) \xrightarrow{w.e.} (W, \prec)\big) \to \mathsf{wwo}(A, <_A)$$

holds. In fact, if NBG *is extended by (3.24), then it is provable that there are linear orders* $(A, <_A)$ *and* $(B, <_B)$ *and a function* F *such that* $\neg\mathsf{wwo}(A, <_A)$ *and* $\mathsf{wo}(B, <_B)$ *and* $F : (A, <_A) \xrightarrow{w.e.} (B, <_B)$.

While these observations make it clear that a pseudo hierarchy argument to obtain fixed points for Γ_φ cannot be obtained as described above, that does not mean that a proper hierarchy with a fixed point of Γ_φ does not exist. In order to prove the existence of a proper hierarchy which contains fixed points of Γ_φ, one needs transfinite recursion and a well order $(W, \prec)$ that is long enough so that a fixed point is eventually reached when Γ_φ is iterated along $(W, \prec)$. Although

[7] A class H such that a linear order $(B, <_B)$ with $\mathsf{hier}_\varphi(H, (B, <_B)) \wedge \neg\mathsf{wo}(B, <_B)$ exists.

possible in principle[8], in view of the next proposition, this is not a consequence of (3.24).

Proposition 247. *Let $(A, <_A)$ be any linear order. If there is a well order $(W, \prec)$ and a function $F : (A, <_A) \overset{w.e.}{\longrightarrow} (W, \prec)$, then there is a weak embedding $F^\star : (A, <_A) \overset{w.e.}{\longrightarrow} (On, \in)$.*

Proof. Let $F, (W, \prec)$ and $(A, <_A)$ be as in the claim. By Corollary 111, we can assume without loss of generality that $(W, \prec) = (On, \in) \oplus (X, \prec)$ for some well order $(X, \prec)$ with $X \cap On = \varnothing$. Further, we can assume that $\{x_\alpha \mid \alpha \in On\} = X$ be any enumeration of X. We use the weak embedding F to define a function $F^\star : A \to On$ as follows

$$F^\star(a) = \begin{cases} F(a) & \text{if } F(a) \in On \\ \alpha & \text{if } F(a) = x_\alpha. \end{cases}$$

First, note that $F^\star$ is a set to one function. To see that $F^\star$ is also weakly order preserving, it suffices to show that it is open. This follows from the fact that for every initial segment Y of $(A, <_A)$, we have that $F[Y] \cap On = F^\star[Y]$ holds. Thus, $F^\star : (A, <_A) \overset{w.e.}{\longrightarrow} (On, \in)$ as desired. □

Corollary 248. *Every well order can be weakly embedded into $(On, \in)$.*

Informally, in the proof of Theorem 245 we have seen that proving the existence of a weak embedding from a weak well order $(A, <_A)$ into some given well order $(W, \prec)$ essentially amounts to "consuming" the whole field of $(A, <_A)$ by removing one minimal subset after another, while $(W, \prec)$ enumerates all the so collected subsets. That is not unlike the situation when proving comparability of well orders; minimal elements are removed from one order and matched to increasing elements of the other order until there is either nothing left to remove or to match respectively. From an outside perspective[9], although the procedures are similar,

[8] Note that by Theorem 209, the existence of fixed points is equivalent to the existence of proper hierarchies that contain fixed points.

[9] When we say that a well order $(W, \prec)$ is of at least the same length as a well order $(W', \prec')$, then we mean to speak in the meta theory and say that the ordinal that corresponds to the order-type of $(W, \prec)$ is greater or equal than the corresponding ordinal of $(W', \prec')$, thus the phrase "from an outside perspective".

a well order can only be embedded into a well order of at least the same length, this is not the case for weak embeddings as demonstrated in Corollary 248. We will now pin down the underlying reason for that difference and then present a more effective version of (3.24). First, an observation:

Remark 249. Let $(w, \prec)$ be a well order (that is a set). The order-type of $(w, \prec)$ is the supremum of all ordinals α such that a weakly order preserving function $f : (w, \prec) \to (\alpha, \in)$ exists.

