

Full and Hat Inductive Definitions Are Equivalent in NBG

SATO Kentaro

the date of receipt and acceptance should be inserted later

Abstract A new research project has, quite recently, been launched to clarify how different, from systems in second order number theory extending **ACA**₀, those in second order set theory extending **NBG** (as well as those in $n+3$ -th order number theory extending the so-called Bernays Gödel expansion of full $n+2$ -order number theory etc.) are. In this article, we establish the equivalence between Δ_0^1 -**LFP** and Δ_0^1 -**FP**, which assert the existence of a least and of a (not necessarily least) fixed point, respectively, for positive elementary operators (or between Δ_0^{n+2} -**LFP** and Δ_0^{n+2} -**FP**). Our proof also shows the equivalence between **ID**₁ and $\widehat{\text{ID}}_1$, both of which are defined in the standard way but with the starting theory **PA** replaced by **ZFC** (or full $n+2$ -th order number theory with global well-ordering).

Keywords subsystems of Morse-Kelley set theory · von Neumann-Bernays-Gödel set theory · higher order number theory · well-foundedness · proof theoretic strength

Mathematics Subject Classification (2000) (Primary) 03F35 · (Secondary) 03B15 · 03D65 · 03E70 · 03F25

1 Introduction

General Programme. A new research project has been launched: Among the known results in second order number theory, which hold in general second order theories (including second order set theory and higher order number and set theories, which can be seen as the second order extensions of one-lower order theories) and which does not (i.e., specific to second order number theory). While the author's previous work [17] in this project was motivated by *relative predicativity*, i.e., predicativity given infinite entities (e.g., the sets of reals or of functions, the universe of sets) other than traditional ω , the project is also motivated by at least three other trends of research.

- (1) The present author is working on a long-term project, *reverse set theory*, which investigates impacts of infinity axiom and analogues (e.g., large cardinal axioms) over the structure of proof-theoretic strengths of set-theoretic axioms (see [15]). In order to see an impact of the existence of an inaccessible cardinal, it is convenient to investigate, via interpretations into second order set theory, the relations among several extensions of *von Neumann-Bernays-Gödel set theory* **NBG** (as, to see that of infinity or ω , the investigation on interpretations into second order number theory were used in [14]). For the existence of a number of regular or inaccessible cardinals, higher order number or set theory plays the role.
- (2) Another is from *axiomatic truth theory*, a field in which the notion of truth is investigated with techniques from mathematical logic. While the truth axiomatized over number theory has been of central interest, some truth-theorists now consider the notion over set theory. As interpretations into

This research was partially supported by Hasler Foundation.

K. Sato
Institut für Informatik und angewandte Mathematik, Universität Bern, Neubrückstrasse 10, Bern 3012, Switzerland
E-mail: sato@iam.unibe.ch

or from second order number theory have played central roles, those into or from second order set theory play the same roles in the new theory, as in [5]. The notions of truth over the theories of reals and of functions also seem worthy to investigate.

- (3) The last but not least is from proof theory for *operational set theory*, a framework introduced in order to extract the core from different styles of mathematics. Jäger [6] identified the proof-theoretic (or, consistency) strengths of some extensions of the basic operational set theory, using second order set theory as measure, and he and Krähenbühl [7] enhanced it by obtaining a result whose analogue was known in second order number theory.

Although second order set theory will be treated as if it is the dominant example in the following expositions (and even in the title), and although Trends (2) and (3) above motivates only the investigations on second order set theory¹, it must be emphasized that the extension of our result to higher order theories is not the icing on the cake but is significant on its own right. For, our investigation of higher order number theories is to higher order recursion theory what second order number theory is to ordinary recursion theory (as pointed out in [17, §7]).

Results from Preceding Researches. On the one hand, the preceding researches show that many results known in second order number theory survive also in second order set theory: By *partial cut elimination* and *partial truth predicate* we can show that Π_{n+1}^1 comprehension axiom $\Pi_{n+1}^1\text{-CA}$ implies the consistency of $\Pi_n^1\text{-CA}$ with a suitable base theory; Krähenbühl [9] shows that several applications of proof-theoretic techniques survive, among which is the conservation in the presence of full foundation, between Σ_1^1 -class-collection $\Sigma_1^1\text{-Coll}$ and $(\Delta_0^1\text{-CA})_{\triangleleft E_0}$, i.e., Δ_0^1 transfinite recursion up to E_0 , the set-theoretic analogue of ε_0 (see [7]).

On the other hand, the previous work [17] of the present author shows that the straightforward analogues do not always hold: The consistency of $\Delta_0^1\text{-TR}$, Δ_0^1 transfinite recursion for all well-founded relations, is implied by $\Delta_0^1\text{-FP}$ (which should be called Δ_0^1 hat inductive definition) which asserts the existence of a fixed point of any positive Δ_0^1 operator, and also by Π_1^1 reduction principle $\Pi_1^1\text{-Red}$ in second order set theory (and $\text{Con}(\Delta_0^{n+2}\text{-TR})$ is by $\Delta_0^{n+2}\text{-FP}$ and by $\Pi_1^{n+2}\text{-Red}$ in higher order theories) whereas in second order number theory all these three are equivalent (see [18, Chapter V]). Moreover, Fujimoto [5] pointed out that the schemata that we have to consider in second order set theory include set-separation and set-collection (see Example 10 and Remark 12), not only foundation, whereas, in second order number theory, the analogues of set-separation and of set-collection are implied by induction.

Additionally, the previous work [17] of the author discovers a new kind of axiom schema, *dependent transfinite recursion* $\Delta_0^1\text{-TR}^n$ or dependent iteration of elementary comprehension. It is shown there that $\Delta_0^1\text{-TR}^n$ for $n \geq 2$ is strictly stronger than the usual transfinite recursion in second order set theory (and in the higher order theories) whereas they are equivalent in second order number theory. As mentioned in the work, this new kind of iteration can apply not only to elementary comprehension but also to other constructions, e.g., fixed-point recursion. It is also worth mentioning that this avoids the general limit of extensions of transfinite recursion scheme, which Flumini and the present author [4] gave in a quite general second order framework.

¹ Trend (2), nonetheless, could motivate our programme in higher order number theory, especially in the third order for the same reason as relative predicativity. Namely, the truth axiomatized over full second order number theory should be an interesting subject. Feferman, one of the founders of axiomatic truth theory, states:

one might argue for an intermediate position between that of conceptual structuralism, which rejects the continuum as a definite totality, and the set-theoretical account which not only accepts that but also much, much more. Namely, one may grant as a working apparently robust idea the concept of $S(N)$ [the power set of the set ω of natural numbers], but nothing higher in the cumulative hierarchy. This would justify the assumption of Dedekind or Cantor completeness of the real line with respect to all sets definable by quantification over the continuum, thus going far beyond predicative mathematics into the domain of descriptive set theory. In logical terms, that would justify working in a system of strength full 2nd order number theory, or “analysis” as it is justly called. [3, p.22]

This view of his motivates well also the research on the “predicativity, given the totality of $\mathcal{P}(\omega)$ (i.e., real numbers)”, one of the main themes of the present author’s previous work [17].

Moreover, even Trend (3) motivates our programme in third order number theory: Feferman, who is also the founder of operational set theory, questioned in [2, Section 4] the strength of $\text{OST}+(\text{Uni})$, (or $\text{OST}(\mathbf{E})$ in the terminology of [6]). At this point it seems the most plausible that this system and variants (e.g., $\text{OST}'(\mathbf{E})$ defined with restricted foundation) can be characterized by extensions of Bernays Gödel expansions of full second order number theory.

Our Results. In this article, we show that the following hold both in second order number theory where $\mathbf{BT}^2 = \mathbf{ACA}_0$, and in second order set theory where $\mathbf{BT}^2 = \mathbf{NBG}$ (as well as in the aforementioned higher order theories where $\mathbf{BT}^2 = \Delta_0^{n+2}\text{-CA}_0$ from [17, Definition 5]), although in the former these results are less significant or weaker than previously known results (e.g., it is well known that $\Pi_1^0\text{-LFP}$ is equivalent to $\Delta_0^1\text{-LFP}$, and that $\Pi_1^1(\text{WF})$ is essentially Π_2^1 , etc.).

1. $\Delta_0^1\text{-LFP}$, the existence of a least fixed point of a positive Δ_0^1 operator, and a strengthened variant $\Delta_0^1\text{-SLFP}$ are both equivalent, over $\mathbf{BT}^2 + \Pi_1^0\text{-LFP}$, to $\Delta_0^1\text{-FP}$ (Corollary 9 (i)).
2. $\Pi_1^1(\text{WF})\text{-Red}$, reduction scheme for those formulae that are Π_1^1 in well-foundedness WF, is both logically and proof-theoretically strictly stronger than $\Delta_0^1\text{-LFP}$ (Corollaries 8 (2) and 12 (1)), and, moreover, than $\Delta_0^1\text{-LFTR}$, the transfinite iteration of least fixed-point construction, (Corollary 13); and even than $\Delta_0^1\text{-LFTR}^n$, the dependent transfinite iteration (see Definitions 16 and 17).
3. $\mathbf{BT}^2 + \Delta_0^1\text{-FP}$ cannot prove $\Delta_1^1(\text{WF})\text{-CA}$ nor $\Sigma_1^1(\text{WF})\text{-Coll}$, a second order collection for Σ_1^1 in WF (Corollary 12 (2)), while $\mathbf{BT}^2 + \Delta_0^1\text{-FP}$, $\mathbf{BT}^2 + \Delta_0^1\text{-FP} + \Delta_1^1\text{-CA}$ and $\mathbf{BT}^2 + \Delta_0^1\text{-FP} + \Sigma_1^1\text{-Coll}$ are all Π_2^1 -equivalent (Remark 9).
4. $\Delta_0^1\text{-Ref}$, the existence of a first order model reflecting a Δ_0^1 -formula, or even weaker $\Pi_2^0\text{-Ref}$, implies both $\Pi_1^0\text{-LFP}$ (Proposition 4) and the Δ_0^1 -ness of WF (Lemma 3).

A contrast appears in $\Delta_0^1\text{-Ref}$, which does not hold in second order number theory (because WF is Π_1^1 -complete) and which does in second order set theory (as well as in the higher order theories).

First order analogues. As in number theory, the research should involve not only second order systems, but also “first order analogues”, e.g., ID_α and \widehat{ID}_α , since they provide finer analysis. We can straightforwardly define the set-theoretic (or real- and function-theoretic) analogues of them by starting with **ZFC** (or full higher order number theory with global well-ordering) instead of **PA**, and more generally, $\mathcal{F}\text{-}ID_\alpha[T]$ and $\mathcal{F}\text{-}\widehat{ID}_\alpha[T]$ starting with theory T , for classes \mathcal{F} of formulae. In Section 8 by examining proofs of our main results, we will see:

5. Reflection principle (in T) allows us to interpret $\Pi_1^0\text{-ID}_\alpha[T]$ in T ;
6. ID_1 is interpretable in $\Pi_1^0\text{-ID}_1[\widehat{ID}_1]$ where $\mathcal{F} = \mathcal{L}^1$, the starting language, is omitted;
7. ID_ω and $\mathbf{BT}^2 + \Delta_0^1\text{-LFP}$ plus the axiom schemata extended to \mathcal{L}^2 are \mathcal{L}^1 -equivalent.

In particular, over **ZFC** (as well as full higher order theories) where reflection principle holds, ID_α and \widehat{ID}_α are mutually interpretable. This contrasts the well known result over **PA** that ID_1 is proof-theoretically strictly stronger than \widehat{ID}_ω (and more general \widehat{ID}_α for reasonable α).

Comparison between second order number and set theories. Combining with the results from the author’s previous work [17], we can summarize the contrasts between second order number theory and second order set theory (as well as higher order number and set theories with global well ordering), as in Figure 1, where one-head arrows denote logical implications and where two-head ones denote both logical and consistency-wise strict implications.

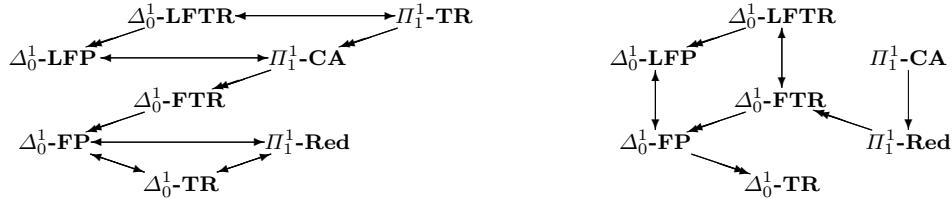


Fig. 1 Comparison between extensions of \mathbf{ACA}_0 (left) and of \mathbf{NBG} (right)

We will see another contrast between second order number and set theories, the way in which Nemoto hierarchy of least fixed point principles collapses: Nemoto [13, §2.5] defined syntactical analogues of Wadge classes, by imitating the description of Wadge hierarchy given in [11, §1], and considered a hierarchy of determinacy statements for such classes. We can consider *Nemoto hierarchy* also for least fixed point principles, the segment below Δ_0^1 of which collapses into two equivalent classes, one consisting of those provable in the base theory and the other of those implying full Δ_0^1 least fixed point principle. The border between the two is shown in Figure 2.

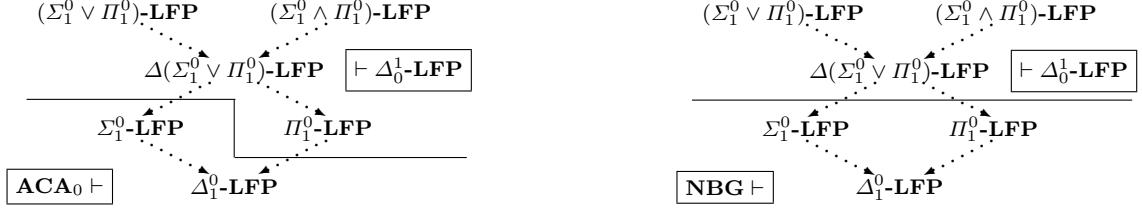


Fig. 2 Nemoto hierarchy of least fixed point principles in second order number/set theory

Outline. Despite the importance of the other higher order number and set theories mentioned above, for readability we at first concentrate on second order set theory and on the comparison with the already well investigated second order number theory, and in Section 9 we will make several comments on the generalizations of our results to the other higher order cases.

2 Preliminaries

As mentioned before, despite the importance of the other higher order number and set theories, at first we concentrate on second order set theory or on the difference between the already well investigated second order number theory and the new second order set theory. Therefore we present here several notations and definitions for second order set theory (while those for second order number theory are well known), and we will generalize them in Section 9.

Definition 1 (The language \mathcal{L}_S^2) The language \mathcal{L}_S^1 is the one-sorted first-order language that has no function symbols and only one predicate symbol \in of arity 2, besides equality $=$.

The language \mathcal{L}_S^2 is the two-sorted first-order language with equality $=$ only for the first sort, with no function symbols and with one predicate symbol \in , where the arguments of \in are either two sets (objects of the first sort) or a pair of a set and a class (an object of the second).

\mathcal{L}_N^2 denotes the well known language of second order number theory with exponentiation \exp .

The first and second sorts are referred to as *first order* and *second order*, respectively. By convention, lower- and upper- cases letters denote first and second, respectively, order objects.

Note that the equality $=$ for the second order is *not* included in \mathcal{L}_S^2 as a primitive symbol. Nevertheless we let $X = Y$ abbreviate $(\forall x)(x \in X \leftrightarrow y \in Y)$, and $X \subset Y$ similarly.

Definition 2 (Δ_0^i and Π_1^i) A first order quantifier is called *bounded* if it occurs in the way $\exists x(x \in a \wedge \varphi(x, a))$ or $\forall x(x \in a \rightarrow \varphi(x, a))$, denoted by $(\exists x \in a)\varphi(x, a)$ or $(\forall x \in a)\varphi(x, a)$ respectively.

An \mathcal{L}_S^2 -formula φ is Δ_0^0 (or called *bounded*) if it contains no unbounded first order nor second order quantifiers. An \mathcal{L}_S^2 -formula is Π_1^0 if it is of the form $(\forall x)\varphi(x)$ for some Δ_0^0 -formula $\varphi(x)$.

An \mathcal{L}_S^2 -formula φ is Δ_0^1 (or called *elementary*) if it contains no second order quantifiers. An \mathcal{L}_S^2 -formula is Π_1^1 if it is of the form $(\forall X)\varphi(X)$ for some Δ_0^1 -formula $\varphi(X)$.

More generally, we can define Σ_n^0 , Π_n^0 , Σ_n^1 and Π_n^1 by numbers of first or second order quantifiers as usual. These are straightforward analogues of those in second order number theory.

For a formula $\varphi(x)$, and $\psi(\{x \mid \varphi(x)\})$ denotes the result of replacing all the subformulae $t \in X$ of $\psi(X)$ by $\varphi(t)$ simultaneously. $\psi(\{x \mid \varphi(x)\})$ is sometimes denoted also by $\psi(\varphi)$ when it is clear from the context by which variable we abstract the substituted formula φ . Sometimes second order X is regarded as the abstract $\{z \mid z \in X\}$.

Definition 3 ($\Delta_0^1(\mathcal{F})$, $\Sigma_n^1(\mathcal{F})$ and $\Pi_n^1(\mathcal{F})$) For a class \mathcal{F} of formulae, $\Delta_0^1(\mathcal{F})$ ($\Sigma_n^1(\mathcal{F})$ and $\Pi_n^1(\mathcal{F})$) is the class consisting of all formulae of the form $\varphi(\{x \mid \psi_1(x, \theta_0, \dots, \theta_m)\}, \dots, \{x \mid \psi_k(x, \theta_0, \dots, \theta_m)\})$ for some formulae ψ_1, \dots, ψ_k from \mathcal{F} , some formulae $\theta_0, \dots, \theta_m$ from Δ_0^0 and $\varphi(X_1, \dots, X_k)$ from Δ_0^1 (Σ_n^1 and Π_n^1 , respectively).

$\{x, y\}$ denotes the unique set z , whose existence is guaranteed by the axiom below, such that $(\forall w)(w \in z \leftrightarrow w = x \vee w = y)$; and $\langle x, y \rangle$ denotes $\{\{x, x\}, \{x, y\}\}$. $(X)_x$ denotes $\{y \mid \langle x, y \rangle \in X\}$. “ W well orders X ” stands for $\text{WF}[W] \wedge (\forall x, y \in X)(\langle x, y \rangle \in W \vee x = y \vee \langle y, x \rangle \in W)$, where

Definition 4 $\text{WF}[W] \equiv (\forall X)((\forall x)((\forall y)(\langle x, y \rangle \in W \rightarrow y \in X) \rightarrow x \in X) \rightarrow (\forall x)(x \in X))$.

Now we introduce *von Neumann-Bernays-Gödel class theory* **NBG**, which is our base theory.

Definition 5 **NBGC** is the \mathcal{L}_S^2 theory generated (over the classical two-sorted first order logic with equality for the first sort) by (the universal closures of) the following axioms:

Extensionality $(\forall z)(z \in x \leftrightarrow z \in y) \rightarrow x = y$;

Emptyset, Pair, Union, Powerset, Infinity (as usual in **ZFC**);

Sep $(\exists y)(\forall z)(z \in y \leftrightarrow (z \in x \wedge z \in X))$;

Repl $(\forall y \in x)(\exists!z)(\langle y, z \rangle \in X) \rightarrow (\exists w)(\forall y \in x)(\exists z \in w)(\langle y, z \rangle \in X)$;

Δ_0^1 -CA $(\exists X)(\forall y)(y \in X \leftrightarrow \varphi(y, \mathbf{x}, \mathbf{Y}))$ for any Δ_0^1 -formula $\varphi(x)$ in which X is not free;

Found $(\forall x)((\forall y \in x)(y \in X) \rightarrow x \in X) \rightarrow (\forall x)(x \in X)$;

Choice $(\exists W)(W \text{ well orders } V = \{x \mid x = x\})$.

NBG includes **Found** but not **Choice** (i.e., **NBGC** minus **Choice**); and the superscript “ $-$ ” means the removal of **Found**, like **NBG** $^-$ and **NBGC** $^-$.

In literature von Neumann - Bernays - Gödel class theory contains **Found**, while there seems to be no agreement on whether it contains **Choice**. For our purpose in what follows, such differences are not so important and we consider all the four variants.

