On interpretations of arithmetic and set theory*

Richard Kaye and Tin Lok Wong
School of Mathematics
University of Birmingham
Birmingham
B15 2TT
U.K.

15 July, 2006

1 Introduction

The work described in this article starts with a piece of mathematical 'folklore' that is 'well known' but for which we know no satisfactory reference.[1]

Folklore Result. The first-order theories Peano arithmetic and ZF set theory with the axiom of infinity negated are equivalent, in the sense that each is interpretable in the other and the interpretations are inverse to each other.

This would make an excellent starting point for any beginning research student working in the area of models of Peano arithmetic, since an understanding of how these interpretations work enables all coding techniques from set theory to be employed in arithmetic in a uniform way, and also places PA 'on the map' (in the sense of consistency and interpretation strength) relative to set theory. This, combined with the lack of suitable references, permits us, we trust, to omit any apology for investigating such folklore. In fact, when the details were finally uncovered, there were surprises for both authors—this is the reason for the quotation marks when we said that the folkloric result above is 'well known'.

We should be more precise and highlight exactly where the imprecision in the folkloric result lies. There are two places: firstly, the notion of 'ZF set theory with the axiom of infinity negated' turns out to be dependent in quite an important way on the preferred choice of the initial presentation of ZF; secondly, the idea of interpretations

*This is a slightly expanded version of a paper to appear in the *Notre Dame Journal of Formal Logic* under the same title. We have taken the opportunity in this web version of the paper to add a few additional comments and references to some additional material.

[1] **Further references.** Świerczkowski [20] wrote a detailed exposition of the equivalence of PA and "the theory of hereditarily finite sets", which uses an additional *adduction operator*. This theory of sets is said be adapted from Tarski and Givant [25] with slight variations. The adduction operator $\lhd$, defined by

$$x \lhd y = x \cup \{y\},$$

is also studied by Kirby [13] in the context of ZF without the axiom of infinity.

Another theory of sets that was proved to be "equivalent" to PA is PS introduced by Previale [17]. However, the interpretations there seem to be the Ackermann one and the ordinal one (cf. Kirby [13, §2] and Sections 3, 4).

being 'inverse to each other' admits important variations. For example, Chang and Keisler [5, §A.31] specify one particular choice of axiomatisation of ZF; for this axiomatisation a weak form of interpretation-equivalence of 'ZF with infinity negated' and PA can be proved, but for stronger notions of interpretation-equivalence a different axiomatisation of ZF seems to be required.

Our notation and terminology is standard. For a text on set theory we have used Drake [7], though many others would be suitable. For background on Peano arithmetic we refer the reader to Kaye [12], and especially Chapter 10 for background on subsystems of PA.

2 Interpretations

The first issue to address here is what we shall mean by 'interpretation' and when interpretations are 'inverse' to one another. In fact, for the folkloric result we are addressing, the notion of interpretation required is very straightforward and concrete, but in general the word 'interpretation' is used in many contexts and with many different meanings in logic and model theory.

As pointed out elsewhere, notably by Visser [26], notions of interpretation can be best brought together using a category theoretic framework, though in this paper we use category theory simply as a notational device. For us, an *interpretation* $\mathfrak{i}$ *of a theory* T_2 *in a theory* T_1 is a morphism $\mathfrak{i}\colon T_2 \to T_1$ in a particular category being studied.

To define the particular category we will use in this paper, let us agree that an $\mathscr{L}$*-theory* is a consistent set of $\mathscr{L}$-sentences (not necessarily complete or closed under deduction) and we make the convenient assumptions that all languages are purely relational, equality ($=$) being one of the relation symbols, and a further unary relation Dom() for the domain is always present. Full first-order logic is assumed, including all the rules for equality and the additional logical axiom $\forall x\, \mathrm{Dom}(x)$. For us here, an interpretation $\mathfrak{i}\colon T_2 \to T_1$ of an $\mathscr{L}_2$-theory T_2 in an $\mathscr{L}_1$-theory T_1 is given by a mapping of atomic formulas $R(x_1, x_2, \ldots, x_n)$ of $\mathscr{L}_2$ to formulas $R(x_1, x_2, \ldots, x_n)^{\mathfrak{i}}$ of $\mathscr{L}_1$ in the same free variables. In particular, the domain and equality on this domain are defined by $\mathrm{Dom}(x)^{\mathfrak{i}}$ and $(x = y)^{\mathfrak{i}}$. The mapping extends to all first-order formulas in the natural way by setting $(\neg\theta(\bar{x}))^{\mathfrak{i}}$ to be $\neg\,\theta(\bar{x})^{\mathfrak{i}}$, $(\phi(\bar{x}) \to \psi(\bar{x}))^{\mathfrak{i}}$ to be $\phi(\bar{x})^{\mathfrak{i}} \to \psi(\bar{x})^{\mathfrak{i}}$, and $(\forall y\, \theta(\bar{x}, y))^{\mathfrak{i}}$ to be $\forall y\, (\mathrm{Dom}(y)^{\mathfrak{i}} \to \theta(\bar{x}, y)^{\mathfrak{i}})$. For this to define an interpretation, we shall insist that $T_1 \vdash \exists x\, \mathrm{Dom}(x)^{\mathfrak{i}}$ and $T_1 \vdash \sigma^{\mathfrak{i}}$ for every axiom $\sigma \in T_2$. Visser calls such interpretations *relative interpretations*.

The category of theories and interpretations is obtained by taking theories as objects and such mappings $\mathfrak{i}$ as morphisms, where we choose to say that $\mathfrak{f}\colon T_2 \to T_1$ and $\mathfrak{g}\colon T_2 \to T_1$ are *equal* if $T_1 \vdash \forall \bar{x}\, (R(\bar{x})^{\mathfrak{f}} \leftrightarrow R(\bar{x})^{\mathfrak{g}})$ for all R in $\mathscr{L}_2$. This makes a category, where the identity interpretation $1 = 1_T\colon T \to T$ is the morphism given up to this equality by the mapping $R(\bar{x})^1 = R(\bar{x})$ for all R. The composition $\mathfrak{g}\mathfrak{f}\colon T_3 \to T_1$ of morphisms $\mathfrak{f}\colon T_3 \to T_2$ and $\mathfrak{g}\colon T_2 \to T_1$ is given by $(R(\bar{x}))^{(\mathfrak{g}\mathfrak{f})} = (R(\bar{x})^{\mathfrak{f}})^{\mathfrak{g}}$, and associativity holds as can easily be checked. (It is somewhat annoying that the convenient notation of writing interpretation-applications as superscripts is at odds with the usual maps-on-left convention for morphisms.) As usual, two morphisms $\mathfrak{f}\colon T \to S$ and $\mathfrak{g}\colon S \to T$ are said to be *inverse* to each other if $\mathfrak{f}\mathfrak{g} = 1_S$ and $\mathfrak{g}\mathfrak{f} = 1_T$. An *interpretation* is then a morphism in this category, i.e., a mapping as described in the last paragraph modulo the equivalence given by this notion of equality.