Proof. Let α_w be the order-type of $(w, \prec)$ and let α_s be the supremum as described in the claim. Since there is an order isomorphism (and thus a weakly order preserving function) between $(w, \prec)$ and $(\alpha, \in)$, we know that $\alpha_w \leq \alpha_s$. For the converse inequality, assume that α is some ordinal such that a weakly order preserving function $f : w \to \alpha$ exists. We have to show that the order-type of $(w, \prec)$ is greater or equal than α. Since $(w, \prec)$ is a well order and f is weakly order preserving, we can define a strictly increasing sequence $\{x_\mu \mid \mu < \alpha\}$ in $(w, \prec)$ by stipulating $x_\mu = \min_\prec f'\mu$. Thus, the order-type of $(w, \prec)$ is at least α. $\square$

From an outside perspective, the previous remark tells us that the order-type of a well order $(W, \prec)$ can be seen as the supremum of all ordinals into which $(W, \prec)$ can be weakly embedded. The reason that this characterization is needlessly complicated when one is only interested in order-types of well orders, is the fact that, in case of well orders, the aforementioned supremum actually is a maximum. Thus, it is much more natural to directly describe the order-type of a well order $(W, \prec)$ as this maximum; the least ordinal that is order isomorphic to $(W, \prec)$. However, in case of weak well orders the supremum mentioned before is in general[10] not a maximum. Thus, in order to formalize the notion of order-types for weak well orders, we will have to use the characterization given in the previous remark. First, in Definition 250, we give an explicit meaning to the phrase "removing minimal subsets of a weak well order up to the point where the remaining part is empty". Thereafter, in Proposition 251 we give a formal variant of our observation that constructing a weak embedding of a weak well

[10] In fact, the supremum described above is a maximum exactly in case that the order is a well founded

order into a well order basically consists of removing minimal subsets of the weak well order up to the point where the remaining part is empty, while the well order enumerates the so collected sets in order of their removal.

In the remainder of this chapter, we work in the theory Δ^1_0-TR.

Definition 250. Let $(A, <)$ be a linear order and $(W, \prec)$ a well order. Further, let $\lessdot$ be any global well ordering. We define the function

$$\mathfrak{F}_{\lessdot}((W, \prec), (A, <_A)) : W \to \mathcal{P}(A)$$

by recursively stipulating

$$F(x) = \begin{cases} \min_{\lessdot}\{m \subset A \mid m \text{ is minimal in } A \setminus \cup F[W_{\prec x}]\} & \text{if } \{\dots\} \neq \varnothing \\ \varnothing & \text{otherwise,} \end{cases}$$

where F is an abbreviation for $\mathfrak{F}_{\lessdot}((W, \prec), (A, <_A))$. In the following, if there is no ambiguity, we will write $\mathfrak{F}_{\lessdot}$ to mean $\mathfrak{F}_{\lessdot}((W, \prec), (A, <_A))$.

Proposition 251. *Let $(A, <)$ be a weak well order and $(W, \prec)$ a well order. The following assertions are equivalent for every function $F : A \to W$.*

1. $F : (A, <_A) \xrightarrow{w.e.} (W, \prec)$
2. *There is a global well ordering $\lessdot$ such that $\forall x \in W \left(F'x = \mathfrak{F}_{\lessdot}(x)\right)$ and $A = \cup(\mathfrak{F}_{\lessdot}[W])$*

Proof. First, assume that $F : (A, <_A) \xrightarrow{w.e.} (W, \prec)$. Let $(X, <_X)$ be the well ordering where $X = \{F'x \mid x \in W\}$ and $F'x <_X F'y \leftrightarrow x \prec y$. By Theorem 104, we can extend $<_X$ to a global well order $\lessdot$ of which X is an initial segment. It is now easy to verify by induction along $(W, \prec)$ that $\mathfrak{F}_{\lessdot}(x) = F'x$ holds for all $x \in W$. For the converse direction, note that for every global well ordering $\lessdot$, the assertion $F(a) = x \leftrightarrow a \in \mathfrak{F}_{\lessdot}(x)$ defines a unique function F with domain A and that this function is a weak embedding[11]. □

The now formalized process of constructing a weak embedding of a non well founded weak well order into a well order is erratic in the sense that the method

[11] Compare also to the proof of Theorem 245, where the sequence $\{m_\alpha \mid \alpha \in On\}$ acts in the place of the function $\mathfrak{F}_{\lessdot}$.

of how to choose the "next" subset to remove is guided by the axiom of choice. For well orders, on the other hand, the "next" minimal subset to be removed can always be chosen "canonically" as the (uniquely determined) singleton that contains the minimal element of the remaining class. The crucial property of this specific strategy to remove minimal subsets from a well order, is that the ordinal needed to enumerate these subsets in the order of their removal is maximal. We will now present our order-type system that resolves this problem at the cost of a second order universal quantification.