Fact 1 **NBG** $^-$, **NBG**, **NBGC** $^-$ and **NBGC** are conservative extension of **ZF** $^-$, **ZF**, **ZF** $^- + \text{AC}$ or **ZFC**, respectively, where **ZF** $^-$ denotes **ZF** minus foundation axiom.

This fact is the analogue of the conservation between **ACA** $_0$ and **PA**, according to the correspondence between the two sorts of \mathcal{L}_N^2 and those of \mathcal{L}_S^2 .

We are interested in similarities and dissimilarities to second order number theory, and the analogy to second order number theory will help the readers to understand our results. Therefore we will treat the both uniformly as far as possible. We denote, by **BT** 2 , both **ACA** $_0$ and **NBG** (as well as the aforementioned variants) when we describe results for both, and by **BT** 1 , both **PA** and **ZF** (as well as the corresponding variants). \mathcal{L}^2 denotes the language of **BT** 2 (i.e., \mathcal{L}_N^2 or \mathcal{L}_S^2 accordingly), and \mathcal{L}^1 denotes that of **BT** 1 . In the following discussion, **BT** 2 can also be **BGZ** \equiv **NBG** – **Repl**, Bernays Gödel expansion of Zermelo set theory **Z** (which will be used in the proof of Proposition 21). In later section 9 we will consider a generalization of **BT** 2 .

Next we prepare additional axiom schemata, which are to be added to our base theory **BT** 2 . Note that, by virtue of the abbreviations, these are literally the same as in second order number theory (and easily generalized to other second order frameworks with the notion of pair).

Definition 6 For a class \mathcal{F} of formulae, define the following schemata.

\mathcal{F} -Coll $(\forall x)(\exists Y)\varphi(x, Y) \rightarrow (\exists Z)(\forall x)(\exists y)\varphi(x, (Z)_y)$;

\mathcal{F} -TI $\text{WF}[R] \rightarrow \text{TI}[\varphi](R)$ where $\text{TI}[\varphi](R) \equiv (\forall x)((\forall y \in (R)_x)\varphi(y) \rightarrow \varphi(x)) \rightarrow (\forall x)\varphi(x)$;

\mathcal{F} -CA $(\exists X)(\forall x)(\varphi(x) \leftrightarrow x \in X)$;

\mathcal{F} -TR $\text{WF}[W] \rightarrow (\exists H)(\forall u)((H)_u = \{x \mid \varphi(u, x, \{\langle v, y \rangle \mid v \in (W)_u \wedge \langle v, y \rangle \in H\})\})$;

$\Delta(\mathcal{F})$ -CA $(\forall x)(\varphi(x) \leftrightarrow \neg\psi(x)) \rightarrow (\exists X)(\forall x)(\varphi(x) \leftrightarrow x \in X)$;

\mathcal{F} -Red $(\forall x)(\varphi(x) \vee \psi(x)) \rightarrow (\exists Z)(\forall x)((x \in Z \rightarrow \varphi(x)) \wedge (x \notin Z \rightarrow \psi(x)))$;

\mathcal{F} -FP $(\exists F)(\forall x)(\varphi(x, F) \leftrightarrow x \in F)$;

\mathcal{F} -LFP $(\exists F)((\forall x)(\varphi(x, F) \rightarrow x \in F) \wedge (\forall Y)((\forall x)(\varphi(x, Y) \rightarrow x \in Y) \rightarrow F \subset Y))$,

for \mathcal{F} -formulae φ, ψ in which neither Z nor H occurs freely and in which F occurs only positively.

Remark 1 Although we officially defined only $\Delta(\mathcal{F})$ -CA, we can define $\Delta(\mathcal{F})$ -Coll, $\Delta(\mathcal{F})$ -TR, $\Delta(\mathcal{F})$ -Red, $\Delta(\mathcal{F})$ -FP and $\Delta(\mathcal{F})$ -LFP etc., similarly. Thus we treat $\Delta(\mathcal{F})$ as if it were a class of formulae, following the standard convention in which Δ_n^1 abbreviates $\Delta(\Pi_n^1)$.

Here we introduce \mathcal{F} -reduction schema, instead of $(\neg\mathcal{F})$ -separation schema defined as in [18], which is equivalent over classical logic. For, the term *separation* has the different meaning in the context of set theory (as in Definition 5). This is only a matter of name.

Lemma 1 **BT** $^2 + \Pi_n^1$ -CA + Π_n^1 -Coll proves Σ_{n+1}^1 -Red and hence Δ_{n+1}^1 -CA.

Proof Let φ and ψ be Π_n^1 , and assume $(\forall x)((\exists Y)\varphi(x, Y) \vee (\exists Y)\psi(x, Y))$. Then $(\forall x)(\exists Y)(\varphi(x, Y) \vee \psi(x, Y))$, and Π_n^1 -Coll yields Z with $(\forall x)(\exists y)(\varphi(x, (Z)_y) \vee \psi(x, (Z)_y))$. Thus, $X = \{x \mid (\exists y)\varphi(x, (Z)_y)\}$ satisfies $(\forall x)(x \in X \rightarrow (\exists Y)\varphi(x, Y))$ and $(\forall x)(x \notin X \rightarrow (\exists Y)\psi(x, Y))$.

To describe some of the schemata above, it is convenient to introduce the following notion.

Definition 7 (operator form) An operator from $\lambda y.\lambda \mathbf{X}, \mathbf{Y}. \{z \mid \varphi(z, y, \mathbf{X}, \mathbf{Y})\}$ of arity (j, k, l) is an \mathcal{L}^2 -formula $\varphi(z, y, \mathbf{X}, \mathbf{Y})$ (which may have parameters other than $z, y, \mathbf{X}, \mathbf{Y}$) in which \mathbf{X} occur only positively, with two kinds of abstractions, one applied to variables \mathbf{X}, \mathbf{Y} and y , and the other to z , where $|y| = j, |\mathbf{X}| = k, |\mathbf{Y}| = l$. For a class \mathcal{F} of formulae, $\lambda y.\lambda \mathbf{X}, \mathbf{Y}. \{z \mid \varphi(z, y, \mathbf{X}, \mathbf{Y})\}$ is called an \mathcal{F} operator if $\varphi(z, y, \mathbf{X}, \mathbf{Y})$ is in \mathcal{F} . We follow the convention for λ -notation, e.g., identification up to α -equivalence and β -equivalence. $\Delta_n^i(\Gamma), \Sigma_n^i(\Gamma), \Pi_n^i(\Gamma)$ denotes $\Delta_n^i(\{\varphi\}), \Sigma_n^i(\{\varphi\}), \Pi_n^i(\{\varphi\})$ respectively,

Using this, the following reformulation is possible, where $(H)^{Wx} = \{\langle y, u \rangle \in H \mid y \in (W)_x\}$:

\mathcal{F} -TR $\text{WF}(W) \rightarrow (\exists H)\text{Hier}[\Gamma](H, W)$ for any \mathcal{F} operator Γ of arity $(1, 0, 1)$ without free H ,

where $\text{Hier}[\Gamma](H, W) \equiv (\forall x)((H)_x = \Gamma(x, (H)^{Wx}))$;

\mathcal{F} -FP $(\exists F)(F = \Gamma(F))$ for an F -free \mathcal{F} operator Γ of arity $(0, 1, 0)$;

\mathcal{F} -LFP $(\exists F)(F = \Gamma(F) \wedge \forall Z(Z = \Gamma(Z) \rightarrow F \subset Z))$ for an F -free \mathcal{F} operator Γ of arity $(0, 1, 0)$.

\mathcal{F} -TR can be seen as iteration, along any well founded relation, of \mathcal{F} comprehension axiom. Similarly fixed point and least fixed point axioms can also be iterated (cf. [19, §4]):

Definition 8 Let us define

\mathcal{F} -FTR $\text{WF}(W) \rightarrow (\exists H)\text{FHier}[\Gamma](H, W)$ for a H -free \mathcal{F} operator Γ of arity $(1, 1, 1)$,

where $\text{FHier}[\Gamma](H, W) \equiv (\forall x)((H)_x = \Gamma(x, (H)_x, (H)^{Wx}))$;

\mathcal{F} -LFTR $\text{WF}(W) \rightarrow (\exists H)\text{LFHier}[\Gamma](H, W)$ for a H -free \mathcal{F} operator Γ of arity $(1, 1, 1)$,

where $\text{LFHier}[\Gamma](H, W) \equiv \text{FHier}[\Gamma](H, W) \wedge (\forall x, Z)(Z = \Gamma(x, Z, (H)^{Wx}) \rightarrow (H)_x \subset Z)$.

For a specific order type α (e.g., ω, ε_0 in number theory, and $\Omega = \text{Ord}, E_0$ in set theory from [7, §3]), $\mathcal{F}\text{-TI}_\alpha, \mathcal{F}\text{-TR}_\alpha, \mathcal{F}\text{-FTR}_\alpha$ and $\mathcal{F}\text{-LFTR}_\alpha$ denote $\mathcal{F}\text{-TI}, \mathcal{F}\text{-TR}, \mathcal{F}\text{-FTR}$ and $\mathcal{F}\text{-LFTR}$, respectively, but “ $\text{WF}(W) \rightarrow$ ” removed and W substituted by the standard well-ordering for α .

Remark 2 Notice that the notion of order type does not always make sense in the formal theories, because we do not know if the principle called comparison of well-orderings is available (and even we do not have the well-foundedness of them in the theories).

Here we take an attitude similar to that in second order arithmetic: Even though we do not have comparability nor well-foundedness in the target theories we are talking about the order types or ordinals (like ε_0) on the meta-level (which can be formalized in e.g., **ZF**) and we refer, as “the standard well-ordering”, to specific binary relations representing them in the standard model of second order number theory. Similarly, here we are talking about the order types or ordinals (like E_0) on the meta-level (which can be formalized in e.g., **ZF** plus the existence of a strongly inaccessible cardinal), and we refer, as “the standard well-ordering”, to specific binary relations representing them in the ‘standard model’ of second order set theory in this context, that is, $(V_\kappa, V_{\kappa+1})$ for a strongly inaccessible cardinal κ .

This is inessential and just a matter of taste. Actually we can avoid this problem by considering α as a (class-size) binary relation in the following practice.

In any case, since we can easily generalize the method employed in [7, §3] to define canonically, from the code of α , a binary relation coding ω^α , we can freely use the notation ω^α (as indeed we will in Section 6) as well as $2 \cdot \alpha$, etc., without any danger of confusion.

3 Reduction of fixed point principles

Our central theme is the investigation of $\Delta_0^1\text{-FP}$, $\Delta_0^1\text{-LFP}$ and their variants. Here as a preparation for later sections, we show the finite axiomatizability of these schemata (over **BT**²).

For this, in Subsection 3.1 we point out that some fragments of the schemata, namely $\Delta(\Sigma_1^0 \vee \Pi_1^0)\text{-FP}$ and $\Delta(\Sigma_1^0 \vee \Pi_1^0)\text{-LFP}$, are strong enough to imply the whole schemata. In the following two subsections, we answer to the naturally arising question: Whether these reductions to $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ are optimal or not.

3.1 Reduction to $\Delta(\Sigma_1^0 \vee \Pi_1^0)$

Lemma 2 Over \mathbf{BT}^2 , (a) both $(\Sigma_1^0 \vee \Pi_1^0)$ -**FP** and $(\Sigma_1^0 \wedge \Pi_1^0)$ -**FP** imply Δ_0^1 -**FP**; (b) both $(\Sigma_1^0 \vee \Pi_1^0)$ -**LFP** and $(\Sigma_1^0 \wedge \Pi_1^0)$ -**LFP** imply Δ_0^1 -**LFP**.

The basic idea is as follows. Assume an operator Γ is decomposed as $\Gamma = \Gamma_1 \circ \Gamma_0$. Consider a “parallelizing” operator $\tilde{\Gamma} : X \mapsto (\{0\} \times \Gamma_0((X)_1)) \cup (\{1\} \times \Gamma_1((X)_0))$. Then a fixed point F of $\tilde{\Gamma}$ satisfies $(F)_0 = \Gamma_0((F)_1)$ and $(F)_1 = \Gamma_1((F)_0)$, and thus $(F)_1 = \Gamma_1(\Gamma_0((F)_1)) = \Gamma((F)_1)$, namely $(F)_1$ is a fixed point of the original operator Γ . Conversely if G is a fixed point of Γ then $(\{0\} \times \Gamma_0(G)) \cup (\{1\} \times G)$ is a fixed point of $\tilde{\Gamma}$. We can generalize this argument for n many factors, for a natural number n . E.g., if $\Gamma = \Gamma_2 \circ \Gamma_1 \circ \Gamma_0$ then the “parallelizing” operator is $X \mapsto (\{0\} \times \Gamma_0((X)_2)) \cup (\{1\} \times \Gamma_1((X)_0)) \cup (\{2\} \times \Gamma_2((X)_1))$. Now, any elementary operator $\lambda X. \{z \mid \forall x_n \exists x_{n-1} \dots \psi(x_n, \dots, z, X)\}$ can be decomposed as $\Gamma = \Gamma_\forall \circ \Gamma_\exists \circ \dots \circ \Gamma_\psi$, by a Δ_0^0 operator $\Gamma_\psi : X \mapsto \{\langle x_n, \dots, x_0, z \rangle \mid \psi(x_n, \dots, z, X)\}$, a Σ_1^0 operator $\Gamma_\exists : X \mapsto \{z \mid \exists x(\langle x, z \rangle \in X)\}$ and a Π_1^0 operator $\Gamma_\forall : X \mapsto \{z \mid \forall x(\langle x, z \rangle \in X)\}$. Thus what we have to check is, basically, the “parallelizing” operator according to this decomposition is of the required complexity.

Proof Let $\lambda X. \{z \mid \varphi(z, X)\}$ be a Δ_0^1 operator of arity $(0, 1, 0)$. We may assume that, for a Δ_0^0 formula ψ ,

$$\varphi(z, X) \equiv \forall x_{2n+2} \exists x_{2n+1} \dots \forall x_0 \psi(x_{2n+2}, x_{2n+1}, \dots, x_0, z, X).$$

Define $\tilde{\varphi}$ as follows, which is equivalently $\Sigma_1^0 \vee \Pi_1^0$ since Δ_0^0 is closed under the quantifiers of the form $(\exists y)(\langle x, y \rangle = z \wedge \dots)$:

$$\begin{aligned} \tilde{\varphi}(u, Z) \equiv & (\exists x_0, \dots, x_{2n+2}, z)(u = \langle 0, \langle x_0, \dots, \langle x_{2n+2}, z \dots \rangle \rangle \wedge \psi(x_{2n+2}, \dots, x_0, z, (Z)_{2n+3})) \\ & \vee \bigvee_{i \leq n+1} (\forall x_{2i})(\exists y)(u = \langle 2i+1, y \rangle \wedge \langle x_{2i}, y \rangle \in (Z)_{2i}) \\ & \vee \bigvee_{i \leq n} (\exists x_{2i+1})(\exists y)(u = \langle 2i+2, y \rangle \wedge \langle x_{2i+1}, y \rangle \in (Z)_{2i+1}). \end{aligned}$$

Since “ $\bigvee[(Qx_i)(\exists y)(u = \langle k, y \rangle \wedge \dots)]$ ” can be replaced by “ $\bigwedge[(Qx_i)(\forall y)(u = \langle k, y \rangle \rightarrow \dots)]$ ”, over the Δ_0^0 region $\{u \mid (\exists k \leq 2n+3)(\exists y)(u = \langle k, y \rangle)\}$, $\tilde{\varphi}(u, Z)$ is equivalently $\Sigma_1^0 \wedge \Pi_1^0$ as well.

Now we have Z with $(\forall u)(u \in Z \leftrightarrow \tilde{\varphi}(u, Z))$. By (meta-)induction on $k \leq 2n+3$:

$$\langle k, \langle x_k, \dots, \langle x_{2n+2}, z \dots \rangle \rangle \in Z \leftrightarrow (Q_{k-1}x_{k-1}) \dots (\forall x_0)\psi((x_{2n+2}, x_{2n+1}, \dots, x_0, z, (Z)_{2n+3})).$$

In particular, $\langle 2n+3, z \rangle \in Z$ iff $\varphi(z, (Z)_{2n+3})$. Thus $(Z)_{2n+3}$ is a fixed point of $\lambda X. \{z \mid \varphi(z, X)\}$.

The ‘least-ness’ of $(Z)_{2n+3}$ follows from that of Z as follows. If F is a fixed point of $\lambda X. \{z \mid \varphi(z, X)\}$, then \tilde{F} defined below is a fixed point of $\lambda X. \{u \mid \tilde{\varphi}(u, X)\}$:

$$\tilde{F} = (\{2n+3\} \times F) \cup \bigcup_{k < 2n+3} \{\langle k, \langle x_k, \dots, \langle x_{2n+2}, z \dots \rangle \rangle \mid (Q_{k-1}x_{k-1}) \dots (\forall x_0)\psi((x_{2n+2}, \dots, x_0, z, F))\}.$$

If Z is a least fixed point of $\lambda X. \{u \mid \tilde{\varphi}(u, X)\}$, then $Z \subset \tilde{F}$ and so $(Z)_{2n+3} \subset (\tilde{F})_{2n+3} = F$.

Remark 3 (1) The proof above actually shows slightly stronger statements: Over \mathbf{BT}^2 , (i) $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ -**FP** implies Δ_0^1 -**FP** and (ii) $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ -**LFP** implies Δ_0^1 -**LFP** (cf. Remark 1).

(2) We have the same results for Δ_0^1 -**FTR**, Δ_0^1 -**LFTTR**, Δ_0^1 -**FTR** $_\alpha$ and Δ_0^1 -**LFTTR** $_\alpha$.

(3) Although we did not define officially the first order systems, the proof actually shows that $(\Sigma_1^0 \vee \Pi_1^0)$ - \widehat{ID}_1 and $(\Sigma_1^0 \vee \Pi_1^0)$ - ID_1 are equivalent to the full systems \widehat{ID}_1 and ID_1 respectively.

This allows us to axiomatize, over \mathbf{BT}^2 , the schemata Δ_0^1 -**FP**, Δ_0^1 -**LFP** and variants with single sentences, provided that we have a universal formula.

Definition 9 Assume that there is a universal $\Upsilon_1^0(e, x, X, Y)$ among the Π_1^0 -formulae whose free variables are among x, X, Y and in which X occurs only positively. Define:

$$\Upsilon(e, z, y, X, Y) \equiv (\exists e_0, e_1)(e = (e_0, e_1) \wedge \Upsilon_1^0(e_0, \langle z, y \rangle, X, Y) \wedge \neg \Upsilon_1^0(e_1, \langle z, y \rangle, \{x \mid x \notin X\}, Y)),$$

$$\tilde{\Upsilon} \equiv \lambda Y. \lambda X. \{u \mid (\exists e, z, y)(u = \langle \langle e, y \rangle, z \rangle \wedge \Upsilon(e, z, y, (X)_{\langle e, y \rangle}, Y))\}.$$

Then obviously Υ is a universal $\Sigma_1^0 \wedge \Pi_1^0$ -formula positive in X , i.e., it is $\Sigma_1^0 \wedge \Pi_1^0$ and X occurs only positively and, for any $(\Sigma_1^0 \wedge \Pi_1^0)$ -formula $\varphi(x, y, X, Y)$ with at most x, y, X, Y free in which X occurs only positively, our base theory \mathbf{BT}^2 proves $(\forall x, y, X, Y)(\varphi(x, y, X, Y) \leftrightarrow \Upsilon(\Gamma^\neg \varphi, x, y, X, Y))$.

Corollary 1 If there is such universal formula Υ_1^0 , Δ_0^1 -**FP** and Δ_0^1 -**LFP** are axiomatizable by single Π_2^1 - and Π_3^1 -sentences, respectively, over \mathbf{BT}^2 . Similarly Δ_0^1 -**FTR** and Δ_0^1 -**LFTTR** are axiomatizable, over \mathbf{BT}^2 , by single sentences and so are Δ_0^1 -**FTR** $_\alpha$ and Δ_0^1 -**LFTTR** $_\alpha$.

3.2 Reduction of Π_1^0 inductive definition to reflection

Now it is natural to ask: Is the reduction of Δ_0^1 -**(L)FP** to the $\Delta(\Sigma_1^0 \wedge \Pi_1^0)$ -fragment optimal?