As for many kinds of interpretations, ours respects derivability in first-order logic because of the way our mappings were defined on non-atomic formulas. Thus we have,

Proposition 1. Let $\mathfrak{f}\colon T_2 \to T_1$ and suppose $T_2 \vdash \sigma$. Then $T_1 \vdash \sigma^{\mathfrak{f}}$. □

The next proposition follows from the definitions just given and an induction on formulas.

Proposition 2. Let $\mathfrak{f}\colon T_2 \to T_1$ and $\mathfrak{g}\colon T_2 \to T_1$ be equal as interpretations. Then for each formula $\theta(\bar{x})$ of $\mathscr{L}_2$ we have $T_1 \vdash \forall \bar{x}\,(\theta(\bar{x})^{\mathfrak{f}} \leftrightarrow \theta(\bar{x})^{\mathfrak{g}})$. In particular, for each sentence σ of $\mathscr{L}_2$, $T_1 \vdash \sigma^{\mathfrak{f}}$ if and only if $T_1 \vdash \sigma^{\mathfrak{g}}$. □

3 The Ackermann interpretation

By ZF–inf we mean the theory in the first-order language $\mathscr{L}_\in$ of set theory with all the usual axioms of ZF except infinity, which is negated. More specifically, and following Baratella and Ferro [2], let the set theory EST have the usual axioms or axiom schemes of extensionality, existence of the empty set, pair set, sum set, separation and replacement[2]; denote the axiom of power set by Pow, the axiom of foundation, expressed as $\forall x\,(x \neq \varnothing \to \exists y \in x\, \forall z \in x\; z \notin y)$, by Found; and the usual axiom of infinity, i.e., $\exists w\,(\varnothing \in w \,\wedge\, \forall x \in w\; (x \cup \{x\} \in w))$, by Inf. Then ZF–inf is EST together with Pow, Found, $\neg$Inf.

It was observed in 1937 by Wilhelm Ackermann [1] that $\mathbb{N}$ with the membership relation defined by

$$n \in m \text{ iff the } n^{\text{th}} \text{ digit in the binary representation of } m \text{ is } 1$$

satisfies ZF–inf. This interpretation, formalised in ZF with ω in place of $\mathbb{N}$ yields a bijection between ω and the collection V_ω of hereditarily finite sets.

This interpretation has been given extensive treatment in the literature. Working in PA (or a suitable fragment of PA) one may define the interpretation by setting $\mathrm{Dom}(x)^{\mathfrak{a}}$ to be '$\mathrm{Dom}(x)$', $(x = y)^{\mathfrak{a}}$ to be '$x = y$', and $(x \in y)^{\mathfrak{a}}$ to be

$$\exists w < y\; \exists p \leqslant y\; \exists r < p\; (\; p = 2^x \;\wedge\; y = (2w+1)p + r\;).$$

We will refer to this interpretation throughout the rest of the paper as the *Ackermann interpretation*.

All that is required for this to work is a suitable formula $p = 2^x$ that represents 'p is the x^{th} power of 2'. In fact many such formulas are known which have the necessary inductive definition

$$1 = 2^0 \;\wedge\; \forall x \forall p\,(p = 2^x \to 2p = 2^{x+1}) \tag{1}$$

and partial-function nature

$$\forall x \forall p \forall q\,(p = 2^x \wedge q = 2^x \to p = q) \tag{2}$$

provable in PA or even much weaker systems, including Δ_0 formulas $p = 2^x$ for which the above statements are provable in the subsystem of PA with induction on Δ_0 formulas only (see Gaifman and Dimitracopoulos [9] or Hájek and Pudlák [11, §I.1(b)]).

Making extensive use of the axiom scheme of induction, it is straightforward to check that every model of PA with this binary relation defined satisfies ZF–inf.

Theorem 3. $\mathfrak{a}\colon$ ZF–inf $\to$ PA. □

More information on this interpretation in subsystems of PA is given in Section 7 below.

[2] The axiom scheme of separation is redundant in the presence of replacement.

4 The ordinal interpretation

The finite ordinals in the standard model of ZFC resemble the usual natural numbers. In a world without infinite numbers, one would expect the class of all ordinals to satisfy the ordinary rules of arithmetic.

Definition 4. $\mathrm{Trans}(x)$ ('x is transitive') is the formula

$$\forall y, z\,(z \in y \wedge y \in x \to z \in x),$$

and $x \in \mathrm{On}$ ('x is an ordinal') is

$$\mathrm{Trans}(x) \wedge \forall y, z \in x\,(y \in z \vee y = z \vee z \in y).$$

Most ordinal theory can be done inside our theory of sets just as it is in ZF. See for example Chapter 2 of Drake [7] for details. However, as an easy consequence of $\neg$Inf, every nonzero ordinal is a successor ordinal—there are no limit ordinals in ZF–inf.

Theorem 5. ZF–inf $\vdash \forall x\,\big(x \in \mathrm{On} \leftrightarrow x = \varnothing \vee \exists y \in \mathrm{On}\,(x = y \cup \{y\})\big)$. □

As a corollary, every set of ordinals in ZF–inf has a maximum element.

Corollary 6. ZF–inf $\vdash \forall x \subseteq \mathrm{On}\,(x \neq \varnothing \to \bigcup x \in x)$. □

Arithmetic of ordinals can be defined in the usual way, and there is no problem at all in showing that the ordinals with usual arithmetic yields an interpretation of PA in ZF–inf.

Definition 7. Let $\mathrm{Dom}(x)^{\mathfrak{o}}$ be '$x \in \mathrm{On}$', $(x = y)^{\mathfrak{o}}$ be '$x = y$', $(x < y)^{\mathfrak{o}}$ be '$x \in y$', $(x + y = z)^{\mathfrak{o}}$ be '$x + y = z$' (ordinal addition), and $(x \cdot y = z)^{\mathfrak{o}}$ be '$x \times y = z$' (ordinal multiplication).

Theorem 8. $\mathfrak{o}\colon \mathrm{PA} \to$ ZF–inf. □

Thus PA and ZF–inf have the same consistency strength, but this interpretation is clearly not inverse to the Ackermann interpretation.