Definition 252 (order-type of weak well orders)**.** Let $(A, <_A)$ be any linear order and let $(W, \prec)$ be a well order. If for every global well ordering $\lessdot$ it is the case that $\cup\mathfrak{F}_{\lessdot}[W] = A$ holds, then we write $\mathsf{ot}(A, <_A) \leq (W, \prec)$. If further there is no element $x \in W$ such that $\mathsf{ot}(A, <_A) \leq (W_{\prec x}, \prec)$ holds, then we say that $(A, <_A)$ is of order-type $(W, \prec)$ and we write $\mathsf{ot}(A, <_A) \simeq (W, \prec)$.

Proposition 253. *The following assertions can be proved in* Δ^1_0-TR*.*

1. *For every well order* $(W, \prec)$ *we have that* $\mathsf{ot}(W, \prec) \simeq (W, \prec)$*.*
2. *If* $(A, <_A)$ *and* $(B, <_B)$ *are order isomorphic and* $\mathsf{ot}(A, <_A) \simeq (W, \prec)$ *and* $\mathsf{ot}(B, <_B) \simeq (W', \prec')$ *then* $(W, \prec)$ *and* $(W', \prec')$ *are order isomorphic.*
3. *Order-types of weak well orders are unique up to order isomorphism.*

Proof. For *1.*, let $(W, \prec)$ be any well order. By Proposition 251, there is a global well order $\lessdot_1$ such that $\mathfrak{F}_{\lessdot_1}(x) = \{x\}$. We have to prove the following two statements:

i) For all global well orderings $\lessdot$ it is the case that $\cup\mathfrak{F}_{\lessdot}[W] \supset W$.

ii) For all $x \in W$ there is a global well ordering $\lessdot$ such that $\cup\mathfrak{F}_{\lessdot}[W_{\prec x}] \neq W$.

Statement ii) follows from the fact that for every $x \in W$ it is the case that $x \notin \cup\mathfrak{F}_{\lessdot_1}[W_{\prec x}]$. Claim i) follows from the fact that for every global well ordering $\lessdot$ we have that $W \subset \cup\mathfrak{F}_{\lessdot_1}[W] \subset \cup\mathfrak{F}_{\lessdot}[W]$. For the second part of the proposition, let $(A, <_A)$, $(B, <_B)$, $(W, \prec)$ and $(W', \prec')$ be as in the claim. Since Δ^1_0-TR proves the comparability of well orders, we can assume without loss of generality that $(W', \prec')$ can be embedded in $(W, \prec)$. It is thus enough to show that there is no element $x \in W$ such that $(W_{\prec x}, \prec)$ is order isomorphic to $(W', \prec')$. By way of contradiction assume such an $x \in W$ exists. Let $H : W' \to W_{\prec x}$ and $F : A \to B$ be order isomorphisms. Since $\mathsf{ot}(A, <_A) \simeq (W, \prec)$, there exists a global well ordering $\lessdot$ such that an $a \in A$ with $a \notin \cup\mathfrak{F}_{\lessdot}[W_{\prec x}]$ exists[12]. Since minimality of sets is preserved under order isomorphism, there is a global well ordering $\lessdot_2$ such that for every y in W', $\mathfrak{F}_{\lessdot_2}(y) = F[\mathfrak{F}_{\lessdot}(H(y))]$ holds. From $\mathsf{ot}(B, <_B) \simeq (W', \prec')$ it follows that $F(a) \in \cup\mathfrak{F}_{\lessdot_2}[W']$. Now consider:

$$\begin{aligned} F(a) \in \mathfrak{F}_{\lessdot_2}[W'] &\Rightarrow \exists y \in W' \,(F(a) \in \mathfrak{F}_{\lessdot_2}(y)) \\ &\Rightarrow F(a) \in F[\mathfrak{F}_{\lessdot_1}(H(y))] \\ &\Rightarrow a \in \mathfrak{F}_{\lessdot_1}(H(y)) \\ &\Rightarrow a \in \cup\mathfrak{F}_{\lessdot_1}[W_{\prec x}], \end{aligned}$$

in contradiction to our choice of a. The third part of the proposition is an immediate consequence of the second. $\square$

In the next proposition, we observe that weak well orders are the only linear orders that can possibly have an order-type in the sense of Definition 252.