The answer to LFP, in number theory, is no: The smaller fragment Π_1^0 -**LFP** is equivalent to Δ_0^1 -**LFP**. More precisely, since the accessible part of a binary relation W is a least fixed point of the Π_1^0 operator $\lambda X.\{z \mid (W)_z \subset X\}$, Π_1^0 -**LFP** allows us to check elementarily the well-foundedness of relations in a given family; and since well-foundedness is Π_1^1 -complete in number theory, Π_1^0 -**LFP** implies, over **ACA**₀, Π_1^1 -**CA**, which implies Δ_0^1 -**LFP** (actually Π_1^1 -**LFP** with the help of Π_1^1 -**CA** $\vdash \Sigma_1^1$ -**Coll**).

How about in set theory? Is Π_1^0 -**LFP** equivalent to Δ_0^1 -**LFP** also over **NBG**? Since well-foundedness is no longer Π_1^1 -complete, the same argument does not work. Actually **NBG** proves Π_1^0 -**LFP**, while, as shown in [17, Proposition 26, Corollary 33], Δ_0^1 -**LFP** has stronger consistency than **NBG**.

In order to clarify the similarities and the dissimilarities, we extract the source of this difference between number theory and set theory. It is reflection principle. Though there are several variants depending on various notions of “model” and of “reflect”, our version is as follows.

Definition 10 (reflection principle) Let φ^a denote the result of replacing (Qx) in φ by $(Qx \in a)$, and $\text{Trans}(a) \equiv (\forall x \in a)(\forall y \in x)(y \in a)$. For a class \mathcal{F} of formulae, \mathcal{F} reflection principle is:

$$\mathcal{F}\text{-Ref } (\forall y)(\exists a)[y \in a \wedge \text{Trans}(a) \wedge (\varphi(y) \rightarrow \varphi^a(y)) \wedge (\forall X)(\exists x)(\forall u)(u \in x \leftrightarrow u \in a \wedge u \in X)]$$

for any \mathcal{F} formula $\varphi(\mathbf{x})$ whose first order free variables are all among \mathbf{x} .

Remark 4 For uniformity with number theory, we define $x \in y$ in number theory as “ x -th digit of binary expansion of y is 1”, which is $\Delta_0^0(\text{exp})$. Then φ^a is $\Delta_0^0(\text{exp})$ as in set theory. However, even Π_2^0 -**Ref** is inconsistent, since there is no finite set reflecting $(\forall x)(\exists y)(y = x + 1)$.

In both number and set theories, the clause $(\forall X)(\exists x)(\forall u)(u \in x \leftrightarrow u \in a \wedge u \in X)$ is redundant. However, to clarify that this fact is essential in the following discussion (and for the extension in the later section), we explicitly include this clause in the definition of the schema.

Lemma 3 For a Π_1^0 operator Γ arity (0, 1, 0), the following hold in **BT**² + Π_2^0 -**Ref**:

- (1) (i) $(\forall x)(\Gamma(x^c) \subset x^c \rightarrow F \subset x^c)$, with $x^c = \{z \mid z \notin x\}$, implies (ii) $(\forall X)(\Gamma(X) \subset X \rightarrow F \subset X)$.
- (2) Both (i) “ F is a least fixed point of Γ ” and (ii) $\text{WF}(W)$ are equivalently Π_3^0 .

Proof (1) Let $\Gamma = \lambda X.\{z \mid \varphi(z, X)\}$ be of Π_1^0 . Assume \neg (ii), say $y_0 \in F \setminus X$ but $(\forall y)(\varphi(y, X) \rightarrow y \in X)$.

$$(\forall y \in a)(\varphi^a(y, X) \rightarrow y \in X) \tag{*}$$

for some transitive $a \ni y_0$ containing all the set parameters in φ . By the redundant clause in Π_2^0 -**Ref**, with Δ_0^1 -**CA**, we can take $x = a \setminus X$. For $y \in a$, if $\varphi(y, x^c)$, by downward persistence of Π_1^0 , $\varphi^a(y, x^c)$ which implies $y \in x^c$ by (*) with $a \cap x^c = a \cap X$. For $y \notin a$, obviously $y \in x^c$. We have seen $(\forall y)(\varphi(y, x^c) \rightarrow y \in x^c)$. If (i) holds, $F \subset x^c$, contradicting $y_0 \in a \cap (F \setminus X)$.

(2)(i) is by (1), since “ $\Gamma(x^c) \subset x^c$ ” is Π_2^0 . (ii) $\text{WF}(W)$ is equivalent to “ $\{x \mid \top\}$ is a least fixed point of a positive Π_1^0 operator $\lambda X.\{z \mid (W)_z \subset X\}$ ”.

It might be worthwhile to emphasize that (2) can be proved without **Choice** in set theory, since Δ_0^1 -**Ref** is provable in **NBG** without choice, by the same proof as in **ZF** (see, e.g., [10, IV.7.4 Theorem]). In the absence of the axiom of foundation, on the other hand, **Choice** can substitute, i.e., Δ_0^1 -**Ref** is provable in **NBGC**⁻ by Skolemization argument. Thus the proof of (2) based on Δ_0^1 -**Ref** clarifies the common feature shared by **NBG** and **NBGC**⁻.

Proposition 4 **BT**² + Π_2^0 -**Ref** $\vdash \Pi_1^0$ -**LFP**. Moreover, **BT**² + Π_2^0 -**Ref** + Δ_0^1 -**TR** $\vdash \Pi_1^0$ -**LFTR**.

Proof Let Γ be Π_1^0 of arity (0, 1, 0). Define $F = \{z \mid (\forall x)((\Gamma(x^c) \subset x^c) \rightarrow z \in x^c)\}$. Then (i) in Lemma 3 (1) holds, and hence so does (ii). Thus it remains to show that $\Gamma(F) = F$.

For x with $\Gamma(x^c) \subset x^c$, since $F \subset x^c$, $\Gamma(F) \subset \Gamma(x^c) \subset x^c$, and, by arbitrariness of x , $\Gamma(F) \subset F$ and so $\Gamma(\Gamma(F)) \subset \Gamma(F)$. Applying Lemma 3 (1)(ii) to $X = \Gamma(F)$, we have $F \subset \Gamma(F)$.

Remark 5 (1) Lemma 3 (2) leads us to define an elementary formula Wf :

$$\text{Wf}(\succ) \equiv (\forall z)\text{TI}[\{x \mid x \notin z\}](\succ) \equiv (\forall z)[(\forall x)((\forall y \prec x)(y \notin z) \rightarrow x \notin z) \rightarrow (\forall x)(x \notin z)],$$

where \prec denotes $\{\langle x, y \rangle \mid \langle y, x \rangle \in \succ\}$, and the proof above shows that $\text{Wf}(\succ) \rightarrow \text{TI}[X](\succ)$.

(2) Proposition 4 means that Π_1^0 -**ID** _{α} is interpretable in **BT**¹ = **ZF(C)** with \mathcal{L}^1 preserved.

3.3 Reduction of Σ_1^0 inductive definition

We saw that $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ -**LFP** implies the whole Δ_0^1 -**LFP** (and $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ -**FP** implies the whole Δ_0^1 -**FP**) in Remark 3 (1), and hence so does \mathcal{F} -**LFP** (or \mathcal{F} -**FP**, respectively) for any class \mathcal{F} between $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ and Δ_0^1 . Proposition 4 asserts, on the other hand, \mathcal{F} -**LFP** (and hence \mathcal{F} -**FP**) for any class \mathcal{F} below Π_1^0 is provable in the base theory **BT**², if $\mathbf{BT}^2 \vdash \Pi_2^0$ -**Ref**. Is there any other “reasonable” class of formulae? (Strictly speaking, $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ is not a class of formulae, as remarked in Remark 1.) Obviously Σ_1^0 is among such classes. Is it the only one?

As is well known, the hierarchy consisting of Σ_n^0 's and Π_n^0 's (with parameters) corresponds to Borel hierarchy in $\omega^n \times (2^\omega)^m$. Among those hierarchies treated in descriptive set theory which are finer than Borel hierarchy is Wadge hierarchy, which is defined as follows. A subset A of a Polish space (e.g., $\omega^n \times (2^\omega)^m$ equipped with the product topology of the discrete spaces 2 and ω) is said to be Wadge reducible to another subset B , if there is a continuous function f such that $f^{-1}[B] = A$. Wadge hierarchy consists of those classes of subsets which are closed under Wadge reducibility (such classes are called topological classes).

Louveau [11, §1] gave a complete description of Wadge hierarchy up to some point (but on ω^ω), and Nemoto [13, §2.5] defined several determinacy schemata which formalize the determinacy for such topological classes, by rewording Louveau's description into terms of formula (where the difference of the base spaces are not so important). The hierarchy consisting of such classes of formulae should be called *Nemoto hierarchy*.

Since complement (corresponding to \neg), binary and countable unions (to \vee and first order \exists , resp.) and binary and countable intersections (to \wedge and first order \forall , resp.) are preserved under continuous preimages, i.e., Wadge reducibility, any class of formulae, if defined by forms of formulae, must correspond to a topological class. Thus, if we agree that the “reasonable” classes of formulae must be defined by forms of formulae, Nemoto hierarchy, the corresponding to Wadge hierarchy, is the finest “reasonable” hierarchy of classes of formulae.

Since Σ_1^0 is known to be the only Wadge (and so Nemoto) class not above $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ nor below Π_1^0 , we can say yes to the question above at least in number theory. And if we further require that “reasonable” formula-classes be defined uniformly both for number and set theories, we can answer yes in general.

Thus in order to have a complete description of Nemoto hierarchy of least fixed principles, it remains to determine Σ_1^0 -**LFP**. It is known that Σ_1^0 -**LFP** is provable in **ACA**₀, which is enough to claim that the general reduction (i.e., the uniform proof for both number and set theories) to Σ_1^0 is impossible. Actually, by more or less the same proof as in number theory we can prove the following.

Lemma 5 **NBG**[−] + Π_2^0 -**Ref** ⊢ Σ_1^0 -**LFP**. Moreover, **NBG**[−] + Π_2^0 -**Ref** + Δ_0^1 -**TR** ⊢ Σ_1^0 -**LFTR**.

Proof Let $\Gamma = \lambda X. \{z \mid \varphi(z, X)\}$ be Σ_1^0 of arity (0, 1, 0). Define $F = \{z \mid (\exists \langle \alpha, f \rangle \in G)(\exists \xi < \alpha)(z \in f(\xi))\}$,

$$\text{where } G = \left\{ \langle \alpha, f \rangle \mid \alpha \in \text{Ord} \wedge \text{“}f \text{ is a function on } \alpha\text{”} \wedge (\forall \xi < \alpha) \left(f(\xi) \subset \Gamma \left(\bigcup_{\eta < \xi} f(\eta) \right) \right) \right\}.$$

For X with $\Gamma(X) \subset X$ and $\langle \alpha, f \rangle \in G$, the induction on $\xi < \alpha$ shows $f(\xi) \subset X$ and hence $F \subset X$. Thus, if this F is a fixed point, then it is the least. It remains to show $\Gamma(F) \subset F$.

Take $z \in \Gamma(F)$. Π_2^0 -**Ref** provides us a transitive set a with $z \in a$ and $\varphi^a(z, F)$, i.e., $\varphi^a(z, F \cap a)$ such that $(\forall x \in a)[x \in F \leftrightarrow (\exists \langle \alpha, f \rangle \in G \cap a)(\exists \xi < \alpha)(x \in f(\xi))]$. Thus

$$(\forall x \in F \cap a)(\exists \langle \alpha, f \rangle \in G \cap a)(\exists \xi < \alpha)(x \in f(\xi)) \tag{*}$$

Define $\beta = \sup\{\alpha \mid (\exists f)(\langle \alpha, f \rangle \in G \cap a)\}$, and

$$g(\xi) = \bigcup\{f(\xi) \mid (\exists \eta \leq \beta)(\xi < \eta \wedge f \in (G \cap a)_\eta)\} \text{ for } \xi < \beta; \quad g(\beta) = F \cap a; \quad g(\beta + 1) = \{z\}.$$

Then $g(\beta) = F \cap a \subset \Gamma(\bigcup_{\xi < \beta} g(\xi))$, by (*). By upward persistence of Σ_1^0 , $\varphi^a(z, F \cap a)$ implies $z \in \Gamma(F \cap a)$, i.e., $g(\beta + 1) = \{z\} \subset \Gamma(F \cap a) = \Gamma(g(\beta))$. Thus $\langle \beta + 2, g \rangle \in G$ witnesses $z \in F$.

The only Nemoto class which changes its side is Π_1^0 : The border between those Nemoto classes \mathcal{F} with \mathcal{F} -**LFP** ↔ Δ_0^1 -**LFP** and those \mathcal{F} with **NBG** ⊢ \mathcal{F} -**LFP** is drawn in Figure 2 at page 4.

For \mathcal{F} -**FP**, since Σ_1^0 -**FP** is equivalent to Π_1^0 -**FP** (actually $(\neg\mathcal{F})$ -**FP** and \mathcal{F} -**FP** are equivalent), the border is immediately below $\Delta(\Sigma_1^0 \vee \Pi_1^0)$ both in number theory and in set theory.

4 Stage Comparison

The description of a least fixed point in the standard proof of Δ_0^1 -**LFP** in Π_1^1 -**CA** is “top-down”. It is known that alternative “bottom-up” description is more convenient in some occasions, i.e., the union of H_α ’s where $H_\alpha = \Gamma(\bigcup_{\xi < \alpha} H_\xi)$, or, H_α is the result of α -time application of Γ to \emptyset . However, in $\mathbf{BT}^2 = \mathbf{N BG}$ the von Neumann ordinals are not always enough to reach fixed points of operators on classes. Thus to formulate this description, we need a (class-size) well-order long enough to reach the fixed point, and define H by recursion along it. Since such well-orders might not be unique, the natural question is: What is the canonical? An answer is *stage comparison*: Intuitively, this compares the first stages α at which the given two elements are included in H_α . More precisely, letting $\|x\|_\Gamma = \min(\{\alpha \mid x \in H_\alpha\} \cup \{\infty\})$, we define $x \prec_\Gamma y$ by $\|x\|_\Gamma < \|y\|_\Gamma$. However, since we do not have single well-order coding α ’s above, such a definition would be circular. We need to extract, without mentioning α ’s, properties of stage comparison which guarantee the canonicity.

4.1 Formalizing stage comparison strict preorder

The properties, that we extract, of stage comparison \prec of Γ are (i) \prec is well-founded ($\text{WF}(\succ)$); (ii) if $y \notin \Gamma(\prec z) \exists x$ for some z then $x \prec y$; and (iii) $(\prec x) = \bigcup_{y \prec x} \Gamma(\prec y)$, where $(\prec x) = \{z \mid z \prec x\}$. We now formalize the notion with these properties except (i):

Definition 11 (SCSP) For an operator Γ of arity $(0, 1, 0)$, a $\Delta_0^1(\Gamma)$ -formula $\text{SCSP}[\Gamma](A)$ (standing for Stage Comparison Strict Preorder) is the conjunction of (a) $(\forall r \in A)(\exists u, v)(r = \langle u, v \rangle)$; (b) $(\exists w)(v \notin \Gamma((A)_w) \wedge u \in \Gamma((A)_w)) \rightarrow u \in (A)_v$ and (c) $u \in (A)_v \leftrightarrow (\exists w)(w \in (A)_v \wedge u \in \Gamma((A)_w))$.

We will see that $\text{SCSP}[\Gamma](A) \wedge \text{WF}(A)$ determines A uniquely, if exists.

Lemma 6 For an operator Γ of arity $(0, 1, 0)$, if $\text{SCSP}[\Gamma](A) \wedge \text{WF}(A)$, then $\{x \mid (\exists v)(x \in \Gamma((A)_v))\}$ exists as a second order object (i.e., $(\exists X)(X = \{x \mid (\exists v)(x \in \Gamma((A)_v))\})$) and is a least fixed point of Γ .

Proof Let $\Gamma(X) \subset X$. We prove $(A)_u \subset X$ by induction on u along A . (Note that Δ_0^0 transfinite induction is enough, regardless of the complexity of Γ .) By the last clause (c) of $\text{SCSP}[\Gamma](A)$, $(A)_u = \bigcup_{v \in (A)_u} \Gamma((A)_v) \subset \Gamma(X) \subset X$ by induction hypothesis.

If $(\forall x)(\exists v)(x \in \Gamma((A)_v))$, then $\{x \mid (\exists v)(x \in \Gamma((A)_v))\} = \{x \mid x = x\}$ is a least fixed point.

Otherwise $(\forall v)(y \notin \Gamma((A)_v))$ for some y . For any u, v , if $u \in \Gamma((A)_v)$, v witnesses $u \in (A)_y$. Thus $\{x \mid (\exists v)(x \in \Gamma((A)_v))\} \subset (A)_y$, which implies $\Gamma((A)_y) \subset (A)_y$. By what we have shown at the first with $X = (A)_y$, we have $(\forall v)((A)_v \subset (A)_y)$. Therefore $(A)_y = \bigcup_{v \in (A)_y} \Gamma((A)_v) \subset \{x \mid (\exists v)(x \in \Gamma((A)_v))\} \subset \Gamma((A)_y) \subset (A)_y$. $(A)_y$ is a least fixed point and is $\{x \mid (\exists v)(x \in \Gamma((A)_v))\}$.

Thus we can define a seemingly stronger variant of least fixed point principle.

Definition 12 (\mathcal{F} -SLFP, \mathcal{F} -SLFTR) The *strong least fixed point principle* and *strong least fixed point transfinite recursion* for a class \mathcal{F} of formulae are

\mathcal{F} -SLFP $(\exists O)(\text{SCSP}[\Gamma](O) \wedge \text{WF}(O))$ for any \mathcal{F} operator Γ of arity $(0, 1, 0)$;

\mathcal{F} -SLFTR $\text{WF}(W) \rightarrow (\exists H)\text{SLFHier}[\Gamma](H, W)$ for any \mathcal{F} operator Γ of arity $(1, 1, 1)$,

where $\text{SLFHier}[\Gamma](H, W) \equiv (\forall x)(\text{SCSP}[\lambda X.\Gamma(x, X, (H)^{W_x})]((H)_x) \wedge \text{WF}((H)_x))$;

\mathcal{F} -SLFTR $_\alpha$ is defined accordingly.

If a class \mathcal{F} of operators are closed under compositions with Δ_0^0 operators, since by composing an operator with $X \mapsto \{x \mid (x \in \omega \rightarrow x \neq 0 \wedge x - 1 \in X) \vee (x \notin \omega \wedge x \in X)\}$ and with the converse, we can restrict ourselves to consider only those Γ with $0 \notin \Gamma(X)$ without loss of generality, and then the least fixed point is $(A)_0$, defined by a uniform formula from the given stage comparison strict preorder.

Corollary 2 For any class \mathcal{F} of formulae,

(i) $\mathbf{BT}^2 + \mathcal{F}\text{-SLFP} \vdash \mathcal{F}\text{-LFP}$,

and if \mathcal{F} is closed under compositions with Δ_0^0 operators,

(ii) $\mathbf{BT}^2 + \mathcal{F}\text{-SLFTR} \vdash \mathcal{F}\text{-LFTR}$ and

(iii) $\mathbf{BT}^2 + \mathcal{F}\text{-SLFTR}_\alpha \vdash \mathcal{F}\text{-LFTR}_\alpha$.

Remark 6 If we define a first order correspondence of strong least fixed point principle, say ID_α^+ , the proof shows that ID_α can be interpreted into ID_α^+ in such a way that \mathcal{L}^1 is preserved.

4.2 Uniqueness of stage comparison

Lemma 7 For an operator Γ of arity $(0, 1, 0)$, $\text{SCSP}[\Gamma](A) \wedge \text{WF}(A)$ implies:

- (i) $(\forall u, v, w)(u \in (A)_v \wedge v \in (A)_w \rightarrow u \in (A)_w)$ (transitivity of A),
- (ii) $(\forall u, v)[(\exists w)(v \notin \Gamma((A)_w) \wedge u \in \Gamma((A)_w)) \leftarrow u \in (A)_v]$ (the converse of (b)).