5 Epsilon induction and transitive containment

In order to define an inverse to the Ackermann interpretation we shall need to consider the principle of $\in$-induction.[3]

Definition 9. For an $\mathscr{L}_{\in}$-formula $\phi(x, \bar{y})$, $\mathrm{I}_{\in}\phi$ is the sentence

$$\forall \bar{y}\,\big(\forall x\,(\forall w \in x\,\phi(w, \bar{y}) \to \phi(x, \bar{y})) \to \forall x\,\phi(x, \bar{y})\big).$$

$\in$-Ind denotes the scheme $\{\mathrm{I}_{\in}\psi : \psi(x, \bar{y})$ is an $\mathscr{L}_{\in}$-formula$\}$.

It turns out that not every model of ZF–inf admits $\in$-induction. This is proved using the closely related notion of the *transitive closure* of a set.

[3]**Further references.** The adduction operator previously mentioned gives rise to another scheme of induction

$$\phi(\varnothing) \wedge \forall x \forall y\big(\phi(x) \wedge \phi(y) \to \phi(x \lhd y)\big) \to \forall x\,\phi(x),$$

which is said to have originated from Givant and Tarski [10] (cf. Kirby [13, §2]).

Definition 10. Define $y = \mathrm{TC}(x)$ ('y is the transitive closure of x') to be

$$y \supseteq x \wedge \mathrm{Trans}(y) \wedge \forall y' \, (y' \supseteq x \wedge \mathrm{Trans}(y') \to y' \supseteq y).$$

TC is the axiom $\forall x \, \exists u \supseteq x \, \mathrm{Trans}(u)$.

It is easy to check (using extensionality) that the transitive closure $y = \mathrm{TC}(x)$ of a set x, if it exists, is unique. The $\mathscr{L}_\in$-sentence TC says that every set is *contained* in a transitive set. This turns out to be equivalent to the apparently stronger statement that every set has a transitive closure over a weak fragment of ZF–inf, because the transitive closure of a set x can alternatively be defined as the intersection of all transitive sets containing x.

Lemma 11. $\mathrm{EST} \vdash \forall x \, (\exists y \, (y \supseteq x \wedge \mathrm{Trans}(y)) \to \exists y \, (\mathrm{TC}(x) = y))$. □

Over $\mathrm{EST} + \mathrm{Found}$, the axiom of transitive containment or closure is equivalent to $\in$-induction. In fact, a single instance of $\in$-induction is enough to prove the whole schema $\in$-Ind.

Proposition 12. For all $V \models \mathrm{EST} + \mathrm{Found}$,

$$V \models \in\text{-Ind} \Leftrightarrow V \models \mathrm{TC}.$$

Proof. (Sketch.) Let $V \models \mathrm{EST} \cup \in\text{-Ind}$. We prove the equivalent statement

$$V \models \forall x \, \exists y \, (y = \mathrm{TC}(x))$$

by $\in$-induction on x in V.

Let $x \in V$ such that every element of x has a transitive closure. By the axioms of EST and the induction hypothesis, the set

$$\bigcup \{ z : \exists x' \in x \, (z = \mathrm{TC}(x')) \} \cup x$$

exists and is a transitive superset of x. So Lemma 11 implies x has a transitive closure, completing the induction.

The proof of the converse requires the axiom of foundation but is the same as the standard proof of $\in$-induction in ZF. See, for example, Chapter 2 of Drake [7]. □

The statement TC is not however provable from ZF–inf. We sketch a proof here due to Mancini. (See Mancini and Zambella's article [15] for more details.)

Theorem 13. $\mathrm{ZF{-}inf} \cup \{\neg \mathrm{TC}\}$ is consistent.

Proof. (Sketch.) Consider the hereditarily finite sets $(V_\omega, \in)$, and the set of ordinals ω in V_ω. Let

$$\omega^* = \{ \{ x \cup \{x\} \} \in V_\omega : x \in \omega \},$$

the set of singletons of nonzero elements of ω. Define $F \colon V_\omega \to V_\omega$ by

$$F(x) = \{ x \cup \{x\} \} \text{ and } F(\{ x \cup \{x\} \}) = x$$

for $x \in \omega$ and $F(x) = x$ for $x \notin \omega \cup \omega^*$. Thus, F is an involution, i.e., a permutation of V_ω which is a disjoint product of 2-cycles. Now define the binary relation $\in_F$ by

$$\forall x, y \in V_\omega \; (x \in_F y \Leftrightarrow x \in F(y)).$$

It can be verified that $(V_\omega, \in_F) \models \mathrm{ZF{-}inf}$. In addition, $(V_\omega, \in_F) \models \neg \mathrm{TC}$. In particular,

$$(V_\omega, \in_F) \models \neg \exists u \, (u \supseteq \varnothing \wedge \mathrm{Trans}(u))$$

since any such set is necessarily infinite. □

Thus, one really needs to add the extra hypothesis TC for the $\in$-induction argument to work.

It seems that this result, or ones like it, were known before Mancini and Zambella's paper. Kunen [14, p149(29)] attributes a result like this to Barwise though this model fails to satisfy the power set axiom, it seems. And Baratella and Ferro [2, p345] point out that the independence of transitive containment relative to other axioms of finite set theory was proved in the context of the Alternative Set Theory by Sochor. They also construct a related model (which they attribute to Kunen) in which transitive containment fails, but one in which the axiom of foundation also fails.

The following easy but important proposition shows that any inverse to the Ackermann interpretation must use TC in some form, even for rather weak systems of arithmetic.

Proposition 14. Let $\mathfrak{a}\colon \text{ZF–inf} \to \text{PA}$ be the Ackermann interpretation. Then $\text{I}\Delta_0 \vdash \text{TC}^{\mathfrak{a}}$.

Proof. For $x \in M \vDash \text{I}\Delta_0$, let $y \in M$ be one less than the smallest power of 2 that is bigger than x. (Such a y exists by Δ_0 induction as there must be a greatest u such that the Δ_0 formula $\exists z \leqslant x\; 2^u = z$ holds.) If $M \vDash (u \in v \wedge v \in y)^{\mathfrak{a}}$ then $u < v < y$ so the u^{th} binary digit of y is 1. $\square$

In other words, $\mathfrak{a}\colon \text{ZF–inf} + \text{TC} \to \text{PA}$. If $\mathfrak{b}\colon \text{PA} \to \text{ZF–inf}$ is inverse to $\mathfrak{a}$ then we would have $\text{ZF–inf} \vdash (\text{TC}^{\mathfrak{a}})^{\mathfrak{b}}$ hence $\text{ZF–inf} \vdash \text{TC}$ by Proposition 2, which is impossible. Thus we must consider the theory $\text{ZF–inf}^* = \text{ZF–inf} + \text{TC}$ rather than ZF–inf itself.