Proposition 254. *It is provable in Δ^1_0-TR that for all linear orders $(A, <_A)$,*

$$\exists (W, \prec)\, (\mathsf{wo}(W, \prec) \wedge \mathsf{ot}(A, <_A) \leq (W, \prec))$$

implies that $(A, <_A)$ is a weak well order.

Proof. Let $(A, <_A)$ be a linear order such that $\mathsf{ot}(A, <_A) \leq (W, \prec)$ for some well order $(W, \prec)$. Let X be any nonempty coinitial subclass of A. By Theorem 104, a global well ordering $\lessdot$ that satisfies

$$\forall x, y \left((x \cap X = \varnothing \wedge y \cap X \neq \varnothing) \to x \lessdot y\right)$$

exists. Let x be a element of X. It follows from our choice of $\lessdot$ and the definition of $\mathfrak{F}_{\lessdot}$ that for all $y \in W$ either $\mathfrak{F}_{\lessdot}(y) \subset X$ or $\mathfrak{F}_{\lessdot}(y) \cap X = \varnothing$. Thus, $x \in \cup\mathfrak{F}_{\lessdot}[W]$

[12] Note that, because of the uniqueness of hierarchies defined by transfinite recursion, there is no need to distinguish between the functions $\mathfrak{F}_{\lessdot}((W, \prec), (A, <_A))$ and $\mathfrak{F}_{\lessdot}((W_{\prec x}), (A, <_A))$.

implies that the class $\{z \in W \mid \mathfrak{F}_{\lessdot}(z) \subset X\}$ is nonempty and therefore has a least element, say z_X. We claim that $\mathfrak{F}_{\lessdot}(z_X)$ is a minimal subset of X. It follows from the definition of $\mathfrak{F}_{\lessdot}$ that $\mathfrak{F}_{\lessdot}(z_X)$ is a minimal (w.r.t. $(A, <)$) subset of $A \backslash \mathfrak{F}_{\lessdot}[W_{\prec z_X}]$. Since $\mathfrak{F}_{\lessdot}[W_{\prec z_X}] \subset A_{<X}$, this implies that $\mathfrak{F}_{\lessdot}(z_X)$ is minimal in X, as desired. □

Definition 255. Let CWO^+ stand for Δ^1_0-TR plus

$$\mathsf{wwo}(A, <_A) \to \exists (W, \prec)\, (\mathsf{wo}(W, \prec) \wedge \mathsf{ot}(A, <_A) \simeq (W, \prec)).$$

Theorem 256. *It is provable in* CWO^+ *that the following two assertions are equivalent:*

1. $\mathsf{wwo}(A, <)$
2. $\exists (W, \prec)\, (\mathsf{wo}(W, \prec) \wedge \mathsf{ot}(A, <) \simeq (W, \prec))$

Proof. The direction from *1.* to *2.* is CWO^+ and the converse is a direct consequence of Proposition 254. □

Corollary 257. *It is provable in* CWO^+ *that there is no* Σ^1_1 *formula* $\varphi(X, Y)$ *that satisfies*

$$\forall (A, <_A), (W, \prec) \big(\mathsf{ot}(A) \simeq (W, \prec) \longleftrightarrow \varphi((A, <_A), (W, \prec))\big).$$

Summary

In the first part, we introduced the theory CWO that guarantees that any two well orders are comparable. In Theorem 225 a straightforward proof that the theory Δ^1_0-TR includes CWO was presented. The second part was devoted to an attempt to interpret CWO in terms of weak well orders. First, we observed that for the purpose of comparing anything not well founded, order isomorphisms and embeddings are too restrictive. We continued to analyze which properties of embeddings have to be weakened in order to obtain functions more suitable for the task. We introduced weakly order preserving functions and weak embeddings in Definition 232 and Definition 242. After we proved some general facts about weak embeddings, we showed in Theorem 245 that weak embeddings potentially exist not only between weak well orders but also that it is consistent to assume that all weak well orders can be weakly embedded into well orders. However, in Proposition 247 we gave evidence for the fact that merely assuming the existence of weak embeddings from all weak well orders into well orders does not constitute a strong theory. Subsequently we introduced our notion of order-types for weak well orders (cf. Definition 252); the order-type of a weak well order can be seen as the supremum of all ordinals in which $(A, <_A)$ can be weakly embedded. In Proposition 253 we showed that our framework satisfies the most basic properties that one would expect from an order-type system. In Definition 255 we introduced the theory CWO^+ which consists of Δ^1_0-TR together with the assertion that every weak well order has an order-type. In Theorem 256 we proved that CWO^+ is as strong as possible in the sense that if a linear order has an order type, then it is already a weak well order, i.e. the axioms of CWO^+ cannot be extended to linear orders that are not weakly well founded. Finally, the big question that remains open is whether or not FP is contained in CWO^+, thus the following conjecture:

Conjecture. *It is provable in* CWO *that for every monotone elementary operator* Γ *there is a well order* $(W, \prec)$*, such that* Γ *reaches a fixed point when iterated along* $(W, \prec)$.