Proof (i) Let $u \in (A)_v \wedge v \in (A)_w$. Suppose for contradiction $u \notin (A)_w$. Then by $v \in (A)_w$ and (c) in $\text{SCSP}[\Gamma](A)$ we have $v' \in (A)_w$ with $v \in \Gamma((A)_{v'})$. Then, again by (c), $u \notin \Gamma((A)_{v'})$ which witnesses $v \in (A)_u$, contradicting $u \in (A)_v$ and $\text{WF}(A)$.

(ii) Let $u \in (A)_v$. Then by the clause (c) of $\text{SCSP}[\Gamma](A)$ we have $w \in (A)_v$ with $u \in \Gamma((A)_w)$. If $v \in \Gamma((A)_w)$, then, by (c) again, $v \in (A)_v$, contradicting $\text{WF}(A)$.

Definition 13 For an operator Γ of arity $(0, 1, 0)$, define $\Delta_0^1(\Gamma)$ -formulae

$$\begin{aligned}\text{ItApp}[\Gamma](H, W) &\equiv (\forall u)((H)_u = \Gamma((H)_{Wu})) \text{ where } (H)_{Wu} = \{x \mid (\exists v \in (W)_u)(x \in (H)_v)\}; \\ \theta_1[\Gamma](y, x, H, W) &\equiv \text{ItApp}[\Gamma](H, W) \wedge (\exists u)(x \in (H)_u \wedge y \notin (H)_u); \\ \theta_2[\Gamma](y, x, H, W) &\equiv \text{ItApp}[\Gamma](H, W) \rightarrow (\forall u)(y \in (H)_u \rightarrow x \in (H)_{Wu}).\end{aligned}$$

Notice the difference between $(H)_{Wu}$ and $(H)^{Wx}$ (introduced below Definition 7).

Lemma 8 For an operator Γ of arity $(0, 1, 0)$, if $\text{WF}(R) \wedge \text{WF}(S)$ then:

- (i) $\text{ItApp}[\Gamma](G, R) \wedge \text{ItApp}[\Gamma](H, S) \rightarrow (\forall r, s)((G)_r \subset (H)_{Ss} \vee (H)_s \subset (G)_r)$,
- (ii) $(\forall x, y)(\theta_1[\Gamma](y, x, G, R) \rightarrow \theta_2[\Gamma](y, x, H, S))$.

Proof (i) We prove $((G)_r \subset (H)_{Ss} \vee (H)_s \subset (G)_r) \wedge ((G)_r \subset (H)_s \vee (H)_s \subset (G)_{Rr})$ by induction on r, s . If there is $s' \in (S)_s$ with $(G)_r \subset (H)_{s'}$, then $(G)_r \subset (H)_{Ss}$. Otherwise for any $s' \in (S)_s$, by induction hypothesis for (r, s') , we have $(H)_{s'} \subset (G)_{Rr}$. Thus $(H)_{Ss} \subset (G)_{Rr}$ and $(H)_s = \Gamma((H)_{Ss}) \subset \Gamma((G)_{Rr}) = (G)_r$. Similarly we have $(G)_r \subset (H)_s \vee (H)_s \subset (G)_{Rr}$.

(ii) Let $\theta_1[\Gamma](y, x, G, R)$, say $x \in (G)_r \not\ni y$, and assume $\text{ItApp}[\Gamma](H, S)$ and $y \in (H)_s$. y witnesses $(H)_s \not\subset (G)_r$, and hence $(G)_r \subset (H)_{Ss}$. Since $x \in (G)_r$, we have $x \in (H)_{Ss}$.

Corollary 3 $(\forall x, y)[(\exists G, R)(\text{WF}(R) \wedge \theta_1[\Gamma](y, x, G, R)) \rightarrow (\forall H, S)(\text{WF}(S) \rightarrow \theta_2[\Gamma](y, x, H, S))]$.

Lemma 9 $\text{SCSP}[\Gamma](O)$ implies $\text{ItApp}[\Gamma](H, O)$ if $H = \{\langle u, x \rangle \mid x \in \Gamma((O)_u)\}$.

Proof By $(O)_u = \bigcup_{v \in (O)_u} \Gamma((O)_v) = \bigcup_{v \in (O)_u} (H)_v = (H)_{Ou}$, we have $(H)_u = \Gamma((O)_u) = \Gamma((H)_{Ou})$.

So far the argument requires only **BT**² even if the operator Γ is of higher complexity. In the next theorem, however, we need some principle for the complexity of Γ . Since the use of such principles is in transfinite induction for G, H in Lemma 8 and in the minimal element principle in the following proof, $\Delta_0^1(\Gamma)$ -**CA** below can actually be weakened to $\Delta_0^1(\Gamma)$ -**TI**.

Theorem 4 For an operator Γ of arity $(0, 1, 0)$,

$$\mathbf{BT}^2 + \Delta_0^1(\Gamma)\text{-}\mathbf{CA} \vdash \text{WF}(A) \wedge \text{WF}(B) \wedge \text{SCSP}[\Gamma](A) \wedge \text{SCSP}[\Gamma](B) \rightarrow A = B.$$

Proof Assume the antecedents and define $G = \{\langle u, x \rangle \mid x \in \Gamma((A)_u)\}$ and $H = \{\langle u, x \rangle \mid x \in \Gamma((B)_u)\}$ by $\Delta_0^1(\Gamma)$ -**CA**. Then $\text{ItApp}[\Gamma](G, A)$ and $\text{ItApp}[\Gamma](H, B)$ by the last lemma.

Let $u \in (A)_v$. By Lemma 7 (ii), $(\exists w)(v \notin \Gamma((A)_w) = (G)_w \ni u)$. Thus $\theta_1[\Gamma](v, u, G, A)$. By Lemma 8 (ii), $\theta_2[\Gamma](v, u, H, B)$ and so $(\forall w)(v \in (H)_w = \Gamma((B)_w) \rightarrow u \in (H)_{Bw})$.

If $v \in \Gamma((B)_w)$ for some w , we can take a B -minimal such w . Then $u \in (H)_{Bw}$, say $u \in (H)_{w'} = \Gamma((B)_{w'})$ with $w' \in (B)_w$. By the choice of w , we have $v \notin \Gamma((B)_{w'})$, and thus, by (b) of the definition of $\text{SCSP}[\Gamma](B)$ we have $u \in (B)_v$. Otherwise v is outside of the least fixed point $\{z \mid (\exists w)(z \in \Gamma((B)_w))\} = \{z \mid (\exists w)(z \in \Gamma((A)_w))\}$ which contains u , and so $u \in (B)_v$ by (b) of the definition of $\text{SCSP}[\Gamma](B)$.

4.3 Stage comparison as fixed point

Informally, we have the following equivalence for x, y with $\|x\|_\Gamma, \|y\|_\Gamma < \infty$:

$$x \prec_\Gamma y \leftrightarrow \neg(\|y\|_\Gamma \leq \|x\|_\Gamma) \leftrightarrow \neg(y \in \Gamma^{\|x\|_\Gamma}(\emptyset)) \leftrightarrow y \notin \Gamma(\{z \mid z \prec_\Gamma x\}).$$

Therefore, one may think that the stage comparison can be obtained as a fixed point of Γ' , where $\Gamma'(R) = \{\langle y, x \rangle \mid \neg(y \in \Gamma(\{z \mid \langle x, z \rangle \in R\}))\}$. However, this Γ' is negative. Nevertheless, replacing “ $z \prec x$ ” in the right-hand side by using this equivalence itself, we have

$$x \prec_\Gamma y \leftrightarrow y \notin \Gamma(\{z \mid z \prec_\Gamma x\}) \leftrightarrow y \notin \Gamma(\{z \mid x \notin \Gamma(\{z' \mid z' \prec_\Gamma z\})\}).$$

Thus we can expect that \prec_Γ can be obtained as a fixed-point of the positive operator $(\Gamma')^2$.

Actually, the famous proof of Stage Comparison Theorem (e.g., [12, 2A.2], where “fixed point” means “least fixed point” in our terminology) tells us that the stage comparison \prec_Γ of Γ is the least fixed point of $(\Gamma')^2$. However $\mathcal{F}\text{-}\mathbf{FP}$ does not give us a least fixed point, but only a fixed point. Thus we have to check how close to the stage comparison we can reach only with a (not necessarily least) fixed point.

Definition 14 (derivation of operator) For an operator Γ of arity $(0, 1, 0)$, the *derived operator* is $\Gamma' \circ \Gamma'$, where $\Gamma' = \lambda R. \{u \mid (\exists x, y)(u = \langle x, y \rangle \wedge x \notin \Gamma((R)_y))\}$.

Lemma 10 For an operator Γ of arity $(0, 1, 0)$, if $\Gamma'(F) = G, \Gamma'(G) = F$, $\text{WF}(W)$ and $\text{ItApp}[\Gamma](H, W)$ hold, then $(\forall w)((R)_w \subset F \wedge G \subset (S)_w)$, where

$$(R)_w = \{\langle y, x \rangle \mid x \in (H)_w \wedge y \notin (H)_w\} \text{ and } (S)_w = \{\langle y, x \rangle \mid y \in (H)_w \rightarrow x \in (H)_{Ww}\}.$$

Proof We prove (a) $(R)_w \subset F$ and (b) $G \subset (S)_w$ simultaneously by induction on w along W .

For (b), let $\langle y, x \rangle \in G = \Gamma'(F)$, i.e., $y \notin \Gamma((F)_x)$. To show the contraposition, assume $x \notin (H)_{Ww}$. The induction hypothesis (a) implies $\bigcup_{u \in (W)_w} (R)_u \subset F$ and hence

$$(H)_{Ww} = \bigcup_{u \in (W)_w} (H)_u = \bigcup_{u \in (W)_w} \{(H)_u \mid u \in (W)_w \wedge x \notin (H)_u\} \subset \bigcup_{u \in (W)_w} ((R)_u)_x \subset (F)_x.$$

Then $y \notin \Gamma((F)_x) \supset \Gamma((H)_{Ww}) = (H)_w$.

For (a), let $\langle y, x \rangle \in (R)_w$. By (b) and by $x \in (H)_w$, $(G)_x \subset ((S)_w)_x = (H)_{Ww}$ and so $y \notin (H)_w = \Gamma((H)_{Ww}) \supset \Gamma((G)_x)$, which implies $\langle y, x \rangle \in \Gamma'(G) = F$.

Theorem 5 For an operator Γ of arity $(0, 1, 0)$, \mathbf{BT}^2 proves: if $(\Gamma')^2(F) = F$ and $\text{WF}(W)$ then

$$(\forall x, y)[(\theta_1[\Gamma](y, x, H, W) \rightarrow \langle y, x \rangle \in F) \wedge (\langle y, x \rangle \in F \rightarrow \theta_2[\Gamma](y, x, H, W))].$$

Proof Assume $(\Gamma')^2(F) = F$ and $\text{WF}(W)$. Let $\text{ItApp}[\Gamma](H, W)$. We have to show that, for any w ,

$$(x \in (H)_w \wedge y \notin (H)_w) \rightarrow \langle y, x \rangle \in F \text{ and } \langle y, x \rangle \in F \rightarrow (y \in (H)_w \rightarrow x \in (H)_{Ww}),$$

i.e., $(R)_w \subset F \subset (S)_w$ in terms of the last lemma. $(R)_w \subset F$ is from the last lemma with $G = \Gamma'(F)$. Since F and G are interchangeable in the lemma, we also have $F \subset (S)_w$.

This seems to be the best we can say about a fixed point F of $(\Gamma')^2$, and F itself is not necessarily the stage comparison of Γ . Nevertheless, as Stage Comparison Theorem tells us (on the informal level), the stage comparison strict preorder is a least fixed point and therefore a subset of F . Main Lemma in the next section tells us how to cut out the stage comparison from F and the proof is based on this “sandwich” property. (However, Main Lemma states more: We can cut out the stage comparison from any F with this “sandwich” property, and we do not need to check if such F is a fixed point of $(\Gamma')^2$.)

Corollary 6 For an operator Γ of arity $(0, 1, 0)$, if $\Gamma' \circ \Gamma'(R) = R$ and if O is defined from R as in Main Lemma then $\text{SCSP}[\Gamma](O)$.

This is not significant in number theory (i.e., the case $\mathbf{BT}^2 = \mathbf{ACA}_0$) since to take the accessible part of R , as required in Main Lemma, we need some principle as strong as Δ_0^1 inductive definition. However this is significant in set theory, since, as we saw or as is well known, the accessible part is Δ_0^1 definable and therefore available in the base theory.

As for the first order systems, the proofs above and the proof of Main Lemma given below actually show that, if $\mathbf{BT}^1 \vdash \mathcal{L}^1\text{-}\mathbf{Ref}$, ID_α^+ , defined as in Remark 6 over \mathbf{BT}^1 , is interpretable into \widehat{ID}_α (and therefore so is ID_α) defined over \mathbf{BT}^1 in such a way that \mathcal{L}_S^1 is preserved.

5 Main Results

5.1 Main Lemma and its consequences

Theorem 7 (Main Lemma) *Let Γ be of $(0, 1, 0)$. The following is provable in $\mathbf{BT}^2 + \Delta_0^1(\Gamma)\text{-}\mathbf{TI}$.
For given R , if R^*, P, O and H satisfy the following:*

$$\begin{aligned} R^* &= \{\langle x, y \rangle \mid (\exists \sigma \in \text{Finseq})(\sigma(0) = x \wedge \sigma(|\sigma| - 1) = y \wedge (\forall k < |\sigma| - 1)(\langle \sigma(k), \sigma(k + 1) \rangle \in R))\}; \\ P &= \{x \mid "x \text{ is in the accessible part of } R" \wedge (\forall z \in (R^*)_x)(z \in \Gamma((R)_z))\}; \\ O &= (R \cap (P \times P)) \cup (P^c \times P) = \{\langle x, y \rangle \mid (x, y \in P \wedge \langle x, y \rangle \in R) \vee (x \notin P \wedge y \in P)\}; \\ H &= \{\langle x, y \rangle \mid (x \in P \wedge y \in \Gamma((O)_x)) \vee (x \notin P \wedge y \in \Gamma(\bigcup_{z \in P} \Gamma((O)_z)))\}, \end{aligned}$$

then following holds, where $x \downarrow = (R^*)_x$; $O \downarrow x = O \cap (x \downarrow \times x \downarrow)$; $H \upharpoonright x = \{\langle u, z \rangle \in H \mid u \in x \downarrow\}$:

$$\begin{aligned} (\forall x, y, z)[(\theta_1[\Gamma](y, x, H \upharpoonright z, O \downarrow z) \rightarrow \langle y, x \rangle \in R) \wedge (\langle y, x \rangle \in R \rightarrow \theta_2[\Gamma](y, x, H \upharpoonright z, O \downarrow z))] \\ \rightarrow \text{WF}(O) \wedge \text{SCSP}[\Gamma](O) \wedge P = \{x \mid (\exists v)(x \in \Gamma((O)_v))\}. \end{aligned}$$

Remark 7 (1) As one can see, R^* denotes the reflexive transitive closure of R . The others P , O and H are $\Delta_0^1(\mathcal{F})$ definable from R and the accessible part of R , and hence can be seen as $\Delta_0^1(\mathcal{F})$ -formulae containing the accessible part as a parameter. Thus this theorem provides a sufficient condition for us to construct the stage comparison strict preorder.

(2) In the first order setting mentioned in Remark 6, we can interpret ID_α^+ in \widehat{ID}_α with a new predicate (and axiom) for the accessible parts. More generally, the iterated inductive definition whose last stage is ID^+ -type is interpretable in the same but the last stage replaced by the two stages, the first of which is \widehat{ID} -type and the second is ID -type (restricted to Π_1^0 operators). See Remark 11 (IV) in Section 8.

Corollary 8 (1) For any operator Γ of arity $(0, 1, 0)$, $\mathbf{BT}^2 + \Delta_0^1(\Gamma, \text{WF})\text{-}\mathbf{CA}$ proves:

$$(\exists R)(\forall x, y) \left(\begin{array}{l} ((\exists H, W)(\text{WF}(W) \wedge \theta_1[\Gamma](y, x, H, W)) \rightarrow \langle y, x \rangle \in R) \\ \wedge (\langle x, y \rangle \in R \rightarrow (\forall H, W)(\text{WF}(W) \rightarrow \theta_2[\Gamma](y, x, H, O))) \end{array} \right) \rightarrow (\exists O) \left(\begin{array}{l} \text{SCSP}[\Gamma](O) \\ \wedge \text{WF}(O) \end{array} \right).$$

(2) For a class \mathcal{F} of formulae, $\mathbf{BT}^2 + \Pi_1^1(\mathcal{F}, \text{WF})\text{-}\mathbf{Red} \vdash \mathcal{F}\text{-SLFP}$.

(3) If \mathcal{F} is closed under derivations,

$$\mathbf{BT}^2 + \Pi_1^0\text{-}\mathbf{LFP} \vdash \mathcal{F}\text{-SLFP} \leftrightarrow \mathcal{F}\text{-LFP} \leftrightarrow \mathcal{F}\text{-FP}.$$

Proof (1) Immediate from Theorem 7. (2) Since $\theta_i[\Gamma]$'s are $\Delta_0^1(\Gamma)$, this is by Corollary 3.

(3) By rephrasing Corollary 6, we have this since $\Pi_1^0\text{-}\mathbf{LFP}$ gives us the accessible parts.

Actually, we can prove Corollary 8 (1) directly and it might be more comprehensible than Main Lemma. The reason why we formulate Main Lemma as above is, however, to make it easier to see that our argument survives also in the first order setting. For, there are no first order analogues of $(\exists H, W)(\text{WF}(W) \wedge \theta_1[\Gamma](y, x, H, W))$ and of $(\forall H, W)(\text{WF}(W) \rightarrow \theta_2[\Gamma](y, x, H, O))$. In this sense, Main Lemma is a quite refined and seemingly stronger result.

Let us discuss the significances of this corollary with $\mathcal{F} = \Delta_0^1$. In number theory, these are less significant: $\Pi_1^0\text{-}\mathbf{LFP}$ itself is strong enough to imply $\Pi_1^1\text{-}\mathbf{CA}$; since WF is Π_1^1 -complete, $\Pi_1^1(\text{WF})\text{-}\mathbf{Red}$ is equivalent to $\Pi_2^1\text{-}\mathbf{Red}$ while $\Delta_0^1\text{-}\mathbf{SLFP}$ is proved by $\Pi_1^1\text{-}\mathbf{CA}$. However, in set theory (and in higher order number or set theory with global well-ordering [17, Definition 5]), these are: Since $\Pi_1^0\text{-}\mathbf{LFP}$ is proven and WF is equivalent to Wf which is Δ_0^1 , these mean that $\Delta_0^1\text{-}\mathbf{SLFP}$, $\Delta_0^1\text{-}\mathbf{LFP}$ and $\Delta_0^1\text{-}\mathbf{FP}$ are all equivalent; and that $\Pi_1^1\text{-}\mathbf{Red}$ implies $\Delta_0^1\text{-}\mathbf{SLFP}$ and hence $\Delta_0^1\text{-}\mathbf{LFP}$.

Letting $\mathcal{F} = \Delta_{n+2}^1$ in Corollary 8 (2), we have $\mathbf{BT}^2 + \Pi_{n+2}^1\text{-}\mathbf{Red} \vdash \Delta_{n+2}^1\text{-}\mathbf{SLFP}$, which might be new even in number theory, where $\Delta_{n+2}^1\text{-}\mathbf{SLFP}$ is as in Remark 1. Among examples of \mathcal{F} applicable to Corollary 8 (3) are essential Δ_n^1 's (i.e., non-prenex Δ_n^1 's). Note that [4] shows $\Delta_0^1\text{-}\mathbf{FTR}_\alpha \vdash \text{WF}(\alpha)$.

Corollary 9 (i)(a) $\mathbf{BT}^2 \vdash \Delta_0^1\text{-}\mathbf{SLFP} \leftrightarrow \Delta_0^1\text{-}\mathbf{LFP}$; (b) $\mathbf{BT}^2 + \Pi_1^0\text{-}\mathbf{LFP} \vdash \Delta_0^1\text{-}\mathbf{LFP} \leftrightarrow \Delta_0^1\text{-}\mathbf{FP}$.
(ii)(a) $\mathbf{BT}^2 \vdash \Delta_0^1\text{-}\mathbf{SLFTR} \leftrightarrow \Delta_0^1\text{-}\mathbf{LFTR}$; (b) $\mathbf{BT}^2 + \Pi_2^0\text{-}\mathbf{Ref} \vdash \Delta_0^1\text{-}\mathbf{LFTR} \leftrightarrow \Delta_0^1\text{-}\mathbf{FTR}$.
(iii)(a) $\mathbf{BT}^2 \vdash \Delta_0^1\text{-}\mathbf{SLFTR}_\alpha \leftrightarrow \Delta_0^1\text{-}\mathbf{LFTR}_\alpha$; (b) $\mathbf{BT}^2 + \Pi_2^0\text{-}\mathbf{Ref} \vdash \Delta_0^1\text{-}\mathbf{LFTR}_\alpha \leftrightarrow \Delta_0^1\text{-}\mathbf{FTR}_\alpha$.