6 The inverse Ackermann interpretation

Equipped with $\in$-induction, we shall obtain an inverse interpretation $\mathfrak{b}\colon \text{PA} \to \text{ZF–inf}^*$. The plan is to define a natural bijection $\mathfrak{p}\colon V \to \text{On}$ between the whole universe and the ordinals. The required interpretation can then be obtained by composing this map with the map $\mathfrak{o}$ defined on On.[4]

At first, it appears difficult to see how to use $\in$-induction at all, since the required inductive definition of $\mathfrak{p}$ is $\mathfrak{p}(x) = \sum_{y \in x} 2^{\mathfrak{p}(y)}$ and this seems to need a separate induction on the cardinality of x—just the sort of induction we don't yet have and are trying to justify. However there is a way round this problem using ordinal summation.

Definition 15. Working in $V \vDash \text{ZF–inf}$, let $\mathscr{P}(\text{On})$ denote the class of sets of ordinals, and let $\widehat{\Sigma}\colon \text{On} \times \mathscr{P}(\text{On}) \to \text{On}$ be the function defined recursively by

$$\widehat{\Sigma}(0,x) = 0$$

for all $x \in \mathscr{P}(\text{On})$, and

$$\widehat{\Sigma}(c \cup \{c\},x) = \begin{cases} \widehat{\Sigma}(c,x), & \text{if } c \cup \{c\} \notin x, \\ \widehat{\Sigma}(c,x) + (c \cup \{c\}), & \text{if } c \cup \{c\} \in x \end{cases}$$

for all $c \in \text{On}$ and $x \in \mathscr{P}(\text{On})$. Also, let $\Sigma\colon \mathscr{P}(\text{On}) \to \text{On}$ be the function defined by

$$\forall x \subseteq \text{On}\; \Sigma(x) = \widehat{\Sigma}\big(\bigcup x, x\big).$$

[4] **Further references.** The inverse Ackermann interpretation is said to have been conjectured to exist in 1964 by Beth [4], and it was found by Mycielski [16] in the same year (cf. Kirby [13, §2]).

Informally, this defines

$$\Sigma(x) = \sum_{y \in x} y$$

for a set of ordinals x, where the summation on the right hand side of the equation refers to ordinal addition. Using induction on the ordinals, it can be proved that both $\widehat{\Sigma}$ and Σ are class functions definable over ZF–inf. The following definition is the place where TC comes in.

Definition 16. In $V \models \text{ZF–inf}^*$, define $\mathfrak{p}\colon V \to \text{On}$ recursively by

$$\mathfrak{p}(x) = \Sigma\big(\{2^{\mathfrak{p}(y)} \in \text{On} : y \in x\}\big).$$

As indicated, $\mathfrak{p}$ is again a class function that can be defined by $\in$-recursion on all sets over ZF–inf*. Informally, in ZF–inf* we have

$$\mathfrak{p}(x) = \sum_{y \in x} 2^{\mathfrak{p}(y)}.$$

This is precisely the bijection we are looking for.

Proposition 17. ZF–inf* proves that $\mathfrak{p}$ is a bijective class function $V \to \text{On}$.

Proof. The injectivity part is proved by $\in$-induction on all sets. The surjectivity part is proved by $\in$-induction on the ordinals. □

Definition 18. Define a mapping $\mathfrak{b}$ of atomic sentences of arithmetic into set theory by saying $\text{Dom}(x)^{\mathfrak{b}}$ is '$\text{Dom}(x)$', $(x = y)^{\mathfrak{b}}$ is '$x = y$', $(x < y)^{\mathfrak{b}}$ is '$\mathfrak{p}(x) < \mathfrak{p}(y)$', $(x + y = z)^{\mathfrak{b}}$ is '$\mathfrak{p}(x) + \mathfrak{p}(y) = \mathfrak{p}(z)$', and $(x \cdot y = z)^{\mathfrak{b}}$ is '$\mathfrak{p}(x) \times \mathfrak{p}(y) = \mathfrak{p}(z)$', where the target relations and operations of $<$, $+$ and $\times$ are the usual operations on the ordinals.

Theorem 19. $\mathfrak{b}\colon \text{PA} \to \text{ZF–inf}^*$.

Proof. This follows from Theorem 8 and Propositions 14 and 17. □

It remains to prove the following.

Theorem 20. The interpretations $\mathfrak{a}\colon \text{ZF–inf}^* \to \text{PA}$ and $\mathfrak{b}\colon \text{PA} \to \text{ZF–inf}^*$ are inverse to each other.

Proof. We must show that $\mathfrak{a}\mathfrak{b} = 1_{\text{PA}}$ and $\mathfrak{b}\mathfrak{a} = 1_{\text{ZF–inf}^*}$, i.e., that

$$\text{PA} \vdash \forall \bar{x}\,((\sigma(\bar{x})^{\mathfrak{b}})^{\mathfrak{a}} \leftrightarrow \sigma(\bar{x}))$$

and

$$\text{ZF–inf}^* \vdash \forall \bar{x}\,((\tau(\bar{x})^{\mathfrak{a}})^{\mathfrak{b}} \leftrightarrow \tau(\bar{x}))$$

for all atomic formulas $\sigma(\bar{x})$ of $\mathscr{L}_{\text{A}}$ and $\tau(\bar{x})$ of $\mathscr{L}_{\in}$. The details are routine; we make only a few comments on them here.

Note first that both interpretations preserve the logical symbols (they map domains to domains and do not alter equality), thus we need only look at non-logical symbols.

To see $\mathfrak{a}\mathfrak{b} = 1_{\text{PA}}$, work in PA and write $x <' y$ for $(x < y)^{(\mathfrak{a}\mathfrak{b})}$, $0'$ for the $<'$-least number, $1'$ for the $<'$-least number not equal to $0'$, $x +' y$ for the unique z such that $(x + y = z)^{(\mathfrak{a}\mathfrak{b})}$, and $x \cdot' y$ for the unique w such that $(x \cdot y = w)^{(\mathfrak{a}\mathfrak{b})}$. It can be checked that $0' = 0$, $1' = 1$, and $\forall x\, x + 1 = x +' 1'$, and by an induction on y this implies that $\forall x, y\, x + y = x +' y$. Since $x < y \leftrightarrow \exists z\, y = x + z + 1$ and $x <' y \leftrightarrow \exists z\, x = y +' z +' 1'$, it follows that $\forall x, y\, (x < y \leftrightarrow x <' y)$. Similarly $x \cdot (y + 1) = (x \cdot y) + x$ and $x \cdot' (y +' 1') = (x \cdot' y) +' x$, so by induction on y we can show that $\forall x, y\, x \cdot y = x \cdot' y$, as required.