3.4 Bar induction and V-model reflection

In this section, we will provide further evidence that the situation in class set theory with regard to the relative strength of theories that fundamentally rely on the notion of well foundedness is quite different from the situation in arithmetic. Specifically, we will see that the relationship of corresponding principles of bar induction and ω-model reflection is differs from their relationship in arithmetic. Again, the fact that in set theory, as opposed to arithmetic, the well foundedness of a relation can be expressed as an elementary formula constitutes the dissimilarity between the two settings.

After introducing the notion of V-models, a straightforward translation of ω-models to sets and classes, we introduce the axiom of V-model reflection, which corresponds to ω-model reflection. In a further step, we will present two versions of bar induction, where the first is a verbatim quote of the usual bar induction schema from arithmetic, and the second principle is obtained by generalizing the weak induction principles from classes to predicates. While we cannot give a conclusive argument for which of the two principles is "the right" analog, we will see that, despite the fact that bar induction with weak well orders entails the usual schema of bar induction, both principles yield theories that fall short of the power of their arithmetic counterpart, at least with respect to their capacities relative to V-model reflection.

Definition 258. Let M be any class, and let φ be any formula. We define the formula $\varphi^{(M)}$ by induction on the complexity of φ as follows:

1. If φ is an elementary formula, then $\varphi^{(M)}$ is φ.
2. If φ is $\exists X\,(\psi(X))$, then $\varphi^{(M)}$ is $\exists\alpha\,\big(\psi((M)_\alpha)\big)^{(M)}$.
3. If φ is $\forall X\,(\psi(X))$. then $\varphi^{(M)}$ is $\forall\alpha\,\big(\psi((M)_\alpha)\big)^{(M)}$.
4. If φ is $\varphi_1 * \varphi_2$ for any of the connectives $* \in \{\wedge, \vee, \rightarrow\}$, then $\varphi^{(M)}$ is $\varphi_1^{(M)} * \varphi_2^{(M)}$.
5. If φ is $\neg\psi$, then $\varphi^{(M)}$ is $\neg\,(\psi^{(M)})$.
6. If φ is $\exists x\,\psi$, then $\varphi^{(M)}$ is $\exists x\,(\psi^{(M)})$.

7. If φ is $\forall x\, \psi$, then $\varphi^{(M)}$ is $\forall x\, (\psi^{(M)})$.

We will use the shortcut $X \dot{\in} M$ to mean $(\exists Y\, (X = Y))^{(M)}$, i.e $\exists \alpha\, (X = (M)_\alpha)$.

Stipulating Φ_{NBG} as the conjunction of all formulas that constitute an arbitrary but fixed finite axiomatization of the theory NBG, we introduce V-models, in analogy to ω-models of arithmetic, as classes in which Φ_{NBG} holds[13].

Definition 259. A *V-model* is a class M such that $\Phi_{\mathsf{NBG}}^{(M)}$ holds. For any collection $\mathcal{F}$ of formulas, we say that a class M is a *V-model of $\mathcal{F}$* if M is a V-model that satisfies $\varphi^{(M)}$ for all formulas φ in $\mathcal{F}$.

The idea behind V-model reflection is to assume that for some formulas $\varphi(X)$ and for all classes X that satisfy $\varphi(X)$, a V-model that contains X and models $\varphi(X)$ exists.

Definition 260. Let $\mathcal{F}$ be a collection of formulas. The theory $\mathcal{F}$-RFN, of *V-model reflection* for formulas in $\mathcal{F}$ is obtained from extending NBG by the schema

$$\varphi(X) \to \exists M \left(X \dot{\in} M \wedge \Phi_{\mathsf{NBG}}^{(M)} \wedge \varphi^{(M)} \right),$$

where $\varphi(X)$ ranges over formulas in $\mathcal{F}$ with at most X free. We will use the shorthand notation RFN for the theory $\{\exists x\, (x = x)\}$-RFN, i.e. the assumption that a V-model exists.