Proof For (ii)(a) and (iii)(a), we can replace one step of taking stage comparison strict preorder (in $\Delta_0^1\text{-SLFTR}$) by the following two stages, that is, taking a (least) fixed point of derived operator, and then taking the accessible part as required in Main Lemma (in $\Delta_0^1\text{-LFTR}$). Thus $\Delta_0^1\text{-LFTR}_{2,\alpha} \rightarrow \Delta_0^1\text{-SLFTR}_\alpha$. This is enough, because $2\cdot\alpha = \alpha$ if α is limit, and otherwise $2\cdot\alpha = \alpha + k$ for some $k \in \omega$.

For (ii)(b) and (iii)(b) $\Pi_2^0\text{-Ref}$ allows us to define the accessible parts by the Δ_0^1 -formula Wf .

5.2 Proof of Main Lemma

Now we are proving the Main Lemma. We are working in \mathbf{BT}^2 (as before). Let us assume that R^* , P , O and H are as in the premise. Assume (b) and (‡) below for all x, y, z :

$$(b) \theta_1[\Gamma](y, x, H \upharpoonright z, O \downarrow z) \rightarrow \langle y, x \rangle \in R \quad (\ddagger) \langle y, x \rangle \in R \rightarrow \theta_2[\Gamma](y, x, H \upharpoonright z, O \downarrow z).$$

Lemma 11 (a) If $y \in P$ then $(R)_y = (O)_y$. (b) If $y \in P$ and $y \in (R)_x$ then $y \in (O)_x$. (c) If $x \in (O)_y$ for some y then $x \in P$. (d) $\text{WF}(O)$.

Proof First note that P is R -downward closed, i.e., $x \in P$ implies $(R)_x \subset P$. Then (a) - (c) follow from the definitions immediately. (d) holds since we define O by restricting R to a subset P of its accessible part and by putting all elements outside of P on the top (as maximal elements).

Lemma 12 $\text{ItApp}(H, O)$, and for any $x \in P$, $(H)_{Ox} = (O)_x$.

Proof First we prove (a) $(H)_{Ox} = (O)_x$ and (b) $\text{ItApp}(H \upharpoonright x, O \downarrow x)$ by induction on $x \in P$ along O .

(a) For $y \in (O)_x$, Lemma 11 (c) implies $y \in P$ and, by the definition of P , $y \in \Gamma((R)_y) = \Gamma((O)_y) = (H)_y \subset (H)_{Ox}$ since $(R)_y = (O)_y$ by Lemma 11 (a). Conversely, let $y \in (H)_{Ox}$, say $y \in (H)_z$ and $z \in (O)_x$. If $x \notin (H)_z$ then $\theta_1[\Gamma](x, y, H \upharpoonright z, O \downarrow z)$, by the induction hypothesis (b) for z , and so by (b) and $x \in P$, $y \in (R)_x = (O)_x$. Suppose $x \in (H)_z$ for contradiction. By (‡), $z \in (O)_x = (R)_x$ implies $\theta_2[\Gamma](x, z, H \upharpoonright z, O \downarrow z)$ and, by $\text{ItApp}[\Gamma](H \upharpoonright z, O \downarrow z)$ the induction hypothesis (b), $z \in (H)_{Oz} = (O)_z$ by induction hypothesis (a) for z , contradicting Lemma 11 (d).

(b) We have to show $(\forall y \in x \downarrow)(H)_y = \Gamma((H)_{Oy})$. By induction hypothesis, it remains to show $(H)_x = \Gamma((H)_{Ox})$, which follows from (a) $(H)_{Ox} = (O)_x$ and the definition of H .

To complete the proof, we have to show $(H)_x = \Gamma((H)_{Ox})$ for $x \notin P$. If $x \notin P$ then $(O)_x = P$ and so $(H)_x = \Gamma(\bigcup_{z \in P} \Gamma((O)_z)) = \Gamma(\bigcup_{z \in (O)_x} (H)_z) = \Gamma((H)_{Ox})$.

Lemma 13 (a) For all x , $(H)_x \subset P$ and (b) $P = \bigcup_{x \in P} \Gamma((O)_x) = \Gamma(P)$.

Proof (a) We prove this by induction on x along O . Let $y \in (H)_x$. Since if $y \in (H)_{Ox}$ then by induction hypothesis $y \in P$, we may assume $y \notin (H)_{Ox}$. We have to show $y \in P$.

To see that y in the accessible part, it suffices to show $(R)_y \subset P$. Let $z \in (R)_y$. Then, by (‡), $\theta_2[\Gamma](y, z, H \upharpoonright x, O \downarrow x)$, by which $y \in (H)_x$ implies $z \in (H)_{Ox} \subset P$ by induction hypothesis.

It remains to show $(\forall z \in y \downarrow)(z \in \Gamma((R)_z))$. Since we have shown $(R)_y \subset P \subset \{z \mid z \in \Gamma((R)_z)\}$ and since P is R -downward closed, what we have to show is $y \in \Gamma((R)_y)$.

To prove this, it suffices to show $(H)_{Ox} \subset (R)_y$, since $y \in (H)_x = \Gamma((H)_{Ox}) \subset \Gamma((R)_y)$ follows. Take $z \in (H)_{Ox}$. By $y \notin (H)_{Ox}$, we have $\theta_1[\Gamma](y, z, H \upharpoonright x, O \downarrow x)$ and so $z \in (R)_y$ by (b).

(b) (i) $P \subset \bigcup_{x \in P} \Gamma((O)_x)$: For $x \in P$, by definition and Lemma 11 (a), $x \in \Gamma((R)_x) = \Gamma((O)_x)$.

(ii) $\bigcup_{x \in P} \Gamma((O)_x) \subset \Gamma(P)$: For $x \in P$, by Lemma 12 and (a), $\Gamma((O)_x) = \Gamma((H)_{Ox}) \subset \Gamma(P)$.

(iii) $\Gamma(P) \subset P$: If there is $y \notin P$, by (a) and (i), $P \supset (H)_y = \Gamma(\bigcup_{x \in P} \Gamma((O)_x)) \supset \Gamma(P)$.

Proof of Theorem 7. We first see that $\text{SCSP}[\Gamma](O)$ holds. (a) of Definition 11 is obviously satisfied.

To see (b) of $\text{SCSP}[\Gamma](O)$, assume $x \in \Gamma((O)_z)$ and $y \notin \Gamma((O)_z)$. If $z \notin P$, then $\Gamma((O)_z) = \Gamma(P) = P$ by Lemma 13 (b) and hence $x \in (O)_y$ by the definition of O . Otherwise $x \in (H)_z$ and $y \notin (H)_z$, which means $\theta_1[\Gamma](y, x, H \upharpoonright z, O \downarrow z)$ and $x \in (R)_y$. Now, by Lemma 13 (a), we have $x \in (H)_z \subset P$ and, by Lemma 11 (b), $x \in (O)_y$.

Lemma 12 implies (c) $(O)_x = \bigcup_{y \in (O)_x} \Gamma((O)_y)$ for $x \in P$. If $x \notin P$, Lemma 13 (b) implies $(O)_x = P = \bigcup_{y \in P} \Gamma((O)_y) = \bigcup_{y \in (O)_x} \Gamma((O)_y)$.

Next we see $P = \{x \mid (\exists v)(x \in \Gamma((O)_v)\}$. Lemma 13 (b) implies one inclusion: $P \subset \{x \mid (\exists v)(x \in \Gamma((O)_v)\}$. For the converse, Lemma 13 (a) implies that, for $v \in P$, $\Gamma((O)_v) \subset P$, and Lemma 13 (b) implies that, for $v \notin P$, $\Gamma((O)_v) = \Gamma(P) = P$.

6 Strictness of $\Pi_1^1(\text{WF})\text{-Red} \rightarrow \Delta_0^1\text{-SLFP}$

In Corollary 8, we have seen that $\Pi_1^1(\text{WF})\text{-Red}$ implies $\Delta_0^1\text{-SLFP}$. In this section we prove that this implication is strict, or, more precisely, $\Pi_1^1(\text{WF})\text{-Red}$ proves the consistency of $\Delta_0^1\text{-SLFP}$. The proof is divided into two parts: In the second part, we show the strict hierarchy theorem for $\Delta_0^1\text{-SLFTR}_{\omega^\alpha}$ or, equivalently, for $\Delta_0^1\text{-LFTR}_{\omega^\alpha}$, i.e., if $\alpha \prec \beta$ then $\Delta_0^1\text{-SLFTR}_{\omega^\beta}$ implies the consistency of $(\text{BT}^2 + \Delta_0^1\text{-SLFTR}_{\omega^\alpha})$; in the first part we show that $\Pi_1^1(\text{WF})\text{-Red}$ is located above this hierarchy, i.e., $\Pi_1^1(\text{WF})\text{-Red}$ implies $\Delta_0^1\text{-SLFTR}$ and therefore $\Delta_0^1\text{-SLFTR}_{\omega^\alpha}$ for any reasonable α (which includes $\Delta_0^1\text{-SLFP}$ as the case of $\alpha = 0$, which means $\omega^\alpha = 1$).

6.1 Above the hierarchy

Proposition 14 (1) $\text{BT}^2 + \text{WF}(\alpha) + \Pi_1^1(\mathcal{F}, \text{WF})\text{-Red} \vdash \mathcal{F}\text{-SLFTR}_\alpha$ and
(2) $\text{BT}^2 + \Pi_1^1(\mathcal{F}, \text{WF})\text{-Red} \vdash \mathcal{F}\text{-SLFTR}$.

Proof Let Γ be of arity $(1, 1, 1)$. Below let $\Gamma(x, Y)$ denote $\lambda X.\Gamma(x, X, Y)$.

We prove $(\forall \xi \prec \alpha)(\forall u, v)(\theta_1(\xi, u, v) \rightarrow \theta_2(\xi, u, v))$, where $\theta_1(\xi, u, v)$ and $\theta_2(\xi, u, v)$ are

$$(\exists O, H, W)((\forall \eta \prec \xi)(\text{SCSP}[\Gamma(\eta, (O)^{\prec \eta})]((O)_\eta) \wedge \text{WF}((O)_\eta)) \wedge (\text{WF}(W) \wedge \theta_1[\Gamma(\xi, (O)^{\prec \xi})](u, v, H, W))) ; \\ (\forall O, H, W)((\forall \eta \prec \xi)(\text{SCSP}[\Gamma(\eta, (O)^{\prec \eta})]((O)_\eta) \wedge \text{WF}((O)_\eta)) \rightarrow (\text{WF}(W) \rightarrow \theta_2[\Gamma(\xi, (O)^{\prec \xi})](u, v, H, W))).$$

Let $(\forall \eta \prec \xi)(\text{SCSP}[\Gamma(\eta, (O_1)^{\prec \eta})]((O_1)_\eta) \wedge \text{WF}((O_1)_\eta))$, $\text{WF}(W_1)$ and $\theta_1[\Gamma(\xi, (O_1)^{\prec \xi})](u, v, H_1, W_1)$. To prove $\theta_2(\xi, u, v)$, take O_2, H_2, W_2 with $(\forall \eta \prec \xi)(\text{SCSP}[\Gamma(\eta, (O_2)^{\prec \eta})]((O_2)_\eta) \wedge \text{WF}((O_2)_\eta))$, $\text{WF}(W_2)$. We have to show $\theta_2[\Gamma(\xi, (O_2)^{\prec \xi})](u, v, H_2, W_2)$. Theorem 4 with $\Delta_0^0\text{-TI}$ on $\eta \preceq \xi$ shows $(O_1)^{\prec \eta} = (O_2)^{\prec \eta}$. Thus $(O_1)^{\prec \xi} = (O_2)^{\prec \xi}$, and hence what we have to show is $\theta_2[\Gamma(\xi, (O_1)^{\prec \xi})](u, v, H_2, W_2)$, which is from $\text{WF}(W_1)$, $\text{WF}(W_2)$ and $\theta_1[\Gamma(\xi, (O_1)^{\prec \xi})](u, v, H_1, W_1)$ because of Corollary 3.

Since θ_1 is $\Sigma_1^1(\mathcal{F}, \text{WF})$ and θ_2 is $\Pi_1^1(\mathcal{F}, \text{WF})$, by applying $\Pi_1^1(\mathcal{F}, \text{WF})\text{-Red}$, we have R with

$$(\forall \xi \in \alpha)(\forall u, v)[(\theta_1(\xi, u, v) \rightarrow \langle \xi, \langle u, v \rangle \rangle \in R) \wedge (\langle \xi, \langle u, v \rangle \rangle \in R \rightarrow \theta_2(\xi, u, v))]. \quad (*)$$

Now we apply the procedure of Main Lemma coordinate-wise: Let R^* be such that, for any $\xi \prec \alpha$, $(R^*)_\xi$ is the reflexive transitive closure of $(R)_\xi$ and F such that $(F)_\xi$ is the accessible part of $(R)_\xi$, i.e., $(F)_\xi = \{x \mid (R^*)_\xi \text{ is well founded on } ((R^*)_\xi)_x\}$. Define P and O accordingly. By definition, $\text{WF}((O)_\xi)$ for all $\xi \prec \alpha$. Now we prove $\text{SCSP}[\Gamma(\xi, (O)^{\prec \xi})]((O)_\xi)$ by induction on $\xi \prec \alpha$. By Main Lemma and by the construction above, it suffices to show that, for any u and v ,

$$(\exists H, W)(\text{WF}(W) \wedge \theta_1[\Gamma(\xi, (O)^{\prec \xi})](u, v, H, W)) \rightarrow \langle u, v \rangle \in (R)_\xi; \text{ and} \\ \langle u, v \rangle \in (R)_\xi \rightarrow (\forall H, W)(\text{WF}(W) \rightarrow \theta_2[\Gamma(\xi, (O)^{\prec \xi})](u, v, H, W)),$$

which are immediate from the induction hypothesis and $(*)$.

Similarly to the standard proof of $\text{BT}^2 + \Sigma_1^1\text{-TI} + \Delta_0^1\text{-CA} \vdash \Delta_0^1\text{-TR}$, Theorem 4 implies:

Proposition 15 (i) $\text{BT}^2 + \mathcal{F}\text{-SLFP} + \Sigma_1^1(\mathcal{F}, \text{WF})\text{-TI} + \Delta_0^1(\mathcal{F}, \text{WF})\text{-CA} \vdash \mathcal{F}\text{-SLFTR}$.
(ii) $\text{BT}^2 + \mathcal{F}\text{-SLFP} + \Sigma_1^1(\mathcal{F}, \text{WF})\text{-TI}_\alpha + \Delta_0^1(\mathcal{F}, \text{WF})\text{-CA} \vdash \mathcal{F}\text{-SLFTR}_\alpha$.

6.2 Hierarchy theorem for fixed point transfinite recursion

Our strategy to prove consistency is to construct a coded first order part sharing model (FOPS model, for short). This is a generalization of coded ω -model in [18, Definition VII.2.1] and was called coded LOPS model in [17, Definition 31], where the lower order correspond to the first order in the present setting.

In the rest of this section, we always assume the existence of \mathcal{Y}_1^0 as required in Definition 9. Therefore we do not know if we can apply the following results to $\mathbf{BGZ} \equiv \mathbf{NBG} - \mathbf{Repl}$.

Definition 15 (coded FOPS model) A *coded FOPS model* is second order M , viewed as encoding the \mathcal{L}^2 -model whose first and second order parts are $\{x \mid \top\}$ and $\{(M)_x \mid \top\}$ respectively. Define

$$\begin{aligned} M \models \varphi(\mathbf{x}) &\equiv \varphi(\mathbf{x}) \quad \text{if } \varphi \text{ is atomic;} \\ M \models (Qx)\varphi(x) &\equiv (Qx)(M \models \varphi(x)); \end{aligned} \quad \begin{aligned} M \models \varphi \square \psi &\equiv (M \models \varphi) \square (M \models \psi) \quad \text{for } \square \equiv \wedge, \vee, \rightarrow; \\ M \models (QX)\varphi(X) &\equiv (Qx)(M \models \varphi((M)_x)) \quad \text{for } Q \equiv \forall, \exists. \end{aligned}$$

Proposition 16 $\mathbf{BT}^2 + \Delta_0^1\text{-LFTTR}_{\omega^{\alpha+1}} \vdash \exists M (M \models \mathbf{BT}^2 + \Delta_0^1\text{-LFTTR}_{\omega^\alpha})$.

Proof Let $\text{LFHier}[\tilde{T}](F, \omega^{\alpha+1})$, i.e., for any $\gamma \prec \omega^{\alpha+1}$, $(F)_\gamma$ is a least fixed point of $\tilde{T}((F)^{\prec\gamma})$. Define

$$M = \{\langle \langle \xi, e, y \rangle, z \rangle \mid \Upsilon_1^0(e, \langle z, y \rangle, \emptyset, (F)_\xi), e \in \omega, \xi \prec \omega^{\alpha+1}\}.$$

The only non- Π_1^1 axiom of \mathbf{BT}^2 is $\Delta_0^1\text{-CA}$ (where **Choice** is treated by a new predicate for well-ordering), or equivalently, $\Sigma_1^0\text{-CA}$, which is an instance of $\Sigma_1^0\text{-FP}$, i.e., $\Sigma_1^0\text{-LFP}_{\omega^0}$. Thus we have to show $M \models (\Sigma_1^0 \wedge \Pi_1^0)\text{-LFP}_{\omega^\alpha}$, i.e., for a $(\Sigma_1^0 \wedge \Pi_1^0)$ operator of arity $(1, 1, 1)$ with parameters

$$\Gamma = \lambda y. \lambda X, Y. \{z \mid \varphi(z, y_0, y, X, Y, (M)_{\langle \gamma_1, e_1, y_1 \rangle}, \dots, (M)_{\langle \gamma_k, e_k, y_k \rangle})\}$$

where $\gamma_1, \dots, \gamma_k \prec \omega^{\alpha+1}$, has the hierarchy of least fixed points along ω^α of the form $(M)_{\langle \delta, e, u \rangle}$ with $\delta \prec \omega^{\alpha+1}$, since being a least fixed point of a Δ_0^1 operator is downward persistent for coded FOPS models. Take $\gamma \prec \omega^{\alpha+1}$ so that $\gamma_1, \dots, \gamma_k \prec \gamma$. Then $(M)_{\langle \gamma_i, e_i, y_i \rangle}$'s are Δ_0^0 definable from y_i 's, γ_i 's and $(F)^{\prec\gamma}$. Let $f(\xi) = \langle \mathbf{y}, \gamma, \xi \rangle$ for $\xi \prec \omega^\alpha$. By diagonalization, there is $e \in \omega$ such that, for any $\xi \prec \omega^\alpha$ and z ,

$$\tilde{T}(e, z, f(\xi), X, (F)^{\prec\gamma+\xi}) \leftrightarrow \varphi(z, y_0, \xi, (X)_{\langle e, f(\xi) \rangle}, \{\langle \eta, x \rangle \mid x \in ((F)_{\gamma+\eta})_{\langle e, f(\eta) \rangle}, \eta \prec \xi\}, (M)_{\langle \gamma_1, e_1, y_1 \rangle}, \dots).$$

$((F)_{\gamma+\xi})_{\langle e, f(\xi) \rangle}$ is a least fixed point of $\lambda X. \{z \mid \varphi(z, y_0, \xi, X, \{\langle \eta, x \rangle \mid x \in ((F)_{\gamma+\eta})_{\langle e, f(\eta) \rangle}, \eta \prec \xi\}, M_{\langle \gamma, e, \mathbf{y} \rangle})\}$, since $(F)_{\gamma+\xi}$ is a least fixed point of $\tilde{T}((F)^{\prec\gamma+\xi})$. Now, $\{\langle \xi, x \rangle \mid x \in ((F)_{\gamma+\xi})_{\langle e, f(\xi) \rangle}, \xi \prec \omega^\alpha\} = (M)_{\langle \gamma+\omega^\alpha, e', f \rangle}$ for some $e' \in \omega$. This is what we require, where $\gamma + \omega^\alpha \prec \omega^{\alpha+1}$.