To see $\mathfrak{b}\mathfrak{a} = 1_{\text{ZF–inf}^*}$, work in ZF–inf* and write $x \in' y$ for $(x \in y)^{(\mathfrak{b}\mathfrak{a})}$. Then by $\in$-induction we may show $\forall x \forall y\, (x \in y \leftrightarrow x \in' y)$ and the result then follows. □

7 Fragments of arithmetic and set theory

As with many results for models of Peano arithmetic, the results presented above have hierarchical variations for subsystems of PA and ZF. It seems worthwhile to briefly survey what one can say about this here. Our results along these lines are not really new since they are essentially contained in Ressayre's *Sous-systèmes remarquables de ZF* [19].

First, we fix a Δ_0 formula for $p = 2^x$ which defines a partial-function (formula (2) above), and for which the recursive definition (formula (1) above) and its 'downward' version,

$$\forall x \forall p\, (p = 2^{x+1} \to \exists q\, (p = 2q \wedge q = 2^x)) \tag{3}$$

are all provable in $\mathrm{I}\Delta_0$. This latter sentence appears necessary and more than just a convenience. But we cannot think of a natural example where the upward recursion holds but the downward one fails. However, the additional assumption is quite mild, for if a formula $\theta(x,y)$ for $y = 2^x$ satisfies (1) and (2) provably in $\mathrm{I}\Delta_0$, then $\theta(x,y) \wedge \forall u{<}x\, \exists v{<}y\, \theta(u,v)$ satisfies (1), (2) and (3) provably in $\mathrm{I}\Delta_0$, as the reader may check.

With such a definition of exponentiation chosen, the domain of definition of the exponentiation operation $y \mapsto 2^y$ is an initial segment of the model of $\mathrm{I}\Delta_0$, necessarily closed under successor and addition, but not in general closed under multiplication. We can prove in $\mathrm{I}\Delta_0$ that for each x there is a power of two, 2^y, that is greater than x, and in fact there is a least such power of two. With a slight misuse of notation we denote the least y such that $2^y > x$ by $\log x$.

It is fairly straightforward to see that $\mathrm{I}\Delta_0$ proves the Ackermann interpretation of many set theory axioms, including: extensionality; empty set; sum set; foundation; transitive closure; and the negation of the axiom of infinity.

Adding the further arithmetic axiom exp, $\forall x \exists y\, y = 2^x$, stating the totality of the function $x \mapsto 2^x$ we may prove the Ackermann interpretation of the pair set axiom and the power set axiom. Moreover, exp is necessary as well as sufficient for these: to show $\mathrm{I}\Delta_0 \vdash \mathrm{Pair}^{\mathfrak{a}} \to \mathrm{exp}$ and $\mathrm{I}\Delta_0 \vdash \mathrm{Pow}^{\mathfrak{a}} \to \mathrm{exp}$, take a number y and 2^x, the smallest power of two greater than y. Then as 2^x codes the set $\{x\}$, by pair set or power set the set $\{\{x\},\{\}\}$ must also be coded, and this code can only be $2^{2^x}+1$. Thus 2^{2^x} exists and so therefore does 2^y.[5]

To study the remaining axioms for set theory, we introduce a hierarchy of formulas related to Lévy's but more convenient for weak systems where the power set axiom may fail. Say a formula of the language of set theory is $\Delta_0^{\mathscr{P}}$ if it is formed from atomic formulas by boolean operations and bounded quantifiers of the form $\mathsf{Q}x \in y\ \ldots$ and $\mathsf{Q}x \subseteq y\ \ldots$. A formula is $\Sigma_n^{\mathscr{P}}$ (respectively $\Pi_n^{\mathscr{P}}$) if it is $\exists \bar{x}_1\, \forall \bar{x}_2 \ldots \mathsf{Q}\bar{x}_n\ \theta$ (respectively $\forall \bar{x}_1\, \exists \bar{x}_2 \ldots \mathsf{Q}^*\bar{x}_n\ \theta$) where θ is $\Delta_0^{\mathscr{P}}$. With a symbol for power set, the quantifier $\mathsf{Q}x \subseteq y\ \ldots$ can be replaced by $\mathsf{Q}x \in \mathscr{P}(y)\ \ldots$ and as the power set of a set x is Π_1-definable (in the sense of Lévy) this means that the $\Sigma_n^{\mathscr{P}}$ / $\Pi_n^{\mathscr{P}}$ hierarchy agrees with the Lévy hierarchy for $n \geqslant 2$ in the presence of the power set axiom.

With these definitions in place, we can readily prove by induction that $\mathrm{I}\Delta_0$ proves the Ackermann interpretation of the $\Delta_0^{\mathscr{P}}$-separation axiom scheme, $\Delta_0^{\mathscr{P}}$–Sep, and for $n \geqslant 1$, $\mathrm{I}\Sigma_n$ proves both $\Sigma_n^{\mathscr{P}}$–$\mathrm{Sep}^{\mathfrak{a}}$ and $\Pi_n^{\mathscr{P}}$–$\mathrm{Sep}^{\mathfrak{a}}$. Also, using the least number principle we have that $\mathrm{I}\Delta_0$ proves $\Delta_0^{\mathscr{P}}$–$\mathrm{Ind}^{\mathfrak{a}}$, the Ackermann interpretation of the $\Delta_0^{\mathscr{P}}$-epsilon

[5] **Further references.** As mentioned by Enayat [personal communication], two theories of sets were proved to be equivalent to $\mathrm{I}\Delta_0 + \mathrm{exp}$ by Gaifman and Dimitracopoulos [9], one of which is the *theory of exponentially finite sets*.

induction axiom scheme

$$\forall \bar{a}\, (\forall y\, (\forall x \in y\, \theta(x,\bar{a}) \to \theta(y,\bar{a})) \to \forall y\, \theta(y,\bar{a}))$$

for θ in $\Delta_0^{\mathscr{P}}$, and similarly $\mathrm{I}\Sigma_n$ proves $\Sigma_n^{\mathscr{P}}$–Ind$^{\mathfrak{a}}$ and $\Pi_n^{\mathscr{P}}$–Ind$^{\mathfrak{a}}$.