Lemma 261. *If $\mathcal{F}$ is any extension of the theory* NBG *and if it is provable in* NBG *that every V-model is also a model of $\mathcal{F}$, then it is provable in* RFN *together with elementary transfinite recursion along ω that $\mathcal{F}$ is consistent.*

Proof. This is a direct consequence of Lemma 6.6 in [Sat13]. □

The general idea behind bar induction[14] is to take the transfinite induction principle (cf. Definition 97)

$$\mathsf{wo}(A, <_A) \to \forall X\, (\mathsf{prog}(X, (A, <_A)) \to A \subset X),$$

[13] The phrasing that some formula φ holds in a given class M, or equivalently, that M models φ, means that the formula $\varphi^{(M)}$ is true.

[14] The reader is advised that for reference in [Sim98], the principle corresponding to bar induction is named transfinite induction in [Sim98].

to the form

$$\mathsf{wo}(A, <_A) \to \forall \mathcal{X}\,(\mathsf{prog}(\mathcal{X}, (A, <_A)) \to A \subset \mathcal{X}),$$

where $\mathcal{X}$ ranges over extensions $\{x \mid \varphi(x)\}$ of suitable formulas $\varphi(x)$, as presented formally in the two following definitions.

Definition 262. For every formula φ, we introduce the formulas

$$\begin{aligned}\mathsf{prog}(\varphi, (A, <_A)) \;&\equiv\; \forall a \in A\,\big(\forall b\,(b <_A a \to \varphi(b)) \;\to\; \varphi(a)\big)\\ \mathsf{Ti}(\varphi, (A, <_A)) \;&\equiv\; \mathsf{prog}(\varphi, (A, <_A)) \;\to\; \forall a \in A\,\varphi(a)\end{aligned}$$

Definition 263. Let $\mathcal{F}$ be some collection of formulas. The theory $\mathcal{F}$-BI of *bar induction* for formulas in $\mathcal{F}$ extends NBG by every instance of

$$\mathsf{wo}(A, <_A) \;\to\; \mathsf{Ti}(\varphi, (A, <_A))$$

where φ is an element of $\mathcal{F}$.

As pointed out earlier, it is shown in [JS99] (cf. Theorem 22) that in the case of arithmetic, the theories Π^1_{n+1}-RFN_0 and Π^1_n-BI_0 are equivalent. The next theorem shows that the situation in set theory is quite different; highlighting the pivotal role of the fact that in our setting well orderedness can be expressed as an elementary formula.

Theorem 264. *Every V-model M of* NBG *satisfies bar induction for all formulas of the language $\mathcal{L}_2$.*

Proof. By way of contradiction, assume that M is a V-model of NBG that does not satisfy bar induction. By assumption there exists a class $(A, <)\dot{\in} M$ and a formula $\varphi(x)$ such that $\mathsf{wo}(A, <)^{(M)}$ and $(\neg(\mathsf{Ti}(\varphi, (A, <))))^{(M)}$. Since $\mathsf{wo}((A, <))$ is an elementary formula, the first clause simplifies to $\mathsf{wo}(A, <)$. Written in full, the second clause amounts to

$$\big(\forall a \in A\,(A_{<a} \subset \{x \mid \varphi(x)\} \to \varphi(a)) \to \forall a \in A\,\varphi(a) \wedge \exists a\,\neg\varphi(a)\big)^M, \qquad (3.26)$$

thus we can apply elementary comprehension to form the class $X = \{x \mid \varphi(x)^M\}$

and rewrite (3.26) as

$$\forall a \in A \left(\forall x \in A(x < a \to x \in X) \to a \in X.\right) \land \exists a \in A\, a \notin X,$$

which contradicts $\mathsf{wo}((A, <))$. □

Remark 265. In arithmetic, for every finite collection S of closed $\mathcal{L}_A^2$ formulas that hold in some ω-model, there is an ω-model of S that does not satisfy full bar induction (cf. Corollary VIII.5.8 in [Sim98]). This is in sharp contrast to the situation in set theory as displayed by Theorem 264.

Theorem 266. *It is provable in the theory* RFN *that elementary transfinite recursion along* ω *entails the consistency of the theory* Π^1_∞-BI.