Remark 8 If we are dealing with the first order theories, this proof shows, in terms of Section 8,

$$ID_{\omega^{\alpha+1}} \vdash \{\langle \langle \xi, e, y \rangle, x \rangle \mid \Upsilon_1^0(e, \langle x, y \rangle, \emptyset, (L_{\tilde{T}})_\xi), e \in \omega, \xi \prec \omega^{\alpha+1}\} \models \mathbf{BT}^2 + \Delta_0^1\text{-LFTTR}_{\omega^\alpha}.$$

With the next lemma we can conclude that the implication from $\Delta_0^1\text{-LFTTR}_{\omega^{\alpha+1}}$ to $\Delta_0^1\text{-LFTTR}_{\omega^\alpha}$ cannot be reversed, since $\Delta_0^1\text{-LFTTR}_{\omega^{\alpha+1}}$ implies $\Delta_0^1\text{-TR}_\omega$ (if $\mathbf{BT}^2 + \Delta_0^1\text{-LFTTR}_{\omega^\alpha}$ is consistent).

Lemma 17 For an \mathcal{L}^2 -theory T , $\mathbf{BT}^2 + \Delta_0^1\text{-TR}_\omega \vdash (\exists M)(M \models T) \rightarrow \text{Con}(T)$.

Proof Since only first order quantifiers are involved in the clauses of the recursive definition of $M \models \varphi$ given above, $\Delta_0^1\text{-TR}_\omega$ allows us to define the satisfaction relations S so that $\langle \Gamma \varphi \sqsupseteq, \mathbf{a} \rangle \in S \leftrightarrow M \models \varphi(\mathbf{a})$. Now, by induction on derivation, we can prove that if $T \vdash \varphi$ then $\langle \Gamma \varphi \sqsupseteq, \mathbf{x} \rangle \in S$ for any \mathbf{x} . Thus $T \vdash \perp$ implies $M \models \perp$, a contradiction.

Example 10 As noted in [18], when $\mathbf{BT}^2 = \mathbf{ACA}_0$, any coded FOPS model (or coded ω -model) satisfies full induction scheme $\mathcal{L}_N^2\text{-Ind}$. Similarly, when $\mathbf{BT}^2 = \mathbf{NGB}$, a coded FOPS model satisfies full transfinite induction $\mathcal{L}_S^2\text{-TI}$, full set separation $\mathcal{L}_S^2\text{-sSep}$ and full set collection $\mathcal{L}_S^2\text{-sColl}$, where

$$\mathcal{F}\text{-TI } \text{WF}(W) \rightarrow \text{TI}[\varphi](W) \text{ for any } \mathcal{F}\text{-formula } \varphi; \quad \mathcal{F}\text{-TI}_\alpha \text{ TI}[\varphi](\alpha) \text{ for any } \mathcal{F}\text{-formula } \varphi;$$

$$\mathcal{F}\text{-sSep } (\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow z \in x \wedge \varphi) \text{ for any } \mathcal{F}\text{-formula } \varphi \text{ in which } y \text{ is not free;}$$

$$\mathcal{F}\text{-sColl } (\forall y \in x)(\exists z)\varphi \rightarrow (\exists u)(\forall y \in x)(\exists z \in u)\varphi \text{ for any } \mathcal{F}\text{-formula } \varphi \text{ in which } u \text{ is not free.}$$

Corollary 11 Let \mathbf{Sch} be any scheme which is satisfied by any coded FOPS model. Then $\mathbf{BT}^2 + \Delta_0^1\text{-LFTTR}_{\omega^{\alpha+1}} \vdash \text{Con}(\mathbf{BT}^2 + \Delta_0^1\text{-LFTTR}_{\omega^\alpha} + \mathcal{L}^2\text{-Sch})$.

By combining the results from the last subsection, we have:

Corollary 12 Let \mathbf{Sch} be any scheme which is satisfied by any coded FOPS model. Then

$$(1) \mathbf{BT}^2 + \text{WF}(\alpha) + \Pi_1^1(\text{WF})\text{-Red} \vdash \text{Con}(\mathbf{BT}^2 + \Delta_0^1\text{-SLFTTR}_\alpha + \mathcal{L}^2\text{-Sch}).$$

$$(2) \mathbf{BT}^2 + \text{WF}(\alpha) + \Delta_0^1\text{-LFTTR}_\alpha + \mathcal{L}^2\text{-Sch} \not\vdash \Delta_1^1(\text{WF})\text{-CA} \text{ (under enough consistency).}$$

Proof (1) is by $\text{WF}(\alpha) \rightarrow \text{WF}(\alpha+1)$. (2) We may assume that **Sch** contains $\mathbf{TI}_{\omega^{\beta+1}}$ for some β with $\omega^\beta \succeq \alpha$. If the derivability were true, then by Proposition 15, $\mathbf{BT}^2 + \Delta_0^1\text{-SLFTR}_\alpha + \Sigma_1^1(\text{WF})\text{-TI}_{\omega^{\beta+1}} + \mathcal{L}^2\text{-Sch}$ (which we assume to be consistent) yields a FOPS model of itself and hence its own consistency.

As an instance, we have $\mathbf{NGB} + \Delta_0^1\text{-FTR}_\alpha + \mathcal{L}_S^2\text{-TI}, \text{-sSep}, \text{-sColl} \not\vdash \Delta_1^1\text{-CA}$ (and hence $\not\vdash \Sigma_1^1\text{-Coll}$, by Lemma 1), which contrasts with the results in number theory (see [1], [18, Theorem V.8.3]): Over **ACA**₀, the weakest $\Delta_0^1\text{-FP}$ is equivalent to $\Delta_0^1\text{-TR}$ and hence implies $\Sigma_1^1\text{-Coll}$ and $\Delta_1^1\text{-CA}$.

Remark 9 The natural question here is whether the contrast on the implication of $\Delta_1^1\text{-CA}$ from $\Delta_0^1\text{-FP}$ survives in the sense of consistency-wise implication. Namely, do $\mathbf{BT}^2 + \Delta_0^1\text{-FP}$ and $\mathbf{BT}^2 + \Delta_0^1\text{-FP} + \Delta_1^1\text{-CA}$ have the same proof-theoretic strength? The answer is yes. Either by the method of recursively saturated model (see [18, IX.4]; for the modification required for $\Delta_0^1\text{-FP}$, see [17, Remark 38]) or by the so-called asymmetric interpretation (for the modification required for $\Delta_0^1\text{-FP}$, see [8]) we can prove the Π_2^1 conservation of $\mathbf{BT}^2 + \Delta_0^1\text{-FP} + \Sigma_1^1\text{-Coll}$ over $\mathbf{BT}^2 + \Delta_0^1\text{-FP}$. Here $\Delta_0^1\text{-FP}$ can be replaced by $\Delta_0^1\text{-FTR}_\alpha$ and, in the presence of $\Pi_2^0\text{-Ref}$ (or the Δ_0^1 -ness of WF), by $\Delta_0^1\text{-FTR}$, $\Delta_0^1\text{-LFTR}$ and $\Delta_0^1\text{-SLFTR}$, as well as $\Delta_0^1\text{-LFTR}_\alpha$ and $\Delta_0^1\text{-SLFTR}_\alpha$.

Thus the implication of $\Sigma_1^1\text{-Coll}$ or of $\Delta_1^1\text{-CA}$ is among dissimilarities in the sense of logical implication whereas it is among similarities in the sense of consistency-wise implication.

6.3 Dependent transfinite recursion

Having the results in the last subsections, especially Proposition 14, we naturally wonder if $\Delta_0^1\text{-LFTR}$ is equivalent to $\Pi_1^1\text{-Red}$. We can separate them, by the same method as the author's previous work [17] which separated $\Delta_0^1\text{-TR}$ and $\Delta_0^1\text{-FP}$. The method was a new kind of way of iteration of a given construction (e.g., elementary comprehension), called *dependent transfinite recursion*, which can be seen as a transfinite recursion of transfinite recursion itself. As pointed out there [17], this can be applied not only to the comprehension constructions but also to other constructions, like fixed points and least fixed points. Besides the natural philosophical motivation for this new kind of iteration, this actually helps us to show, in set theory, that $\Pi_1^1\text{-Red}$ is strictly stronger than $\Delta_0^1\text{-LFTR}$ and that $\Delta_0^1\text{-LFTR}$ is not capable of implying $\Delta_1^1\text{-CA}$.

Definition 16 The *fixed point 2-fold dependent transfinite recursion* is the following schema:

$\mathcal{F}\text{-FTR}^2$ $\text{WF}(\succ^{(1)}) \wedge (\forall w, Y)\text{WF}(\succ_{Y \upharpoonright w}^{(0),w}) \rightarrow (\exists H)\text{FHier}[\Gamma](H, \oplus_{w \in \text{fd}(\succ^{(1)})} \succ_{H \upharpoonright w}^{(0),w}),$
 for any binary relation $u \succ^{(1)} v$ and $u \succ_{Y \upharpoonright w}^{(0),w} v$, defined by \mathcal{F} -formulae $\theta^1(u, v)$ and $\theta^0(u, v, w, Y)$ with parameters respectively, and for any \mathcal{F} operator Γ of arity $(1, 1, 1)$,

where $Y \upharpoonright w$ denotes $\{\langle \langle w', u \rangle, x \rangle \in Y \mid w \succ^{(1)} w'\}$ and where

$$\langle w', u \rangle (\oplus_{w \in \text{fd}(\succ^{(1)})} \succ_{H \upharpoonright w}^{(0),w}) \langle w'', v \rangle \leftrightarrow (w' \succ^{(1)} w'') \vee (w' = w'' \wedge u \succ_{H \upharpoonright w'}^{(0),w'} v).$$

The order $\succ^{(1)}$ is called the *preceding order*.

$\mathcal{F}\text{-LFTR}^2$, *least fixed point 2-fold dependent transfinite recursion*, and $\mathcal{F}\text{-SLFTR}^2$, *strong least fixed point 2-fold dependent transfinite recursion*, are defined in the same way with $\text{FHier}[-](-, -)$ replaced by $\text{LFHier}[-](-, -)$ and by $\text{SLFHier}[-](-, -)$ respectively.

Here the well-founded relation $\oplus_{w \in \text{fd}(\succ^{(1)})} \succ_{H \upharpoonright w}^{(0),w}$ depends on the resulting hierarchy H in the following manner: If $*$ is the minimal element of $\succ^{(1)}$, since $H \upharpoonright * = \emptyset$, the well-founded relation $\succ_{H \upharpoonright *}^{(0),*}$ is fixed at first, not depending on H . Then, by non-dependent (least) fixed point transfinite recursion along $\succ_{H \upharpoonright *}^{(0),*}$, we have $(H)_{\langle *, x \rangle}$'s. If we assume that we have a family of $(H)_{\langle w', x \rangle}$'s for all $w' \prec^{(1)} w$ (as induction hypothesis), then, since $\succ_{H \upharpoonright w}^{(0),w}$ is now fixed, we can apply non-dependent (least) fixed point transfinite recursion along it to obtain $(H)_{\langle w, x \rangle}$'s. The schemata formalize such informal construction, and this is why we call such a scheme *dependent transfinite recursion*.

Thus, if the iterant construction is deterministic (like comprehension and least fixed point, as opposed to fixed point, which is not unique), the resulting hierarchy is unique:

Lemma 18 In the same syntactic situation as in the definition above, if $\text{LFHier}[\Gamma](H, \oplus_{w \in \text{fd}(\succ^{(1)})} \succ_{H \upharpoonright w}^{(0), w})$ and if $\text{LFHier}[\Gamma](G, \oplus_{w \in \text{fd}(\succ^{(1)})} \succ_{G \upharpoonright w}^{(0), w})$ then $G = H$, provided that $\text{WF}(\succ^{(1)}) \wedge (\forall w, Y) \text{WF}(\succ_{Y \upharpoonright w}^{(0), w})$.

We can further generalize it by allowing $\succ^{(1)}$ to depend on H . Iterating this generalization,

Definition 17 We define the following schemata for $k \in \omega$:

$$\mathcal{F}\text{-}\mathbf{FTR}^{k+1} (\forall w_k, \dots, w_0, Y) \wedge_{j \leq k} \text{WF}(\succ_{Y \upharpoonright \langle w_k, \dots, w_{j+1} \rangle}^{(j), w_k, \dots, w_{j+1}}) \rightarrow (\exists H) \text{FHier}[\varphi](H, \Pi_{j \leq k} \succ_H^{(j)});$$

for any \mathcal{F} -formulae φ and $\succ_Y^{(j), w^k, \dots, w^{j+1}}$ (for $j < k$), where

$$\langle w_k, \dots, w_0 \rangle (\Pi_{j \leq k} \succ_H^{(j)}) \langle v_k, \dots, v_0 \rangle \text{ iff } (w_k \succ^{(k)} v_k) \vee \dots \vee (w_k = v_k \wedge \dots \wedge w_1 = v_1 \wedge w^0 \succ_{H \upharpoonright \langle w_n, \dots, w_1 \rangle}^{(0), w_k, \dots, w_1} v_0).$$

$\mathcal{F}\text{-}\mathbf{LFTR}^{k+1}$ and $\mathcal{F}\text{-}\mathbf{SLFTR}^{k+1}$ are defined accordingly. Moreover, $\mathcal{F}\text{-}\mathbf{FTR}_\alpha^{k+1}$, $\mathcal{F}\text{-}\mathbf{LFTR}_\alpha^{k+1}$ and $\mathcal{F}\text{-}\mathbf{SLFTR}_\alpha^{k+1}$ denote the schemata with the most preceding order $\succ^{(k)}$ restricted to α .

As also pointed out in the previous work, we can further define $\mathcal{F}\text{-}\mathbf{FTR}^\omega$, $\mathcal{F}\text{-}\mathbf{LFTR}^\omega$ and $\mathcal{F}\text{-}\mathbf{SLFTR}^\omega$, by considering ω -sequences f such that, for all but finite $k \in \omega$, $f(k)$ is minimal in $\succ_X^{f \upharpoonright (\omega \setminus (k+1))}$ (like Veblen hierarchy), and we can replace this ω by any well-order, which, again, depends on the intermediate stages of resulting hierarchy H , and so on.

Our previously obtained results can easily be generalized to such dependent transfinite recursions as follows, where for (iii) we need the same trick as [17, Theorem 33], i.e., we collect all the well-founded relations defined from previous parameters into a disjoint union, by a Δ_0^1 formula.

- (i) $\mathbf{BT}^2 \vdash \Delta_0^1\text{-}\mathbf{SLFTR}^{k+1} \leftrightarrow \Delta_0^1\text{-}\mathbf{LFTR}^{k+1}; \mathbf{BT}^2 + \Pi_2^0\text{-}\mathbf{Ref} \vdash \Delta_0^1\text{-}\mathbf{LFTR}^{k+1} \leftrightarrow \Delta_0^1\text{-}\mathbf{FTR}^{k+1}$.
- (ii) $\mathbf{BT}^2 \vdash \Delta_0^1\text{-}\mathbf{SLFTR}_\alpha^{k+1} \leftrightarrow \Delta_0^1\text{-}\mathbf{LFTR}_\alpha^{k+1}; \mathbf{BT}^2 + \Pi_2^0\text{-}\mathbf{Ref} \vdash \Delta_0^1\text{-}\mathbf{LFTR}_\alpha^{k+1} \leftrightarrow \Delta_0^1\text{-}\mathbf{FTR}_\alpha^{k+1}$.
- (iii) $\mathbf{BT}^2 + \Pi_2^0\text{-}\mathbf{Ref} + \Delta_0^1\text{-}\mathbf{LFTR}_{\omega^{\alpha+1}}^{k+1} \vdash \text{Con}(\mathbf{BT}^2 + \Delta_0^1\text{-}\mathbf{LFTR}_{\omega^\alpha}^{k+1} + \mathcal{L}^2\text{-}\mathbf{Sch})$.
- (iv) $\mathbf{BT}^2 + \Pi_1^1(\text{WF})\text{-}\mathbf{Red} \vdash \Delta_0^1\text{-}\mathbf{SLFTR}^{k+1}; \mathbf{BT}^2 + \Delta_0^1\text{-}\mathbf{SLFP} + \Sigma_1^1(\text{WF})\text{-}\mathbf{TI} + \Delta_1^1(\text{WF})\text{-}\mathbf{CA} \vdash \Delta_0^1\text{-}\mathbf{SLFTR}^{k+1}$.
- (v) $\mathbf{BT}^2 + \Delta_0^1\text{-}\mathbf{LFTR}_\alpha^{k+1} + \mathcal{L}^2\text{-}\mathbf{Sch} \not\vdash \Delta_1^1(\text{WF})\text{-}\mathbf{CA}$ (and so $\not\vdash \Sigma_1^1(\text{WF})\text{-}\mathbf{Coll}$, $\not\vdash \Pi_1^1(\text{WF})\text{-}\mathbf{Red}$), if $\mathbf{BT}^2 \vdash \Pi_2^0\text{-}\mathbf{Ref}$ (and under enough consistency), for any scheme \mathbf{Sch} which is satisfied by any coded FOPS model

In particular, we can conclude, in set theory, that $\Pi_1^1\text{-}\mathbf{Red}$ is strictly stronger than $\Delta_0^1\text{-}\mathbf{LFTR}$ by (iii) and (iv), and that $\Delta_0^1\text{-}\mathbf{LFTR}$ is not capable of implying $\Delta_1^1\text{-}\mathbf{CA}$ by (v) with $k = \alpha = 1$.

There is a result which might support the view that the dependent transfinite recursion is the iteration of transfinite recursion: $\mathbf{BT}^2 + \Delta_0^1\text{-}\mathbf{TR}^n$ plus $\Delta_0^1\text{-}\mathbf{DC}_\Omega$ (Δ_0^1 dependent choice along the global well-ordering whose length is denoted by Ω) is Π_2^1 -conservative over $\mathbf{BT}^2 + \Delta_0^1\text{-}\mathbf{TR}_{<\Omega^\omega}^{n+1}$; and $\mathbf{BT}^2 + \Delta_0^1\text{-}\mathbf{FTR}^n$ plus $\Delta_0^1\text{-}\mathbf{DC}_\Omega$ is Π_2^1 -conservative over $\mathbf{BT}^2 + \Delta_0^1\text{-}\mathbf{FTR}_{<\Omega^\omega}^{n+1}$, which shall be included in the author's future work. Here Ω does not need to be of the order type Ord. These (with $n = 0$) should be compared with (and seen as extensions of) the following well-known results in number theory (e.g., [8]): $\mathbf{ACA}_0 + \Delta_0^1\text{-}\mathbf{DC}$ is Π_2^1 -conservative over $\mathbf{ACA}_0 + \Delta_0^1\text{-}\mathbf{TR}_{<\omega^\omega}$; and $\mathbf{ACA}_0 + \Delta_0^1\text{-}\mathbf{FP} + \Delta_0^1\text{-}\mathbf{DC}$ is Π_2^1 -conservative over $\mathbf{ACA}_0 + \Delta_0^1\text{-}\mathbf{FTR}_{<\omega^\omega}$.

7 Conservation Results for Some Axioms in NBG

All the results in second order set theory whose equivalents in second order number theory do not hold, that we saw so far, are based on $\Pi_2^0\text{-}\mathbf{Ref}$. However, the standard proof of it (or more generally $\Delta_0^1\text{-}\mathbf{Ref}$) heavily depends on **Found** in **NBG**, the status of which as an axiom of set theory has been somehow controversial. Though, even in the absence of it, Skolemization argument allows us to prove $\Delta_0^1\text{-}\mathbf{Ref}$, if we have **Choice**, the status of which has also been controversial. Then the natural question is: What happens if we take, as base theory, **NBG**− with at most the local choice **LC** (or, equivalently, the axiom of choice in the first order formulation)?

We can conclude that the statuses of these axioms do not matter for the investigations of consistency strengths or proof-theoretic strengths of the schematic extensions of \mathbf{BT}^2 (and of corresponding first order systems ID_α or \widehat{ID}_α). More precisely, we can prove the equiconsistency between the most systems with and without these axioms. Let us start with a preparation.