The collection and replacement axiom schemes are a little more delicate. Let $\Sigma_n^{\mathscr{P}}$–Coll be the scheme

$$\forall a, \bar{c}\, (\forall x \in a\, \exists y\, \theta(x,y,\bar{c}) \to \exists b\, \forall x \in a\, \exists y \in b\, \theta(x,y,\bar{c}))$$

for θ in $\Sigma_n^{\mathscr{P}}$, etc. To justify the Ackermann interpretation of axioms like this, we need fast growing functions. For example, taking $a \supseteq \{l-1\}$ where $l = \log a$ and letting $\theta(x,y)$ be the Δ_0 formula $y = \{x\}$ we trivially have $(\forall x \in a\, \exists y\, \theta(x,y))^{\mathfrak{a}}$ in all models of $\mathrm{I}\Delta_0$, but the code for any b with $(\exists b\, \forall x \in a\, \exists y \in b\, \theta(x,y))^{\mathfrak{a}}$ must be at least as large as $2^{2^{l-1}} \approx 2^a$ so exponentiation is required. On the other hand, it is straightforward to see that $(\Sigma_n^{\mathscr{P}}\text{–Coll})^{\mathfrak{a}}$ is provable in $\mathrm{I}\Delta_0 + \exp + \mathrm{B}\Sigma_n$. Similar remarks apply to the replacement axioms, which we might as well take here to be

$$\forall a, \bar{c}\, (\forall x \in a\, \exists y\, \theta(x,y,\bar{c}) \to \exists f\, \forall x \in a\, \theta(x, f(x), \bar{c})).$$

Then $(\Sigma_n^{\mathscr{P}}\text{–Rep})^{\mathfrak{a}}$ is provable in $\mathrm{I}\Sigma_n$, for each $n \geqslant 1$, but for $n = 0$ the best result seems to be $\mathrm{I}\Delta_0 + \exp + \mathrm{B}\Sigma_1 \vdash (\Delta_0^{\mathscr{P}}\text{–Rep})^{\mathfrak{a}}$.

The inverse $\mathfrak{b}$ to the Ackermann interpretation is a somewhat more complicated affair, and we make no effort here to even to attempt to say what happens here in the absence of the axioms corresponding to exponentiation, namely pair set and power set. We shall simply state our results, which are based on a straightforward but rather tedious analysis of the proofs given in Sections 4, 5 and 6, and leave the verification to the reader. (The presentation of Kripke–Platek set theory given in Chapter I of Barwise [3] may be found helpful.) It turns out that for the results that follow, the usual Lévy hierarchy (where bounded quantifiers take the form $\mathsf{Q}x \in y\ \ldots$ only) gives the sharpest result. We denote the levels in this hierarchy by Δ_n^L, Σ_n^L and Π_n^L.

The first stage is to define the ordinals, in particular the arithmetic structure on the ordinals, and to define the bijection $\mathfrak{p}\colon V \to \mathrm{On}$ between the universe and the ordinals. We take as our base theory, the set theory consisting of the axioms of extensionality, empty set, sum set, and the negation of the axiom of infinity, and the axiom schemes of: Δ_0^L-separation, Δ_0^L-collection, and both Σ_1^L-epsilon induction and Π_1^L-epsilon induction. With this base theory, the class of ordinals is defined (as the transitive sets linearly ordered by $\in$) by a Δ_0^L formula, and the arithmetic operations on the ordinals including $+$, $\times$, exp are all functions with Δ_1^L graph. Moreover the scheme of Σ_1^L-induction proves the axiom of foundation and the axiom of transitive containment, and hence we may define the bijection $\mathfrak{p}\colon V \to \mathrm{On}$, also with Δ_1^L graph, and that this map is inverse to the Ackermann interpretation of sets from the ordinals. The base axioms for arithmetic, i.e., the axioms for nonnegative parts of discretely ordered rings, can then be shown to hold in the ordinals.

To prove the induction scheme we need somewhat stronger axioms showing that the ordinals satisfy $\mathrm{I}\Sigma_n$, etc. Perhaps the simplest option is to take stronger instances of the separation scheme. Given a Σ_n-definable subclass A of the ordinals containing 0 and closed under successor, and given $n \in \mathrm{On}$, let a be the set of ordinals in A less than n, by Σ_n^L-separation. Then by ordinal induction $a = n$ so, as n was arbitrary, $A = \mathrm{On}$.

Putting all this together, we have

Theorem 21. Let $n \geqslant 1$ and let Σ_n^L–Sep denote the set theory consisting of Σ_n^L-separation together with base theory Ext, Emp, Sum, $\neg$Inf, Δ_0^L–Coll, Σ_1^L–Ind, and Π_1^L–Ind. Then the Ackermann interpretation $\mathfrak{a}$ and its inverse $\mathfrak{b}$ are defined for $\mathrm{I}\Sigma_n$ and Σ_n^L–Sep, with

$$\mathfrak{a}\colon \Sigma_n^L\text{–Sep} \to \mathrm{I}\Sigma_n$$

and

$$\mathfrak{b}\colon \mathrm{I}\Sigma_n \to \Sigma_n^L\text{–Sep}$$

and these are inverse to each other. □

Thus, the Ackermann interpretation behaves well for the 'traditional' subtheories $\mathrm{I}\Sigma_n$ of PA. For weaker theories the pair set and power set axioms behave like the axiom exp expressing the totality of exponentiation.

To close this section, we should remark that for systems of arithmetic without exp, there are other choices of interpretations of set theory. The Ackermann interpretation is of course valid, and in some sense represents the smallest collection of sets or smallest model of set theory one might take, but we could transform a model M of $\mathrm{I}\Delta_0$ to a different model of set theory by taking as sets all *bounded* Δ_0*-definable subsets of* M, and identifying an $a \in M$ with the set of $x \in M$ such that $M \vDash (x \in a)^{\mathfrak{a}}$. In the (apparent) absence of a formalised notion of truth for Δ_0 formulas in $\mathrm{I}\Delta_0$ (this itself an open problem) this does not appear to be an interpretation in the formal sense used above. It is essentially the approach taken by Diaconescu and Kirby [6] and, as they point out, counting principles and pigeonhole principles become highly relevant. (It is still not known if $\mathrm{I}\Delta_0$ alone can prove the pigeonhole principle for Δ_0 functions.) Of course if $M \vDash \exp$ this interpretation and the Ackermann one coincide since

$$\mathrm{I}\Delta_0 + \exp \vdash \forall a\, \exists y\, \forall x{<}a\, ((x \in y)^{\mathfrak{a}} \leftrightarrow \theta(x))$$

for each Δ_0 formula $\theta(x)$, possibly with further parameters.

8 Conclusions for finite set theories and arithmetic

Perhaps the first and most obvious conclusion is that statements concerning the equivalence of 'Peano Arithmetic' and 'ZF with the axiom of infinity negated' require some care to formulate and prove. It is certainly true that PA and 'ZF with the axiom of infinity negated' are equiconsistent for just about any sensible axiomatisation of the latter, in the sense that interpretations exist in both directions.[6] Probably this is the 'folklore result' that most people remember. But for the finer result with interpretations inverse to each other, careful axiomatisation of the set theory is required. A category theoretic framework for interpretations is useful to direct attention to these refinements.