Proof. We know from Theorem 264 that in every V-model every instance of Π^1_∞-BI holds, thus we obtain the claim as a result of Lemma 261. □

Comparing the situation to the arithmetic case, it is clear from the previous result that in the realm of sets and classes the principle of bar induction falls significantly short in terms of its power relative to V-model reflection. Since the notion of weak well orders cannot be captured by an elementary formula, and since this proved to be the crucial point in the weakness of bar induction, we will give a formulation of the bar induction schema that is based on our principle of weak induction for weak well orders (cf. Lemma 167 and Definition 168). However, while it is clear that our new schema with weak well orders entails bar induction as presented before, we are unable to provide a convincing case of application of our principle, such as for example proving the existence of V-models, let alone V-model reflection for any richer collections of formulas.

The schema of bar induction for weak well orders, BI^+ is a straightforward generalization of the principle of weak induction[15],

$$\mathsf{wwo}(A, <_A) \to \forall X \left(\mathsf{prog}^+(X, (A, <_A)) \to A \subset X\right)$$

[15] Remember that $\mathsf{prog}^+(X, (A, <_A))$ means that X is a superprogressive subclass of A (cf. Definition 165).

to the schema

$$\mathsf{wwo}(A, <_A) \rightarrow \forall \mathcal{X} \left(\mathsf{prog}^+(\mathcal{X}, (A, <_A)) \rightarrow A \subset \mathcal{X}\right),$$

where, just as in the usual case of bar induction, the meta variable $\mathcal{X}$ ranges over collections of the form $\{x \mid \varphi(x)\}$ for suitable formulas $\varphi(x)$. To formalize this principle we proceed as in the case of usual bar induction.

Definition 267. For every formula φ, we introduce the formulas

$$\begin{aligned}\mathsf{prog}^+(\varphi, (A, <_A)) &\equiv \forall m \in \mathcal{P}(A)\backslash\{\varnothing\} \left(\forall b\, (b <_A m \rightarrow \varphi(b)) \rightarrow \exists a \in m\, \varphi(a)\right)\\ \mathsf{Ti}^-(\varphi, (A, <_A)) &\equiv \mathsf{prog}^+(\varphi(x), (A, <_A)) \rightarrow \forall a \in A\, \varphi(a)\end{aligned}$$

For any collection $\mathcal{F}$ of formulas, the theory $\mathcal{F}$-BI^+ consists of NBG together with all instances of the schema

$$\mathsf{wwo}(A, <_A) \rightarrow \mathsf{Ti}^-(\varphi, (A, <_A))$$

where φ is in $\mathcal{F}$.

While the schema bar induction in the form BI^+ applies to a wider class of linear orders than the usual formulation BI (i.e. weak well orders vs. well orders), it also draws a weaker (i.e. $\mathsf{Ti}(\dots)$ vs. $\mathsf{Ti}^-(\dots)$) conclusion from its premise. Thus, it is not clear at prima facie which of the two systems is more powerful. However, at second glance, it is clear that since for any given well order $(A, <_A)$ and any formula φ the formulas $\mathsf{prog}(\varphi, (A, <_A))$ and $\mathsf{prog}^+(\varphi, (A, <_A))$ are equivalent, it follows that $\mathsf{BI} \subset \mathsf{BI}^+$.

Lemma 268. *Let $(A, <_A)$ be a well order and let φ be a formula. It is provable in* NBG *that*

$$\mathsf{prog}^+(\varphi, (A, <_A)) \leftrightarrow \mathsf{prog}(\varphi, (A, <_A))$$

holds.

Proof. The proof is almost identical to the proof of Lemma 122: we assume $\mathsf{prog}(\varphi, (A, <_A))$ for some formula φ and a well order $(A, <_A)$. For every nonempty set $m \subset A$ fix $a_m = \min_{<_A}(m)$. To verify $\mathsf{prog}^+(\varphi, (A, <_A))$, we have to prove that for every nonempty set $m \subset A$ that satisfies $A_{<_A m} \subset \{x \mid \varphi(x)\}$, there is an element $a \in m$ with $\varphi(a)$. This is clearly the case, since for every such set m we can apply the assumption $\mathsf{prog}(\varphi, (A, <_A))$ to obtain $\varphi(a_m)$. $\square$

Corollary 269. *For every collection $\mathcal{F}$ of formulas it is provable in* NBG *that* $\mathcal{F}\text{-}\mathsf{BI} \subset \mathcal{F}\text{-}\mathsf{BI}^+$.