Definition 18 (relativization) Z is called *transitive* (or $\text{Trans}(Z)$) if $(\forall x \in Z)(\forall y \in x)(y \in Z)$. For a formula φ and an operator Γ , φ^Z and Γ^Z are defined as follows:

$$\begin{aligned}\psi(\mathbf{y}, \mathbf{X})^Z &\equiv \psi(\mathbf{y}, \mathbf{X}) \text{ if } \psi \text{ is atomic; } (\psi_1 \square \psi_2)^Z \equiv (\psi_1)^Z \square (\psi_2)^Z \text{ for } \square \equiv \wedge, \vee, \rightarrow; \\ ((Qy)\psi)^Z &\equiv (Qy \in Z)(\psi^Z); \quad ((QX)\psi)^Z \equiv (QX)(\psi^Z); \\ (\lambda \mathbf{y}.\lambda \mathbf{X}. \{z \mid \varphi(z, \mathbf{y}, \mathbf{X})\})^Z &= \lambda \mathbf{y}.\lambda \mathbf{X}. \{z \mid z \in Z \wedge \varphi^Z(z, , \mathbf{y}, \mathbf{X})\}.\end{aligned}$$

Remark 10 (1) $(X = Y)^Z$ is $(\forall x \in Z)(z \in X \leftrightarrow x \in Y)$ and hence does not imply $X = Y$.

(2) Up to now we did not consider the field $\text{fd}(W) = \{x \mid (\exists y)(\langle y, x \rangle \in W \vee \langle x, y \rangle \in W)\}$ by treating $z \notin \text{fd}(W)$ as a minimal element, i.e., $(W)_z = \emptyset$. In this section, however, it is more convenient to restrict the quantifiers in $\text{SCSP}[\Gamma]$ (w in Definition 11), and in $\text{Hier}[\Gamma]$, $\text{FHier}[\Gamma]$, $\text{LFHier}[\Gamma]$ and $\text{SLFHier}[\Gamma]$ (i.e., x in Definitions 8 and 12), to the fields (of A or W). It is easy to see that these modifications preserve the axiom schemata in the sense of provable equivalence.

Lemma 19 (absoluteness) Recall $\text{Wf}(W)$ from Remark 5. Let θ state the Σ_1 -ness of Wf and the decreasingness of Wf (i.e., $X \subset Y \wedge \text{Wf}(Y) \rightarrow \text{Wf}(X)$) (both of which hold in \mathbf{ZF}).

NBG⁻ proves the following: If $\text{Trans}(Z)$ and Z is closed under Mostowski pairs, then

- (i) $\text{Hier}[\Gamma^Z](H \cap Z, W) \leftrightarrow (\text{Hier}[\Gamma](H, W))^Z$ for any $W \subset Z$;
- (ii) $\text{FHier}[\Gamma^Z](H \cap Z, W) \leftrightarrow (\text{FHier}[\Gamma](H, W))^Z$ for any $W \subset Z$;
- (iii) $\text{LFHier}[\Gamma^Z](H \cap Z, W) \leftrightarrow (\text{LFHier}[\Gamma](H, W))^Z$ for any $W \subset Z$;
- (iv) $\text{SCSP}[\Gamma^Z](W) \leftrightarrow (\text{SCSP}[\Gamma](W))^Z$ for any $W \subset Z$;
- (v) $\text{SLFHier}[\Gamma^Z](H \cap Z, W) \leftrightarrow (\text{SLFHier}[\Gamma](H, W))^Z$ for any $W \subset Z$;
- (vi) $\text{Wf}(W) \leftrightarrow \text{Wf}^Z(W)$, if (a) $W, \text{Ord} \subset Z$, (b) $(\forall z \in Z)(z \cap W \in Z)$, (c) $Z = \bigcup_{\alpha \in \text{Ord}} (V_\alpha)^Z$ (where Ord is not relativized to Z) (d) $\Delta_0^1\text{-Ref}$ and (e) θ^Z .

Proof (i)-(v) are from Remark 10 (all the first-order quantifiers are bounded by $\text{fd}(W) \subset Z$), since second order quantifiers are not changed by the relativizations and since Γ is \mathcal{F} then so is $\Gamma^{\mathfrak{W}}$.

(vi) Assume (a)-(e). $\text{Wf}(W) \rightarrow \text{Wf}^Z(W)$ is easy. Suppose, for contradiction, $\neg \text{Wf}(W)$ and $\text{Wf}^Z(W)$. (d) yields transitive a with $(\neg \text{Wf})^a(W)$, i.e., $(\neg \text{Wf})^a(W \cap a)$. By upward persistency of Σ_1 formulae, we have $\neg \text{Wf}(W \cap a)$. $W \cap a \subset (V_\alpha)^Z$ for some $\alpha \in \text{Ord}$ by (c). $\text{Wf}^Z(W)$ implies $\text{Wf}^Z(W \cap (V_\alpha)^Z)$. By (b) we have $W \cap (V_\alpha)^Z \in Z$, and so (e) implies $\text{Wf}(W \cap (V_\alpha)^Z)$, contradicting $(\neg \text{Wf})^a(W \cap a)$ and $W \cap a \subset (V_\alpha)^Z$.

We first treat **Found**. The main idea is to restrict the domain to what is called *well-founded part* $\mathfrak{W} = \{x \mid \text{Wf}(\in \upharpoonright \text{trh}(x))\}$, where $\text{trh}(x)$ is the *transitive hull* of x , the least transitive superset of x .

Proposition 20 Let \mathcal{F} be Δ_0^1 , Σ_n^1 or Π_n^1 . Let T be **NBG⁻** plus any combination of

- **LC**, **Choice**, $\Delta(\mathcal{F})\text{-CA}$, $\mathcal{F}\text{-Red}$, $\mathcal{F}\text{-TR}$, $\mathcal{F}\text{-FTR}$, $\mathcal{F}\text{-LFTR}$, $\mathcal{F}\text{-SLFTR}$, and
 - $\mathcal{F}\text{-TR}_\alpha$, $\mathcal{F}\text{-FTR}_\alpha$, $\mathcal{F}\text{-LFTR}_\alpha$, $\mathcal{F}\text{-SLFTR}_\alpha$
- where the order \succ for α is Δ_0^0 -defined with $\mathbf{ZF}^- \vdash \text{WF}(\succ) \wedge (\succ \subset \mathfrak{W})$.

Then $T + \text{Found}$ can be interpreted into T by relativizing (the first order part) to \mathfrak{W} .

Proof We are working in **NBG⁻**. $(\text{NBG}^-)^{\mathfrak{W}}$, $\text{LC} \rightarrow (\text{LC})^{\mathfrak{W}}$ and $\text{Choice} \rightarrow (\text{Choice})^{\mathfrak{W}}$ are obvious. To see $(\text{Found})^{\mathfrak{W}}$, let $(\forall x \in \mathfrak{W})(\forall y \in x)(y \in X) \rightarrow x \in X$. For $z \in \mathfrak{W}$, since $\{z\} \in \mathfrak{W}$, we have $\text{TI}[\{x \mid x \in \mathfrak{W} \rightarrow x \in X\}] \in \upharpoonright \text{trh}(\{z\})$ (cf. Remark 5) which implies $z \in X$.

To show $\mathcal{F}\text{-LFTR} \rightarrow (\mathcal{F}\text{-LFTR})^{\mathfrak{W}}$, assume $\text{WF}(W)^{\mathfrak{W}}$. Let $W' = W \cap \mathfrak{W}$. Then, by definition, $\text{WF}(W')$. For an \mathcal{F} operator Γ , $\mathcal{F}\text{-LFTR}$ yields H with $\text{LFHier}[\Gamma^{\mathfrak{W}}](H, W')$. $(\text{LFHier}[\Gamma](H, W))^{\mathfrak{W}}$ follows from Lemma 19. We can similarly treat $\mathcal{F}\text{-SLFTR}$, $\mathcal{F}\text{-FTR}$, $\mathcal{F}\text{-TR}$, and α -restricted versions of them.

It is straightforward to treat $\Delta(\mathcal{F})\text{-CA}$. For $\mathcal{F}\text{-Red}$, if $((\forall x)(\varphi(x) \vee \psi(x)))^{\mathfrak{W}}$, we can apply $\mathcal{F}\text{-Red}$ to $(\forall x)[(x \in \mathfrak{W} \rightarrow \varphi^{\mathfrak{W}}(x)) \vee (x \in \mathfrak{W} \rightarrow \psi^{\mathfrak{W}}(x))]$ to get desired Z .

It seems impossible to deal with $\mathcal{F} = \Sigma_n^0$ or Π_n^0 , since the relativization to \mathfrak{W} does not preserve them. With $\mathcal{F}\text{-Coll}$, it also seems impossible to deal in this way, but possible, by a detour via Remark 9, which is free from **Found**: E.g., $\text{NBG}^- + \Delta_0^1\text{-LFTR} + \Sigma_1^1\text{-Coll}$ contains $\text{NBG}^- + \Delta_0^1\text{-FTR}$ and so interprets, by \mathfrak{W} , $\text{NBG} + \Delta_0^1\text{-FTR}$ which is, by Remark 9, Π_2^1 -equivalent to $\text{NBG} + \Delta_0^1\text{-FTR} + \Sigma_1^1\text{-Coll}$ and, by our main result, to $\text{NBG} + \Delta_0^1\text{-SLFTR} + \Sigma_1^1\text{-Coll}$.

Theorem 14 Let T be **NBG**⁻ plus any choice of **LC**, **Choice**, Σ_1^1 -**Coll**, Δ_0^1 -**TR**, Δ_0^1 -**FTR**, Δ_0^1 -**LFR**, Δ_0^1 -**SLFTR**, Δ_0^1 -**TR** $_\alpha$, Δ_0^1 -**FTR** $_\alpha$, Δ_0^1 -**LFR** $_\alpha$ and Δ_0^1 -**SLFTR** $_\alpha$ (with the same condition on α as above). Then T and $T + \mathbf{Found}$ are equiconsistent.

Next we treat the other controversial axioms, **LC** and **Choice**, by the notion of constructibility. Let L_α be the α -th level in the constructable hierarchy. (Though L_α is a set, it can be denoted by a capital letter, for it is identified with $\{x \mid x \in L_\alpha\}$.) Let us restrict the first order domain to $\mathfrak{L} = \bigcup_{\alpha \in \text{Ord}} L_\alpha$. In this case we have to restrict also the second order, for otherwise **Sep** does not hold: For $x \in \mathfrak{L}$, nothing guarantees $x \cap X \in \mathfrak{L}$. How should we do this?

Should we extend the notion of constructibility to classes, in the same way as Simpson [18, VII.3-4] did for sets in \mathcal{L}_N^2 ? In the “standard” model $(V_\kappa, V_{\kappa+1})$ for strongly inaccessible κ , this works well: \mathfrak{L} and the interpretation of constructibility predicate for the classes are L_κ and $L \cap V_{\kappa+1}$ respectively, and famous *Condensation Lemma* tells us $L \cap V_\kappa = L_\kappa$. However, the lemma does not seem to hold in “non-standard” models. We need another way to restrict the second order domain, as below. Note that by the previous result we can freely use **Found**, which makes the three variants of fixed point construction be equivalent.

Proposition 21 Let T be **NBG** plus any combinations of Δ_0^1 -**TR**, Δ_0^1 -**SLFTR**, Δ_0^1 -**TR** $_\alpha$, Δ_0^1 -**SLFTR** $_\alpha$ where α is as in Proposition 20. Then $T + \mathbf{Choice}$ is interpretable into T .

Proof The interpretation \mathfrak{l} defined below obviously interprets **Sep**, i.e., $(\mathbf{Sep})^\mathfrak{l}$.

$$(Qx\varphi)^\mathfrak{l} \equiv (Qx \in \mathfrak{L})\varphi^\mathfrak{l} \text{ and } (QX\varphi)^\mathfrak{l} \equiv (QX \in \mathfrak{L}^2)\varphi^\mathfrak{l} \text{ where } \mathfrak{L}^2 = \{X \mid (\forall z \in \mathfrak{L})(z \cap X \in \mathfrak{L})\}.$$

To see $(\mathbf{Repl})^\mathfrak{l}$, let $((\forall x \in a)(\exists y)(\langle x, y \rangle \in X))^\mathfrak{l}$, i.e., $(\forall x \in a)(\exists \xi)(\exists y \in L_\xi)(\langle x, y \rangle \in X)$. **Repl** yields $\beta \in \text{Ord}$ with $(\forall x \in a)(\exists \xi < \beta)(\exists y \in L_\xi)(\langle x, y \rangle \in X)$. So $((\forall x \in a)(\exists y \in L_\beta)(\langle x, y \rangle \in X))^\mathfrak{l}$ with $L_\beta \in \mathfrak{L}$.

The proof of $(\Delta_0^1\text{-CA})^\mathfrak{l}$ (as well as $(\Delta_0^1\text{-TR})^\mathfrak{l}$ and $(\Delta_0^1\text{-SLFTR})^\mathfrak{l}$) is included in the proof of $(\Delta_0^1\text{-TR})^\mathfrak{l}$ (and $(\Delta_0^1\text{-SLFTR})^\mathfrak{l}$) below, but not depending on (\dagger) for which we need $(\mathbf{N BG})^\mathfrak{l}$. Thus without circularity we have $(\mathbf{N BG})^\mathfrak{l}$.

Note that, by $(\mathbf{N BG})^\mathfrak{l}$, definition of $(\cdot)^\mathfrak{l}$, Lemma 19 (vi) and **N BG** respectively, for $X \in \mathfrak{L}^2$,

$$(\text{WF}(X))^\mathfrak{l} \leftrightarrow (\text{Wf}(X))^\mathfrak{l} \leftrightarrow (\text{Wf}(X))^\mathfrak{L} \leftrightarrow \text{Wf}(X) \leftrightarrow \text{WF}(X). \quad (\dagger)$$

We prove $(\Delta_0^1\text{-TR})^\mathfrak{l}$ and $(\Delta_0^1\text{-SLFTR})^\mathfrak{l}$. Let $(\text{WF}(W))^\mathfrak{l}$ with $W \in \mathfrak{L}^2$. We may assume $W \subset \mathfrak{L}$. By (\dagger) , $\Delta_0^1\text{-TR}$ or $\Delta_0^1\text{-SLFTR}$ yields H with $\text{Hier}[\Gamma^\mathfrak{L}](H, W)$ or $\text{SLFHier}[\Gamma^\mathfrak{L}](H, W)$. By Lemma 19 (i)(v), it remains to show $H \in \mathfrak{L}^2$. Take $z \in \mathfrak{L}$. We have to show $H \cap z \in \mathfrak{L}$.

Since we have replacement for Δ_0^1 with any parameters (not necessarily in \mathfrak{L}^2), the standard proof of reflection within \mathfrak{L} yields $a = (V_\alpha)^\mathfrak{L}$ containing z and all the first-order parameters from Γ with $a \models \theta$ (where θ is from Lemma 19) and either $(\text{Hier}[\Gamma](H, W))^\mathfrak{a}$ or $(\text{SLFHier}[\Gamma](H, W))^\mathfrak{a}$. By Lemma 19 (i)(v) with $Z = a$, $\text{Hier}[\Gamma^a](H \cap a, W \cap a)$ or $\text{SLFHier}[\Gamma^a](H \cap a, W \cap a)$. Since $(a, \mathcal{P}^\mathfrak{L}(a)) \models \mathbf{BGZ} + \Pi_\infty^1\text{-CA}$ (where recall $\mathbf{BGZ} \equiv \mathbf{N BG} - \mathbf{Repl}$), on the other hand, $(a, \mathcal{P}^\mathfrak{L}(a)) \models \Delta_0^1\text{-TR} \wedge \Delta_0^1\text{-SLFTR}$. Thus $(\text{Hier}[\Gamma^a])^\mathfrak{L}(h, W \cap a)$ or $(\text{SLFHier}[\Gamma^a])^\mathfrak{L}(h, W \cap a)$ for some $h \in \mathfrak{L}$, since $W \cap a \in \mathfrak{L}$. Again by Lemma 19 and $(\Gamma^a)^\mathfrak{L} = \Gamma^a$, the superscript \mathfrak{L} can be omitted. By the uniqueness of such h , $H \cap a = h$ and so $H \cap z = z \cap (H \cap a) = z \cap h \in \mathfrak{L}$.

8 First Order Systems

Whereas we have been working on second order languages, as was mentioned several times many of our results can be expressed in first order settings, in the same sense as, in number theory, the first order systems \widehat{ID}_α and ID_α can be seen as refinements of the second order $\Delta_0^1\text{-FP}$ and $\Delta_0^1\text{-LFP}$ respectively. In what follows, \mathcal{L}^1 is the first order part of \mathcal{L}^2 , i.e., the language \mathcal{L}_S^1 of the first order set theory or that \mathcal{L}_N^1 of the first order number theory; and **BT**¹ is (a variant of) Zermelo-Fraenkel set theory or Peano Arithmetic.

Definition 19 Let \mathcal{L} be a language extending \mathcal{L}^1 and T a schematically extension of **BT**¹ in \mathcal{L} . $\mathcal{L}[X_0, \dots, X_k]$ is \mathcal{L} augmented by unary predicates X_0, \dots, X_k where $X_i(y)$ is denoted by $y \in X_i$. Following this convention, Δ_0^1 , Σ_n^0 and Π_n^0 have the obvious meanings in $\mathcal{L}[\mathbf{X}]$. An \mathcal{L} -operator form has of the form $\lambda \mathbf{y}. \lambda \mathbf{X}, \mathbf{Y}. \{z \mid \varphi(z, \mathbf{y}, \mathbf{X}, \mathbf{Y})\}$ as in \mathcal{L}^2 . Let \mathcal{F} be a class of $\mathcal{L}[X, Y]$ -formulae.

(1) $\mathcal{F}\text{-}\widehat{ID}_\alpha[T]$ is the $\mathcal{L}' = \mathcal{L} \cup \{F_\Gamma \mid \Gamma \text{ is a } \mathcal{F} \text{ operator of arity } (1, 1, 1)\}$ -theory generated by

- T with all the schematic axioms extended to the language \mathcal{L}' ;
- $(\forall x \prec \alpha)((F_\Gamma)_x = \Gamma(x, (F_\Gamma)_x, (F_\Gamma)^{\prec x}))$.

(2) $\mathcal{F}\text{-}ID_\alpha[T]$ is the $\mathcal{L}'' = \mathcal{L} \cup \{L_\Gamma \mid \Gamma \text{ is a } \mathcal{F} \text{ operator of arity } (1, 1, 1)\}$ -theory generated by

- T with all the schematic axioms extended to the language \mathcal{L}'' ;
- $(\forall x \prec \alpha)((L_\Gamma)_x = \Gamma(x, (L_\Gamma)_x, (L_\Gamma)^{\prec x}))$; and
- $(\forall x \prec \alpha)((\forall y)(y \in \Gamma(x, \varphi, (L_\Gamma)^{\prec x}) \rightarrow \varphi(y)) \rightarrow (\forall y \in (L_\Gamma)_x) \varphi(y))$ for any \mathcal{L}'' -formula φ .

(3) $\mathcal{F}\text{-}ID_\alpha^+[T]$ is the $\mathcal{L}''' = \mathcal{L} \cup \{O_\Gamma \mid \Gamma \text{ is a } \mathcal{F} \text{ operator of arity } (1, 1, 1)\}$ -theory generated by

- T with all the schematic axioms extended to the language \mathcal{L}''' ;
- $(\forall x \prec \alpha)\text{SCSP}[\lambda X.\Gamma(x, X, (O_\Gamma)^{\prec x})]((O_\Gamma)_x)$; and
- $(\forall x \prec \alpha)\text{TI}[\varphi]((O_\Gamma)_x)$ for any \mathcal{L}''' -formula φ .

Sometimes we omit “ Δ_0^1 ” and “[\mathbf{BT}^1]”: e.g., both ID_α and $ID_\alpha[\mathbf{BT}^1]$ denote $\Delta_0^1\text{-}ID_\alpha[\mathbf{BT}^1]$.

Note that $\mathcal{F}\text{-}ID_\alpha[\mathcal{F}\text{-}ID_\beta[T]]$ is mutually interpretable with $\mathcal{F}\text{-}ID_{\beta+\alpha}[T]$ for reasonable \mathcal{F} .