The reader will have noticed that the axiom of choice (AC) has been completely absent from the discussion here. It wasn't necessary at all for the inverse Ackermann interpretation, though it comes for free with the Ackermann interpretation itself: $\mathfrak{a}\colon \mathrm{ZF}\text{–inf}^* + \mathrm{AC} \to \mathrm{PA}$. This shows $\mathrm{ZF}\text{–inf}^* \vdash \mathrm{AC}$. In fact AC can be proved in much less than this, as Baratella and Ferro [2] point out.

The universe of hereditarily finite sets has been an important source of analogies, inspiration and axioms for set theory with infinite sets, from the time of Cantor onwards. In the case of the work here we have learnt that models of ZF–inf* are categorical with respect to their ordinals, in the sense that the model of set theory can

[6] **Further references.** It is mentioned in Kirby's article [13, §5] that PS (which was proved to be "equivalent" to PA by Previale [17] as previously mentioned) is interpretable in $\mathrm{ZF}\backslash\{\mathrm{Inf}\}$ (attributed to Mycielski).

be reconstructed from the ordinals as a model of arithmetic. The analogous case for models of full ZF is a result due to Harvey Friedman [8] that characterises $\mathrm{ZF}+V{=}L$ as the first-order theory extending ZF in which any two models with the same ordinals are isomorphic. (See Friedman's paper for a precise statement of this result.)[7]

Results like this provide evidence that models of ZF in which $V{=}L$ holds are actually the 'uninteresting' models of ZF, and a similar remark holds for models of finite set theory too. Arguably, the uninteresting models of ZF–inf are the ones in which the axiom of transitive containment holds, and the interesting ones are those for which TC fails. It would be nice to know more about such models, but very few have appeared in the literature. Note that every model M of ZF–inf has a transitive submodel $V \subseteq M$, $V \vDash$ ZF–inf + TC, with the same ordinals, defined to be the image of the function $f\colon \mathrm{On} \to V$ given by the recursion on ordinals

$$f(x) = \{f(y) : \mathrm{On} \vDash (y \in x)^{\alpha}\}.$$

One way to think about the submodel $V \subseteq M$ is that this is the collection of sets in M with a rank, i.e., they are the sets in the cumulative hierarchy. (Hence the name V we have used for this submodel.)

Whether the theory of models with $M \neq V$ will have any consequence for arithmetic itself is not clear, as the ordinals in models of ZF–inf $+\neg$TC are of course just the usual models of PA.

Finite set theories in which the axiom of power set fails are an altogether more difficult proposition—apparently for some very good reasons possibly related to the coding problems and complexity theory issues one gets with models of arithmetic without exponentiation.

It is not quite clear to us exactly how this might arise from the point of view of the Ackermann interpretation considered here. However, one view of the inverse Ackermann interpretation in a model M of some fragment of finite set theory is that it sets out to define suitable arithmetic structure for $+$ and $\cdot$ on the whole universe of sets, in a way that is compatible with the Ackermann interpretation that maps ordinals to sets in the cumulative hierarchy $V \subseteq M$. (V will in general be a subclass of the class of all sets because the transitive containment axiom may fail in M.) But it is clear that if one did have a general process for achieving this, or even for defining $+$ and $\cdot$ on a significant part of the whole universe, this arithmetic model would have to be an end-extension of the image of the ordinals given by the Ackermann interpretation.[8]

Such a general method for producing end-extensions of models of arithmetic might be very significant for models of $\mathrm{I}\Delta_0 + \neg\exp$. There are several open questions concerning when models of $\mathrm{I}\Delta_0 + \neg\exp$ have end-extensions [28]. More significantly, success in the naïve attempt at producing an inverse Ackermann interpretation would give end-extensions $M \subseteq_{\mathrm{e}} K$ of $M \vDash \mathrm{I}\Delta_0 + \neg\exp$ where each element $x \in M$ has an exponential $y = 2^x \in K$. Such a construction might be applied in the case when the model of finite set theory is of the type studied by Diaconescu and Kirby [6], that is formed from Δ_0-definable bounded subsets. This would in turn have consequences for the Π_1-theory of the base model M. In particular this would have consequences for the Π_1-theory provable in $\mathrm{I}\Delta_0$ with a single application of exponentiation. This Π_1-theory contains some statements not known in $\mathrm{I}\Delta_0$ itself, such as the Δ_0 pigeonhole principle,

[7] **Further references.** The first-order theory of ordinals in models of ZF was studied by Takeuti [21, 22, 23, 24] in a language extending $\mathscr{L}_{\mathrm{A}}$.

[8] **Further references.** Kirby's article [13] gives a definition of addition and multiplication on all sets in a model of ZF$\setminus\{$Inf$\}$, which is a apparently not isomorphic to the Ackermann one.

and also contains consistency statements such as the tableau consistency of Q, known not to be provable in $I\Delta_0$ alone. (See for example papers by Pudlák [18], and Paris and Wilkie [27] for much more information.)

Finally, leaving arithmetic aside, the Axiom of Choice becomes interesting for systems of finite set theory that do not have the power set axiom. Apparently power set is required for the equivalence of the well-ordering principle (WO) and AC, though WO $\to$ AC can be proved by the usual proof without power set. However, we were unable to find a model of $\mathrm{AC} + \neg\mathrm{WO}$.[9] Using the notation of Baratella and Ferro introduced above, we can ask

Question 22. Does $\mathrm{EST} + \neg\mathrm{Inf} + \mathrm{Found} \vdash \mathrm{AC}$?

Note that, as Baratella and Ferro point out, $\mathrm{EST} + \neg\mathrm{Inf} + \mathrm{Pow} \vdash \mathrm{WO}$ by a straightforward argument mapping sets bijectively onto ordinals. They also provide a model of $\mathrm{EST} + \neg\mathrm{Inf} + \neg\mathrm{WO}$.

References

[1] ACKERMANN, WILHELM (1937). Die Widerspruchsfreiheit der allgemeinen Mengenlehre. *Mathematische Annalen*, **114**:305–315.

[2] BARATELLA, STEFANO and FERRO, RUGGERO (1993). A theory of sets with the negation of the axiom of infinity. *Mathematical Logic Quarterly*, **39**(3):338–352.

[3] BARWISE, JON (1975). *Admissible Sets and Structures: An Approach to Definability Theory*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin.