Summary

We saw that the arithmetical and the set theoretical setting are dissimilar when comparing the relationship between reflection principles (in the sense of Definition 21 and Definition 260) and bar induction (cf. Definition 263 and Definition 19). In the arithmetic case, the theories $\Pi^1_1\text{-}\mathsf{RFN}_0$ and $\Pi^1_1\text{-}\mathsf{BI}_0$ prove the same formulas (cf. Theorem 22) in contrast to our setting, where the theory RFN together with elementary transfinite recursion along ω is enough to prove the consistency of full bar induction. Further, we gave a straightforward formalization of weak induction for formulas, resulting in the theory $\mathcal{F}\text{-}\mathsf{BI}^+$ (cf. Definition 267). We saw in Corallary 269 that $\mathcal{F}\text{-}\mathsf{BI}^+$ entails bar induction for all formulas in $\mathcal{F}$. Further questions about the capabilities of the theory $\mathcal{F}\text{-}\mathsf{BI}^+$ remain open.

Bibliography

[Avi96] J. Avigad. On the relationships between ATR_0 and $\widehat{ID}_\omega$. *Journal of symbollic logic*, 61(3), 1996.

[Bar75] J. Barwise. *Admissible Sets and Structures: An Approach to Definability Theory.* Perspectives in Mathematical Logic Series. Springer-Verlag, 1975.

[Fel71] U. Felgner. Comparison of the axiom of local and universal choice. *Fundamenta Mathematica*, LXXI, 1971.

[Fri74] H. Friedman. Some systems of second order arithmetic and their use. *JProceedings of the International Congress of Mathematicians Vancouver*, 41:235–242, 1974.

[Fri76] H. Friedman. Systems of second order arithmetic with restricted induction i,ii. *Journal of symbollic logic*, 41:557–559, 1976.

[FS13] D. Flumini and K. Sato. Hierarchies and well foundedness. Submitted, 2013.

[Glo76] K. Gloede. Reflection principles and indescribability. *In [Mül76]*, pages 277–323, 1976.

[Hal11] L. J. Halbeisen. *Combinatorial set theory.* Springer, 2011.

[Jec03] T. Jech. *Set Theory, The Third Millennium Edition.* Springer, 2003.

[JK10] G. Jäger and J. Krähenbühl. Σ^1_1 choice in a theory of sets and classes. In Ralf Schindler, editor, *Ways of Proof Theory*, pages 283 – 314. Ontos Verlag, Ontos Verlag, 2010.

[JS99] G. Jäger and T. Strahm. Bar induction and ω model reflection. *Annals of Pure and Applied Logic*, 97:221–230, 1999.

[Kan09] A. Kanamori. *The higher infinite.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, second edition, 2009.

[Kle55] S. C. Kleene. Hierarchies of number-theoretic predicates. *Bulletin of the American Mathematical Society*, 61, 1955.

[Kun80] K. Kunen. *Set Theory, An Introduction to Independence Proofs.* Studies in Logic and the Foundations of Mathematics; vol. 102. Elsevier, 1980.

[LL13] Logic and Theory Group (LTG). Subsystems of second order arithmetic and set theory, accessed June 03, 2013.

[Men97] E. Mendelson. *Introduction to mathematical logic.* Chapman&Hall/CRC, 4 edition, 1997.

[Mül76] G. H. Müller, editor. *Sets and Classes. On The Work by Paul Bernays.* North-Holland Publishing Company, 1976.

[Oik92] J. Oikkonen. A recursion principle for linear orderings. *The journal of symbolic logic*, 57(1), March 1992.

[Pro05] D. Probst. *Pseudo-Hierarchies in Admissible Set Theory without Foundation and Explicit Mathematics.* Phd, Universität Bern, Institut für Informatik und angewandte Mathematik, 2005.

[Sal05] V. Salipante. *On the Consistency Strength of the Strict Π^1_1 Reflection Principle.* Phd, Universität Bern, Institut für Informatik und angewandte Mathematik, 2005.

[Sat12] K. Sato. Relative predicativity and dependent recursion in second-order set theory and higher-order theories. Submitted, 2012.

[Sat13] K. Sato. Full and hat inductive definitions are equivalent in nbg. Submitted, 2013.

[SF10] R.M. Smullyan and M. Fitting. *Set Theory and the Continuum Problem.* Dover Books on Mathematics Series. Dover Publications, 2010.

[Sim98] S. G. Simpson. *Subsystems of Second Order Arithmetic*. Perspectives in Logic. Cambridge University Press, 11. 1998.