Remark 11 We have the following interpretability results in all of which \mathcal{L} -formulae are preserved:

(I) Remark 3 (3) can now be extended to: $(\Sigma_1^0 \vee \Pi_1^0)\text{-}\widehat{ID}_\alpha[T]$ and $(\Sigma_1^0 \vee \Pi_1^0)\text{-}ID_\alpha[T]$ are mutually interpretable with $\Delta_0^1\text{-}\widehat{ID}_\alpha[T]$ and with $\Delta_0^1\text{-}ID_\alpha[T]$ respectively.

(II) Remark 5 (2) states that $\Pi_1^0\text{-}ID_\alpha[T]$ is interpretable in T if $T \vdash \mathcal{L}\text{-Ref}$.

(III) Remark 6 is now extended to: $\Delta_0^1\text{-}ID_\alpha[T]$ is interpretable into $\Delta_0^1\text{-}ID_\alpha^+[T]$.

(IV) Remark 7 (2) states that $\Delta_0^1\text{-}ID_1^+[T]$ is interpretable into $\Pi_1^0\text{-}ID_1[\Delta_0^1\text{-}\widehat{ID}_1[T]]$.

By (IV) and (II), $ID_1^+[\mathbf{ZF}]$ is interpretable into $\widehat{ID}_1[\mathbf{ZF}]$. More generally, $\widehat{ID}_\alpha[\mathbf{ZF}]$, $ID_\alpha[\mathbf{ZF}]$ and $ID_\alpha^+[\mathbf{ZF}]$ are all pairwise mutually interpretable in such a way that all \mathcal{L}_S -formulae are preserved.

Remark 12 Proposition 16 shows that $\mathbf{BT}^2 + \Delta_0^1\text{-LFTR}_{\omega^\alpha} + \mathcal{L}^2\text{-AxSch}$ can be interpreted into $ID_{\omega^{\alpha+1}}$, where $\mathcal{L}^2\text{-AxSch}$ denotes the axiom schemata (i.e., induction schema in number theory; separation, collection and transfinite induction schemata in set theory). For the converse, we can interpret $ID_{\omega^{\alpha+1}}^+$ into $\mathbf{BT}^2 + \Delta_0^1\text{-SLFTR}_{\omega^\alpha} + \mathcal{L}^2\text{-AxSch}$ as follows: By induction on n , we can show $(\exists! H)\text{SLFHier}[\Gamma](H, \omega^\alpha \cdot n)$. Then O_Γ can be interpreted as

$$\{(\xi, x) \mid (\exists n \in \omega)(\exists H)(\text{SLFHier}[\Gamma](H, \omega^\alpha \cdot n) \wedge (\xi < \omega^\alpha \cdot n) \wedge (x \in (H)_\xi))\}.$$

Thus, by Corollary 9 (iii)(a), the theories $\mathbf{BT}^2 + \Delta_0^1\text{-LFTR}_{\omega^\alpha} + \mathcal{L}^2\text{-AxSch}$, $ID_{\omega^{\alpha+1}}[\mathbf{BT}^1]$, $ID_{\omega^{\alpha+1}}^+[\mathbf{BT}^1]$ and $\mathbf{BT}^2 + \Delta_0^1\text{-SLFTR}_{\omega^\alpha} + \mathcal{L}^2\text{-AxSch}$ are mutually interpretable.

If $\mathbf{BT}^1 \vdash \mathcal{L}\text{-Ref}$, we can add $\mathbf{BT}^2 + \Delta_0^1\text{-FTR}_{\omega^\alpha} + \mathcal{L}^2\text{-AxSch}$ and $\widehat{ID}_{\omega^{\alpha+1}}[\mathbf{BT}^1]$.

9 Generalization of BT²

As mentioned several times before, most of our results have analogues in higher order number theory and in higher order set theory. In other words, although we have been considering \mathbf{BT}^2 as \mathbf{ACA}_0 or (variants of) \mathbf{NGB} (or \mathbf{BGZ}), we can consider \mathbf{BT}^2 as any systems with certain conditions without losing most results. Among such theories is $\Delta_0^{n+2}\text{-CA}_0$ from [17].

In this section we see the certain conditions which guarantee the general treatment of our proofs in the previous sections.

Definition 20 (s.e.(\mathcal{L})) For a first order language \mathcal{L} with at most finitely many sorts and with equalities for all the sorts, the second order extension s.e.(\mathcal{L}) is defined as follows.

(sort) The sorts of s.e.(\mathcal{L}) are those of \mathcal{L} , called *first order*, and a distinct one called *second order*.

(functions) The function symbols of s.e.(\mathcal{L}) are those of \mathcal{L} with the arity assignment preserved.

(predicates) The predicate symbols of s.e.(\mathcal{L}) are those \mathcal{L} with the arity preserved, and a distinct binary one \in , whose first and second arguments are of first order and of second order, respectively.

Note that, although the equalities are included in \mathcal{L} , the equality for the second order is not included as a primitive symbols and is defined by extensionality and the finiteness of sorts.

The languages \mathcal{L}_N^2 of second order number theory and \mathcal{L}_S^2 of second order set theory are our prototypical examples of second order extension. Indeed they are the second order extensions of the language \mathcal{L}_N^1 of first order number theory and of the language \mathcal{L}_S^1 of first order set theory.

Since second order extension is itself a first order language with one more sorts (with the defined equality for the second order), we can apply the operator s.e. repeatedly.

Example 15 *The languages \mathcal{L}_N^{n+1} (or \mathcal{L}_S^{n+1}) of $n+1$ -th order number (or set) theory is the second order extension of that \mathcal{L}_N^n (or \mathcal{L}_S^n , respectively) of n -th order number (or set) theory.*

By convention, lower case Latin letters denote the terms (or objects, from a semantical point of view) of the first order, while upper ones denote those of the second order. We define $\Delta_0^1, \Sigma_n^1, \Pi_n^1$ and their relativizations $\Delta_0^1(\mathcal{F}), \Sigma_n^1(\mathcal{F})$ and $\Pi_n^1(\mathcal{F})$ as usual.

Along the line of Example 15, Σ_k^1 and Π_k^1 in \mathcal{L}_N^{n+1} are, in the traditional terminology, denoted by Σ_k^n and Π_k^n respectively. In our terminology, all the orders strictly below the $n+1$ -th are considered as the first order and only the $n+1$ -th order as the second order.

Assumption 21 ($\mathbf{BT}^2, \Delta_0^0, \Sigma_n^0, \Pi_n^0$) \mathbf{BT}^2 denotes any s.e.(\mathcal{L}^1)-theory satisfying the following:

1. the only non- Π_1^1 -axiom of \mathbf{BT}^2 is $\Delta_0^1\text{-CA}$:
 $\Delta_0^1\text{-CA } (\forall \mathbf{X}, \mathbf{x})(\exists Z)(\varphi(z, \mathbf{x}, \mathbf{X}) \leftrightarrow z \in Z)$, for any Δ_0^1 formula φ without free Z ;
2. there are Δ_0^1 formulae $x \in \omega$ and $x < y$ such that $\mathbf{BT}^2 \vdash (\forall X)((\omega, <, X) \models \mathbf{PA}[X])$, where $\mathbf{PA}[X]$ is \mathbf{PA} , formulated only with $<$, with the induction scheme extended to $\mathcal{L}_N^1[X]$;
3. there are Δ_0^1 formulae $\text{finseq}(x, n)$ and $\text{ev}(x, k; v)$ such that \mathbf{BT}^2 proves the basic properties expected from the intention “ x is finite sequence of length n ” and “ k -th component of x is v ”, e.g., $(\exists x)\text{finseq}(x, 0), (\forall x)(\forall n \in \omega)[\text{finseq}(x, n) \rightarrow (\forall k < n)(\exists!v)\text{ev}(x, k; v) \wedge (\forall u)(\exists z)(\text{finseq}(z, n+1) \wedge \text{ev}(z, n; u) \wedge (\forall k < n)(\exists v)(\text{ev}(z, k; v) \wedge \text{ev}(x, k; v))]$;
4. there is a class \mathcal{F} of formulae, containing atomic formulae and their negations, such that
 - (a) there is a \mathcal{F} -formula pair($x, y; z$) such that \mathcal{F} is closed under the quantifiers of the forms
 $(\exists y)(\text{pair}(x, y; z) \wedge \dots)$ and $(\forall y)(\text{pair}(x, y; z) \rightarrow \dots)$
and that $\mathbf{BT}^2 \vdash (\forall x, y, u, v)[(\exists!z)\text{pair}(x, y; z) \wedge (\forall z)(\text{pair}(x, y; z) \wedge \text{pair}(u, v; z) \rightarrow x = u \wedge y = v)]$;
 - (b) (optionally) there is a \mathcal{F} -formula $\pi(e, w, x, X, Y)$ at most e, w, x, X, Y free, in which X occurs only positively, such that for any \mathcal{F} -formula $\varphi(w, x, X, Y)$ in which X occurs only positively,
 $\mathbf{BT}^2 \vdash (\forall x, X, Y)[(\exists w)\varphi(w, x, X, Y) \leftrightarrow (\exists w)\pi(\neg\varphi^\top, w, x, X, Y)]$.

In what follows, we fix such \mathcal{F} and $\text{pair}(x, y; z)$, which are denoted by Δ_0^0 and “ $\langle x, y \rangle = z$ ”. Then we can define Σ_n^0 and Π_n^0 starting with $\Sigma_0^0 = \Pi_0^0 = \Delta_0^0$, and relativizations $\Sigma_n^0(\mathcal{F})$ and $\Pi_n^0(\mathcal{F})$ accordingly. $(Y)_x$ denotes $\{z \mid \langle x, z \rangle \in Y\}$ as usual. Based on π , we define the X -positive Σ_1^0 universal formula:

$$\Upsilon_1^0(e, x, X^+, Y) \equiv (\exists w)\pi(e, w, x, X, Y).$$

Remark 13 Since Δ_0^0 contains all atomic formulae and their negations, and since \mathbf{BT}^2 is on classical logic, for any elementary formula $\varphi(\mathbf{x})$ there are k and $\psi(\mathbf{x}, y_1, \dots, y_k)$ in Δ_0^0 with $\mathbf{BT}^2 \vdash (\forall \mathbf{x})[\varphi(\mathbf{x}) \leftrightarrow (\forall y_1)(\exists y_2)\dots(Qy_k)\psi(\mathbf{x}, y_1, \dots, y_k)]$. Thus, $\Delta_0^1\text{-CA}$ can be replaced by $\Sigma_1^0\text{-CA}$.

By Assumption 21.1 and 21.2, \mathbf{BT}^2 contains induction on ω for all Δ_0^1 formulae.

Example 16 *In the cases of $\mathcal{L}^2 = \mathcal{L}_N^2$ and $\mathcal{L}^2 = \mathcal{L}_S^2$, \mathbf{ACA}_0 and \mathbf{NBG} (and its variants, with a modification for **Choice**, that is, we introduce a new predicate **W**) can be \mathbf{BT}^2 with 4(b) satisfied. \mathbf{BGZ} can also be \mathbf{BT}^2 but not the optional 4(b).*

For $\mathcal{L}^2 = \mathcal{L}_N^{n+2}$ or $= \mathcal{L}_S^{n+2}$, $\Delta_0^{n+1}\text{-CA}_0$ defined in [17] can be \mathbf{BT}^2 with 4(b), with a modification: We require that k -th order objects be also $k+1$ -th order, and we define $x^{k+1} \in^k y^{k+1}$ by $(\exists u^k \in^k y^{k+1})(x^{k+1} = u^k)$, so that the highest order objects may have objects of all the lower orders. This is inessential, because k -th order part can be embedded into $k+1$ -th order by singleton.

However, we need more assumption for reflection principle: Though Definition 18 allows us to define the relativization for second order objects in general \mathcal{L}^2 , for first order objects as required in $\Delta_0^1\text{-Ref}$, we have to specify the way how to consider first order objects as second order.

Assumption 22 We assume there is an \mathcal{L}^1 formula, denoted by $x \in y$, without other free variables, such that Δ_0^0 is closed under quantifiers bounded by this \in , i.e., of the form $(Qx \in y)$.

Example 17 For $\mathcal{L}^1 = \mathcal{L}_S^1$, the language of set theory, we can choose the elementhood \in as the specified binary relation. For $\mathcal{L}^1 = \mathcal{L}_N^1$, that of number theory, we can choose, say, the $\Delta_0^0(\exp)$ relation “the x -th digit of binary expression of y is 1” (see Remark 4). For $\mathcal{L}^1 = \mathcal{L}_N^{n+3}, \mathcal{L}_S^{n+3}$, those of higher order number and set theories, we take the relation $(\exists u^n)(x^{n+1} = (y^{n+1})_u) \vee x^{n+1} \in^n y^{n+1}$ (cf. Example 16), where $(y)_u$ is the section of y at u , as defined in [17, Definition 3].

Though \in we have introduced must not be confused with the relation symbol also denoted by \in in Definition 20, if we identify a first order a with the second order $\{z \mid z \in a\}$, the two relations can be identified. Accordingly we let φ^a denote φ^Z for $Z = \{x \mid x \in a\}$. Based on this convention, $\mathcal{F}\text{-Ref}$ defined in Definition 10 makes sense in general \mathcal{L}^2 with our assumptions.

Example 18 For $\mathcal{L}^2 = \mathcal{L}_S^2$, as mentioned before $\Delta_0^1\text{-Ref}$ holds in **NBG** and in **NBGC**.

For $\Delta_0^{n+2}\text{-CA}_0$ from [17, Definition 5], Löwenheim-Skolem argument yields Δ_0^{n+2} -reflection.

Note that, in all of these examples, the clause $(\forall X)(\exists x)(x = "a \cap X")$ is redundant.

Thus all the results in Sections 3, 4, 5, 6 and A, except Subsection 3.3, can be generalized for such general **BT**², with literally the same proofs (with Ω being the order-type of the fixed global well-ordering), where Subsections 6.2 and 6.3 require the optional 4(b). We could say that Figure 1 shows the contrast between **BT**² with countable first order and **BT**² with reflection. Note that since the framework of [4] is much more general the results obtained there also hold for our **BT**².

We close this subsection by informally mentioning the generalization of equiconsistency results in Section 7 to the higher order set theory (with the formulation given in [17]).

For the conservation between the extensions of $\Delta_0^{n+1}\text{-CA}_0$ from [17, Def.5] with and without **Found** in the first order, it seems convenient to restrict also the higher order domains (except the highest) as follows: The second order domain to $\mathfrak{W}^2 = \{x^1 \mid x^1 \subset \mathfrak{W}\}$; the third to $\mathfrak{W}^3 = \{x^2 \mid x^2 \subset \mathfrak{W}^2\}$; and so on. Then, since $x^k \in \mathfrak{W}^{k+1}$ is Δ_0^k (while x^k is of type k , i.e., $k+1$ -th order), the argument for Proposition 20 and Theorem 14 works well for $n+3$ -th order set theory.

For global well ordering on the lower order parts, we only need the notion of constructibility for lower orders because then the argument of Proposition 21 survives. This is planned to be discussed in the author’s future work [16].

A Disanalogy to Conventional Ordinal Analysis

In the proof of Proposition 21, we used a fact of the form: Strong principles (e.g., even full comprehension) hold in the sense of a reflecting first order object a , because the last clause of our version of reflection implies $(QX)\varphi^a(X) \leftrightarrow (Qx)\varphi^a(x)$. This explains (only informally) why the analogues of conventional ordinal analysis do not work for **NBG** (and other $\Delta_0^n\text{-CA}_0$ ’s).

One might ask if the supremum of provably well-founded Δ_1^0 relations of **NBG** can be described as E_0 , which Jäger and Krähenbühl [7, §3] introduced as the set-theoretic analogue of the ordinal ε_0 to obtain analogues of some results known in number theory, and if that of **NBG** + $\Delta_0^1\text{-TR}$ can be described by defining the analogue of Γ_0 similarly, and so on. These are not the case as shown below.

Let \prec be the analogue of any known notation system of recursive ordinal (of course, here is some ambiguity, e.g., the definition of E_0 is obtained by replacing some occurrences of ω in the definition of ε_0 with Ω but some ω remaining). It is at least Δ_0^1 . All such notations are provably well-founded in at least full second order number theory. This is the case even for the notations for the strongest systems whose ordinal analysis has ever been done. Thus it is natural to assume **BGZ** + $\Pi_\infty^1\text{-CA} \vdash \text{WF}(\succ)$. By reflection in **NBG**, we have a with $\neg\text{WF}(\succ) \rightarrow (\neg\text{WF}(\succ))^a$. Since $(\Pi_\infty^1\text{-CA})^a$ as explained above and $(\text{BGZ})^a$, we have $(\text{WF}(\succ))^a$ and so $\text{WF}(\succ)$.

Thus we could conclude (informally, not as a formal mathematical theorem) that the supremum of provably well-founded Δ_1^0 relations of **NBG** (and of $\Delta_0^{n+2}\text{-CA}_0$ ’s) is far beyond our knowledge; not to mention that of the extensions.

Acknowledgments The author would like to express his gratitude to the anonymous referee for his or her careful reading and valuable comments. He also greatly appreciates very stimulating discussions with Kentaro Fujimoto. He would like to thank members of Logic and Theory Group at Institut für Informatik und angewandte Mathematik, Universität Bern and all the participants of a reading group on the earlier version of this paper, especially Gerhard Jäger, Thomas Strahm, Dieter Probst and Dandolo Flumini.

References

1. Avigad, J.: On the relationships between ATR_0 and $\widehat{ID}_{<\omega}$. *J. Symbolic Logic* 61(3), 768-779 (1996)
2. Feferman, S.: Operational set theory and small large cardinals. *Inform. Comput.* 207(10), 971-979, (2009)
3. Feferman, S.: Is the Continuum Hypothesis a definite mathematical problem? <http://math.stanford.edu/~feferman/papers/IsCHdefinite.pdf>. Accessed 2011.
4. Flumini, D., Sato, K.: From hierarchy to well-foundedness. *Arch. Math. Logic*, published online (2014)
5. Fujimoto, K.: Classes and truths in set theory. *Ann. Pure Appl. Logic* 163(11), 1484-1523 (2012)
6. Jäger, G.: Full operational set theory with unbounded existential quantification and power set. *Ann. Pure Appl. Logic* 161(1), 33-52 (2009).
7. Jäger, G., Krähenbühl, J.: Σ^1_1 choice in a theory of sets and classes. In: R. Schindler (ed.) *Ways of Proof Theory*, pp.283-314, Ontos Verlag (2010)
8. Jäger, G., Strahm, T.: Fixed point theories and dependent choice. *Arch. Math. Logic* 39(7), 493-508 (2000)
9. Krähenbühl, J.: *On the relationship between choice schemes and iterated class comprehension in set theory*. Ph.D. Thesis, Universität Bern (2011)
10. Kunen, K.: *Set Theory: An Introduction to Independent Proofs*. Elsevier, Amsterdam-New York (1980)
11. Louveau, A.: Some results in the Wadge hierarchy of Borel sets. In: *Cabal seminar '79-'81*, 28-55, Lecture Notes in Math., 1019, Springer, Berlin (1983)
12. Moschovakis, Y.: *Elementary Introduction on Abstract Structures*. North Holland, Amsterdam-London (1974)
13. Nemoto, T.: Determinacy of Wadge classes and subsystems of second order arithmetic, *Math. Log. Q.* 55(2), 154-176 (2009).
14. Sato, K.: The strength of extensionality I: weak weak set theories with infinity, *Ann. Pure Appl. Logic* 157(2-3), 234-268 (2009).
15. Sato, K.: The strength of extensionality II: weak weak set theories without infinity, *Ann. Pure Appl. Logic* 162(8), 579-646 (2011).
16. Sato, K.: The strength of extensionality III: weak weak set theories with higher infinite, in preparation.
17. Sato, K.: Relative predicativity and dependent recursion in second-order set theory and higher-order theories, accepted for publication in *J. Symbolic Logic*. <http://www.iam.unibe.ch/ltpub/2012/sat12.pdf>.
18. Simpson, S.: *Subsystems of Second Order Arithmetic*. Springer, Berlin (1999)
19. Strahm, T.: Autonomous Fixed Point Progressions and Fixed Point Transfinite Recursion. in: Buss, S., Hajek, P., Pudlak, P. (eds.) *Logic Colloquium 98*, pp.449-464 (2000)