[4] BETH, E. W. (1964). Axiomatique de la théorie des ensembles sans axiome de l'infini. *Bulletin de la Société Mathématique de Belgique*, **16**:127–136. In French. Inaccessible.

[5] CHANG, CHEN CHUNG and KEISLER, H. JEROME (1977). *Model Theory*, Studies in Logic and the Foundations of Mathematics, vol. 73. North-Holland Publishing Company, Amsterdam, second edition.

[6] DIACONESCU, RADU and KIRBY, LAURENCE A. S. (1987). Models of arithmetic and categories with finiteness conditions. *Annals of Pure and Applied Logic*, **35**(2):123–148.

[7] DRAKE, FRANK R. (1974). *Set Theory: An Introduction to Large Cardinals*, Studies in Logic and the Foundations of Mathematics, vol. 76. North-Holland Publishing Company, Amsterdam.

[8] FRIEDMAN, HARVEY (1978). Categoricity with respect to ordinals. In *Higher Set Theory: Proceedings, Oberwolfach, Germany, April 13-23, 1977* (MÜLLER, GERT H. and SCOTT, DANA S., eds.), Lecture Notes in Mathematics, vol. 669, pp. 17–20. Springer-Verlag, Berlin.

[9] **Further references.** Zarach [29, §4] proved that if $\mathrm{ZF}\setminus\{\mathrm{Pow}\}$ is consistent, then so is $(\mathrm{ZF}\setminus\{\mathrm{Pow}\}) + \mathrm{AC} + \neg\mathrm{WO}$. He attributed this to Zbigniew Szczepaniak.

[9] GAIFMAN, HAIM and DIMITRACOPOULOS, CONSTANTINE (1982). Fragments of Peano's arithmetic and the MRDP theorem. In *Logic and Algorithmic: Papers from an International Symposium in honor of Ernst Specker held at Eidgenössische Technische Hochschule, Zurich, February 5-11, 1980*, no. 30 in Monographies de L'Enseignement Mathématique, pp. 187–206. Université de Genève, Geneva.

[10] GIVANT, S. and TARSKI, A. (1977). Peano arithmetic and the Zermelo-like theory of sets with finite ranks. *Notices of the American Mathematical Society*, **77T–E51**:A–437. Inaccessible.

[11] HÁJEK, PETR and PUDLÁK, PAVEL (1993). *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin.

[12] KAYE, RICHARD (1991). *Models of Peano Arithmetic*, Oxford Logic Guides, vol. 15. Clarendon Press, Oxford.

[13] KIRBY, LAURENCE (2007). Addition and multiplication of sets. *Mathematical Logic Quarterly*, **53**(1):52–65.

[14] KUNEN, KENNETH (1980). *Set Theory: An Introduction to Independence Proofs*, Studies in Logic and the Foundations of Mathematics, vol. 102. North-Holland Publishing Company, Amsterdam.

[15] MANCINI, ANTONELLA and ZAMBELLA, DOMENICO (2001). A note on recursive models of set theories. *Notre Dame Journal of Formal Logic*, **42**(2):109–115.

[16] MYCIELSKI, JAN (1964). The definition of arithmetic operations in the Ackermann model. *Algebra i Logika Seminar*, **3**:64–65. In Russian. Inaccessible.

[17] PREVIALE, FLAVIO (1994). Induction and foundation in the theory of the hereditarily finite sets. *Archive for Mathematical Logic*, **33**:213–241. Inaccessible.

[18] PUDLÁK, PAVEL (1985). Cuts, consistency statements and interpretations. *The Journal of Symbolic Logic*, **50**(2):423–441.

[19] RESSAYRE, JEAN-PIERRE (1987). Modèles non standard et sous-systèmes remarquables de ZF. In *Modèles non Standard en Arithmétique et théorie des Ensembles*, Publications Mathématiques de l'Université Paris VII, vol. 22, pp. 47–147. Université de Paris VII U.E.R. de Mathématiques, Paris.

[20] ŚWIERCZKOWSKI, S. (2003). Finite sets and Gödel's incompleteness theorems. *Dissertationes Mathematicae (Rozprawy Matematyczne)*, **422**:1–58.

[21] TAKEUTI, GAISI (1954). Construction of the Set Theory from the Theory of Ordinal Numbers. *Journal of the Mathematical Society of Japan*, **6**(2):196–220.

[22] TAKEUTI, GAISI (1957). On the Theory of Ordinal Numbers. *Journal of the Mathematical Society of Japan*, **9**(1):93–113.

[23] TAKEUTI, GAISI (1958). On the Theory of Ordinal Numbers, II. *Journal of the Mathematical Society of Japan*, **10**(1):106–120.

[24] TAKEUTI, GAISI (1965). A Formalization of the Theory of Ordinal Numbers. *The Journal of Symbolic Logic*, **30**(3):295–317.

[25] TARSKI, A. and GIVANT, S. (1987). *A Formalization of Set Theory Without Variables*, American Mathematical Society Colloquium Publications, vol. 41. American Mathematical Society, Providence. Inaccessible.

[26] VISSER, ALBERT (2004). Categories of theories and interpretations. Logic Group Preprint Series 228, Faculteit Wijsbegeerte, Universiteit Utrecht, `http://www.phil.uu.nl/preprints/preprints/PREPRINTS/preprint228.pdf`. Published in *Logic in Tehran: Proceedings of the Workshop and Conference on Logic, Algebra, and Arithmetic, held October 18-22, 2003* (ENAYAT, ALI, KALANTARI, IRAJ and MONIRI, MOJTABA, eds.), Lecture Notes in Logic, vol. 26, pp. 284–341. Association for Symbolic Logic, La Jolla, CA, 2006.

[27] WILKIE, ALEX J. and PARIS, JEFF B. (1987). On the scheme of induction for bounded arithmetic formulas. *Annals of Pure and Applied Logic*, **35**(3):261–302.

[28] WILKIE, ALEX J. and PARIS, JEFF B. (1989). On the existence of end extensions of models of bounded induction. In *Logic, Methodology and Philosophy of Science, VIII (Moscow, 1987)* (FENSTAD, JENS ERIK, FROLOV, IVAN T. and HILPINEN, RISTO, eds.), Studies in Logic and the Foundations of Mathematics, vol. 126, pp. 143–161. North-Holland Publishing Company, Amsterdam.

[29] ZARACH, ANDRZEJ (1982). Unions of ZF^--models which are themselves ZF^--models. In *Logic Colloquium '80* (VAN DALEN, D., LASCAR, D. and SMILEY, J., eds.), Studies in Logic and the Foundations of Mathematics, pp. 315–342. North-Holland Publishing Company, Amsterdam.